NIHR | National Institute for Health and Care Research

Journals Library

Health and Social Care Delivery Research

Volume 13 • Issue 25 • July 2025

ISSN 2755-0079

Consequences, costs and cost-effectiveness of workforce configurations in English acute hospitals

Peter Griffiths, Christina Saville, Jane Ball, David Culliford, Jeremy Jones, Francesca Lambert, Paul Meredith, Bruna Rubbo, Lesley Turner, and Chiara Dall'Ora

DOI 10.3310/ZBAR9152

Health and Social Care Delivery Research

Extended Research Article

Consequences, costs and cost-effectiveness of workforce configurations in English acute hospitals

Peter Griffiths,[1,2,3*] Christina Saville,[1,3] Jane Ball,[1] David Culliford,[1,3]
Jeremy Jones,[1] Francesca Lambert,[1] Paul Meredith,[1,2,3]
Bruna Rubbo,[1] Lesley Turner[1] and Chiara Dall'Ora[1,3]

[1]Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK
[2]Clinical Outcomes Research Group, Portsmouth Hospitals University Trust, NHS, Portsmouth, UK
[3]National Institute for Health and Care Research Applied Research Collaboration (Wessex), University of Southampton, Southampton, UK

[*]**Corresponding author** peter.griffiths@soton.ac.uk

Published July 2025
DOI: 10.3310/ZBAR9152

This report should be referenced as follows:

Griffiths P, Saville C, Ball J, Culliford D, Jones J, Lambert F, *et al.* Consequences, costs and cost-effectiveness of workforce configurations in English acute hospitals. *Health Soc Care Deliv Res* 2025;**13**(25). https://doi.org/10.3310/ZBAR9152

Health and Social Care Delivery Research

ISSN 2755-0079 (Online)

A list of Journals Library editors can be found on the NIHR Journals Library website

Health and Social Care Delivery Research (HSDR) was launched in 2013 and is indexed by Europe PMC, DOAJ, INAHTA, Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA), NCBI Bookshelf, Scopus and MEDLINE.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

This journal was previously published as *Health Services and Delivery Research* (Volumes 1–9); ISSN 2050-4349 (print), ISSN 2050-4357 (online)

The full HSDR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr.

Criteria for inclusion in the *Health and Social Care Delivery Research* journal

Manuscripts are published in *Health and Social Care Delivery Research* (HSDR) if (1) they have resulted from work for the HSDR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HSDR programme

The HSDR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HSDR programme please visit the website at www.nihr.ac.uk/explore-nihr/funding-programmes/health-and-social-care-delivery-research.htm

This article

The research reported in this issue of the journal was funded by the HSDR programme or one of its preceding programmes as award number NIHR128056. The contractual start date was in March 2020. The draft manuscript began editorial review in December 2024 and was accepted for publication in January 2025. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HSDR editors and production house have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Background: The National Health Service faces significant challenges in recruiting and retaining registered nurses. Recruiting unregistered staff is often adopted as a solution to the registered nurse shortage, but recent research found lower registered nurse staffing levels increase hospital mortality with no evidence that higher levels of assistant staff reduced risk.

Objectives: To estimate the consequences, costs and cost-effectiveness of variation in the size and composition of the staff on acute hospital wards in England. To determine if results are likely to be sensitive to staff groups such as doctors and therapists, who are not on ward rosters, associations between staffing and outcomes for multiple staff groups, including medical, are explored at hospital level.

Design: A national cross-sectional panel study and a patient-level longitudinal observational study using routine data.

Setting: All English acute hospital Trusts and a subsample of four Trusts for the patient-level study.

Interventions: Naturally occurring variation in the size and composition of the workforce.

Participants: Patients experiencing a hospital admission with an overnight stay and nursing staff providing care on inpatient wards.

Outcomes: Death, patient and staff experience, length of stay, re-admission, adverse events, incidents (Datix), staff sickness, costs and quality-adjusted life-years.

Data sources: Publicly available records of hospital activity, staffing and outcomes (cross-sectional study) and hospital administrative systems (longitudinal study).

Results: In the cross-sectional study, lower staffing levels from doctors and allied health professionals were associated with increased risk of death. Higher nurse staffing levels were associated with better patient experience and staff well-being. In the longitudinal study, for adult inpatients, exposure to days with lower-than-expected registered nurses or nursing assistant staff was associated with increased hazard of death (adjusted hazard ratio 1.08/1.07, 95% confidence interval 1.07 to 1.09/1.06 to 1.08) and longer hospital stays. Low registered nurse staffing was also associated with increased hazard of re-admission (adjusted hazard ratio 1.01, 95% confidence interval 1.01 to 1.02). Eliminating low staffing cost £2778 per quality-adjusted life-years gained. Avoidance of registered nurse understaffing gave more benefits and was more cost-effective for highly acute patients. Although high bank or agency staffing was associated with increased hazard of death, avoiding low staffing using temporary staff still reduced mortality but was more costly and less effective than using permanent staff. If costs of avoided hospital stays are included, avoiding low staffing generates a net cost saving. Exploration of thresholds for low staffing indicated a greater beneficial effect from registered nurse staffing higher than current norms.

Limitations: This is an observational study. Causal inferences cannot be made from these results in isolation. Quality-adjusted life-years gains were estimated, although conclusions are not sensitive to assumptions or discount rates. We used current ward norms as reference for low staffing.

Conclusions: Our results show the adverse effects of low nurse staffing but also show that medical and allied health professional staffing are important considerations for patient safety. Eliminating low registered nurse staffing gave more benefits than eliminating assistant staffing.

Future work: Research is needed to validate methods to determine nurse staffing requirements, and the interaction between registered nurse and assistant staffing needs further exploration.

Study registration: This study is registered as Current Controlled Trials ClinicalTrials.gov NCT04374812.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme (NIHR award ref: NIHR128056) and is published in full in *Health and Social Care Delivery Research*; Vol. 13, No. 25. See the NIHR Funding and Awards website for further award information.

Contents

List of tables **vi**

List of figures **viii**

List of supplementary material **ix**

List of abbreviations **x**

Plain language summary **xi**

Scientific summary **xii**

Chapter 1 Introduction **1**
Temporary and flexible staffing 1
Skill-mix 1
Cost-effectiveness 2
Causal inferences and estimation of staffing effects 3
Building from previous research 3
Aims and objectives 4

Chapter 2 Literature review **5**
Background and overview of relevant literature 5
Review methods 6
Economic evidence 7
  Volume and quality of economic evidence on nurse staffing 7
  Costs and cost-effectiveness 8
Longitudinal studies 11
Multidisciplinary staffing 13
Review summary and conclusions 14

Chapter 3 Methods **16**
Cross-sectional study (national, hospital Trust level) 16
  Trust and staffing data 16
  Outcomes 17
  Analysis 18
Cross-sectional study (national, ward level) 18
  Data cleaning and linkage 19
  Analysis 19
Longitudinal study (selected Trusts, patient level) 19
  Sampling and data sources 21
  Analysis 23
Economic analysis 26
Ethics and approvals 27
Patient, public and service engagement/involvement 27

Chapter 4 National cross-sectional study results **29**
Hospital mortality 29
Patient experience 29
Staff experience 31

Ward-level associations between nurse staffing and harms (national data) 31
Summary of main results 31

Chapter 5 Trust data results **37**
Adult inpatient ward admissions 37
Wards and staffing 37
Effects of days of low hours per patient day 39
Sensitivity analyses 39
Patient and staffing subgroups 43
Effects of staff-mix 44
Low staffing thresholds 44
Interactions and non-linear effects 44
Patient transfers 46
Adverse events (patient) 46
Incidents (Datix) 47
Paediatrics 50
Staff sickness absence 50
  Effects of exposure to registered nurse and nursing assistant understaffing 50
  Staff mix and working patterns 50
Cost consequences and cost-effectiveness 53
Summary of main results 57

Chapter 6 Discussion and conclusions **61**
Summary of main findings 61
Relationship to existing research 61
Strengths and limitations relative to other research 63
Challenges faced and other considerations 64
Patient, public and service engagement and involvement 64
  Public engagement 65
  Staff and service engagement 65
Equality, diversity and inclusion 66
Impact and learning 66
Implications for decision-makers 67
Research recommendations 68
Conclusions 68

Additional information **69**

References **74**

Appendix 1 Additional material for literature review **84**

Appendix 2 Additional methods **92**

Appendix 3 Additional tables of results **100**

List of tables

TABLE 1 Summary of results of economic studies of nurse staffing 9

TABLE 2 Costs per life saved from studies of increased staffing 12

TABLE 3 Contextual data for participating Trusts (2019) 20

TABLE 4 Occupied bed-per-staff across 138 hospital Trusts (2016–9) 30

TABLE 5 Associations between beds per FTE staff and mortality from the negative binomial random-effect models, adjusted for hospital Trust characteristics 30

TABLE 6 Associations between beds per FTE staff and patient experience, from random-effect models, adjusted for hospital Trust characteristics 32

TABLE 7 Associations between beds per FTE staff and staff experience from the random-effect mixed models, adjusted for hospital characteristics 32

TABLE 8 Associations between beds per FTE staff and staff-reported health and well-being theme score, within between random-effects model (adjusted for hospital characteristics) 34

TABLE 9 Association between ward-level HPPD and harms recorded in the safety thermometer 36

TABLE 10 Adult inpatient admissions 38

TABLE 11 Patient day staffing characteristics for all days and days of low staffing 40

TABLE 12 Exposures and outcomes for those exposed vs not exposed to low staffing 41

TABLE 13 Association between days of understaffing and outcomes, adjusted for risk 42

TABLE 14 Staff mix and mortality 45

TABLE 15 Association between days of understaffing, patient transfers and mortality, adjusted for risk 48

TABLE 16 Frequency of nurse-sensitive adverse events and associations with staffing 48

TABLE 17 Association between reported incident rates and staffing levels for different levels of harm 49

TABLE 18 Exposures and outcomes for those exposed vs not exposed to low staffing (paediatrics) 51

TABLE 19 Association between days of understaffing and outcomes, adjusted for risk (paediatrics) 51

TABLE 20 Staffing configurations in the previous week by sickness cohort 52

TABLE 21 Association between exposure to understaffing and sickness absence 52

TABLE 22 Association between understaffing and sickness absence, adjusted for skill-mix, temporary staffing, shift work 54

TABLE 23 Association between understaffing and sickness absence, adjusted for statistically significant variables added to the core models **55**

TABLE 24 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing (base-case and sensitivity analysis) **56**

TABLE 25 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing for patient/ward subgroups **58**

TABLE 26 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing with agency staff (base-case and sensitivity analysis) **59**

TABLE 27 Search strategies **84**

TABLE 28 Characteristics of economic studies **87**

TABLE 29 Risk of bias in economic studies **89**

TABLE 30 Assessment of economic study quality **90**

TABLE 31 Items contributing to the themes in staff survey benchmark data set **93**

TABLE 32 Candidate nurse-sensitive outcomes based on Blume *et al.* 2021 **95**

TABLE 33 Unit costs to be used in the analysis **98**

TABLE 34 Abridged life table life expectancy estimates, discounted life expectancy, quality-of-life weights and DANQALE **99**

TABLE 35 Sensitivity analysis – different ‘exposure windows’ for low staffing **100**

TABLE 36 Effects of staff mix on mortality – survival models **101**

TABLE 37 Association between net hours of understaffing and outcomes, adjusted for risk **102**

TABLE 38 Nonlinear and interaction effects in mortality – survival models **103**

TABLE 39 Association between days of understaffing and outcomes for ward and patient subgroups **105**

TABLE 40 Recorded harms and levels of harm **107**

List of figures

FIGURE 1 Hierarchical matrix to summarise findings and economic conclusions from economic studies of nurse staffing/skill-mix increase **11**

FIGURE 2 Scatterplot of planned vs. actual CHPPD for acute hospital wards **35**

FIGURE 3 Summary of subgroup analyses **43**

FIGURE 4 Hazard ratio for days of low staffing for different thresholds **46**

FIGURE 5 Nonlinear interactions: low RN staffing conditional on NA staffing **46**

FIGURE 6 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram of sources for the economic review **86**

List of supplementary material

Report Supplementary Material 1	Engagement and involvement report

Supplementary material can be found on the NIHR Journals Library report page (https://doi.org/10.3310/ZBAR9152).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

AHP	allied healthcare professional
AIC	Akaike information criterion
BIC	Bayesian information criterion
CHPPD	care hours per patient day
CMI	Charlson Comorbidity Index
DANQALE	discounted and quality-adjusted life expectancy
DVT	deep vein thrombosis
FTE	full-time equivalent
GDP	gross domestic product
GVIF	Generalised Variance Inflation Factor
HES	Hospital Episode Statistics
HPPD	hours per patient day
HRG	Health Care Resource Group
ICD-10	*International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision
ICU	intensive care unit
LCL	lower confidence level
LOS	length of stay
NA	nursing assistant
NEWS	National Early Warning Score
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
PAS	Patient Administration System
PH	proportional hazards
PPIE	patient and public involvement and engagement
PU	pressure ulcer
QALYs	quality-adjusted life-years
RN	registered nurse
ROBINS–I	Risk Of Bias In Non-randomized Studies – of Interventions
SHMI	Summary Hospital Mortality Index
ST&T	scientific, therapeutic and technical
UCL	upper confidence level
WBRE	Within-Between Random Effects model
WISeRD	Wessex Inclusion in Service Research and Design

Note on currency

US$1 = £0.77 (date of conversion: 13 March 2025)

CN¥ = £0.11 (date of conversion: 13 March 2025)

AU$1 = £0.49 (date of conversion: 13 March 2025)

€ = £0.84 (date of conversion: 13 March 2025)

Plain language summary

Staffing shortages are a major concern for the National Health Service. A lot of research shows that low nurse staffing in hospital is correlated with worse patient outcomes, including an increased risk of death. However, a lot of this research has only looked at hospital average staffing and has not considered other staff, such as doctors and allied health professionals, so it is hard to be sure if improving nurse staffing on wards leads to better outcomes. It is also hard to know the most cost-effective approach to addressing staff shortages. Our study used existing data from national reports and daily staffing data from hospital wards to answer some of the main uncertainties from past research. Using data from national reports, we found low staffing levels from doctors and allied health professionals were linked to increased risk of death. Nurse staffing levels were linked to important aspects of patient experience and staff well-being, but to properly understand the effects of nurse staffing we needed to know the staffing patients experience when on hospital wards. Our study included 626,313 patients in 4 hospitals. We found that when patients spent time on wards with fewer-than-expected registered nurses or nursing assistants, they were more likely to die and their stay in hospital was longer. Low registered nurse staffing was also associated with more re-admissions. We looked at the cost of avoiding low staffing and the cost of gaining the equivalent of 1 year of healthy life. We compared these 'cost-effectiveness' estimates for different ways of avoiding low staffing and for different patient groups. Overall, we concluded that a focus on avoiding low registered nurse staffing gave more benefits than using assistants to fill any gaps, and should be the priority, although it is still not clear what the best level of staff is.

Scientific summary

Background

The consequences of staff shortages in the NHS are potentially serious. Several inquiries, NHS guidance and an extensive body of research indicate that lower registered nurse (RN) staffing levels are associated with adverse patient and staff outcomes. Patient outcomes associated with lower nurse staffing include increased risk of death, hospital-acquired infections, falls, poor patient experience and nursing care omissions. Adverse nurse outcomes include burnout, job dissatisfaction and intention to leave.

Maintaining adequate staffing for hospital wards is challenging. As of March 2022, approximately 10% of nursing posts in acute settings were vacant, totalling 38,972 vacancies. In the aftermath of the Francis inquiry and the publication of guidance for safe staffing in adult wards in acute hospitals by the National Institute for Health and Care Excellence in 2014, the number of RNs employed in acute hospitals increased. However, this followed a period where absolute numbers fell, and the increases have not matched activity growth. Although there appeared to be an immediate uplift in the number of applicants for nursing courses following the start of the COVID-19 pandemic in 2020, this has not been sustained. Steps taken to increase supply will not resolve shortages for some time, assuming they are successful. In the face of such scarcity and the need to manage expenditure on staff, care providers and policy-makers face difficult decisions as they plan how to provide adequate nurse staffing levels.

Despite extensive evidence demonstrating associations between low nurse staffing levels and adverse outcomes, important uncertainties remain. Most evidence is from cross-sectional studies which have not considered staffing by other staff groups. This means that estimates of nurse staffing effects could be biased, and the importance of the multidisciplinary team may not be recognised. Although there is a growing body of longitudinal studies which avoid many of the limitations of cross-sectional studies, these are still limited. Most economic studies rely on studies with a high risk of bias to estimate the effects of changes in staffing configurations.

Aims and objectives

This study aims to provide evidence to inform cost-effective deployment of nursing and other care staff on hospital wards in England and policy decisions to address nursing workforce shortages. We address the question of which combinations of care staff employed by hospitals and deployed on hospital wards provide the most cost-effective care in terms of patient safety, experience and efficient use of resources by undertaking two studies. First, we explore hospital-level cross-sectional associations between staffing and outcomes and, second, we assess how outcomes and costs vary as patients and nurses are exposed to low staffing on hospital wards longitudinally.

Methods

We undertook a national cross-sectional panel study using routine data from all English acute hospital Trusts and a patient-level longitudinal observational study in four Trusts. We used natural variation in workload (beds per staff member) and, for the longitudinal study, staffing shortfalls relative to the expected staffing for the ward to determine the association between staffing levels and outcomes and to estimate the effects of change. For nursing, we considered RNs (band 5+) and nursing assistant staff (bands 2–4), which would include nursing associates (although current numbers are small). In the longitudinal study, we also considered the composition of the nursing team in terms of the staff grade mix (proportion of band 4 assistant staff, proportion of band 6+ RNs) and the proportion of bank and agency staff. Across the two studies, we considered a range of outcomes including death, length of stay, re-admission, patient experience, staff experience and staff sickness. For the economic analysis, we considered costs, consequences and quality-adjusted life-years (QALYs), estimating incremental cost-effectiveness ratios (cost per life saved and cost per QALY gained) for eliminating low staffing. Our data were derived from publicly available records of hospital activity, staffing and outcomes for the cross-sectional study and hospital administrative systems for the longitudinal study. We tested associations with multivariable mixed statistical models including random terms to account for the clustering of observations in Trust or ward, as appropriate. We used the national standardised hospital mortality indicator model to adjust for risk of death and for national patient and the staff experience models included the Trust's mortality rate to adjust for the acuity of the case mix.

Results

Cross-sectional study

We included 138 hospital Trusts. The number of beds per full-time equivalent (FTE) staff member varied considerably between Trusts. The largest variation was in allied healthcare professional (AHP) staff (mean 2.4 beds per FTE) and support to AHP staff (mean 11.1 beds per FTE), where the standard deviation was 38% and 44% of the mean, respectively. RN staffing levels were strongly correlated with staffing levels by doctors ($\rho > 0.71$). Although the number of beds per RN had the largest effect on mortality in models including single staff groups [rate ratio (RR) 1.33, 95% confidence interval (CI) 1.15 to 1.54], this was greatly reduced and no longer statistically significant when all staff groups were included in the model, although it remained the largest effect size (RR 1.07, 95% CI 0.88 to 1.31).

In multiprofessional models, more occupied beds per AHP (RR 1.04, 95% CI 1.02 to 1.06) and per medical doctor (RR 1.04, 95% CI 1.02 to 1.06) were associated with increased risk of death. More beds per nurse support (RR 0.85, 95% CI 0.79 to 0.91) and AHP support (1.00, 95% CI 0.99 to 1.00) were associated with lower death rates. In multiprofessional models, having more beds per RN was associated with lower scores for patient experience, staff health and well-being, and staff reports of quality of care. More beds per nurse support were associated with lower morale scores but more beds per surgical doctor were associated with higher morale scores. Using ward-level reports of nurse staffing we found that wards with more RN hours per patient day reported fewer harms on the national 'safety thermometer' but calculating a staffing shortfall, relative to the Trusts' reported staffing plans, did not strengthen the observed relationship and, for nursing assistant staff, shortfalls were associated with reduced harms.

Longitudinal study

We linked staffing data for 626,313 adult admissions and 57,375 paediatric admissions with overnight stays from four hospital Trusts. We used the ward mean to define the expected staffing level, with staffing below this classified as low. During the first 5 days of hospital stay for adult patients low staffing (relative to the ward mean) occurred on 45% of days, with a mean exposure of 3.32 days. On days when RN staffing was low, the mean shortfall was 0.87 RN hours relative to average staffing of 5.3 hours per patient day. For nursing assistants, the mean shortfall on days of low staffing was 0.49 hours relative to average staffing of 2.9 hours per patient day.

For adult inpatients, exposure to days with lower-than-expected RNs or nursing assistant staff was associated with increased hazard of death [adjusted hazard ratio (aHR) 1.08/1.07, 95% CI 1.07 to 1.09/1.06 to 1.08] and longer hospital stays. Low RN staffing was also associated with increased hazard of re-admission (aHR 1.01, 95% CI 1.01 to 1.02). Results for paediatric admissions were similar, although low nursing assistant staffing was also significantly associated with increased risk of re-admission. Effect sizes were larger, but CIs were much wider. The risk of death was very low (0.25%) and associations with deaths were not statistically significant. Exposure to days of low RN staffing increased the odds of sickness absence for both RNs and assistants (adjusted odds ratio 1.02/1.03, 95% CI 1.00 to 1.03/1.01 to 1.04 for each 10% of the past 7 days worked that were understaffed).

Our primary analyses included staffing in the first 5 days of the hospital stay but results were not sensitive to the exposure window used. We had no measure of the availability of AHPs, doctors or other staff groups, but we tested models that included weekend and seasonal effects, both factors that may cause short staffing from other groups to correlate with low nurse staffing. Addition of these factors did not alter the estimated effect of low staffing. We modelled effects for various subgroups. The effects of low staffing on general wards and for less acute patients [National Early Warning Score (NEWS) < 5] only were similar to the original results. Effects in highly acute patients (NEWS 5+) and older peoples' wards were smaller. For highly acute patients, there was evidence of an adverse effect from nursing assistant staffing, as low staffing was associated with significant decreases in the hazard of death (aHR

0.98, 95% CI 0.96 to 1.00) and increased risk of re-admission (aHR 0.97, 95% CI 0.95 to 0.99). Nursing assistant low staffing was also associated with a lower hazard of re-admission for people over 75. There was no statistically significant association between low staffing and mortality in intensive care or for people over 75, but hospital stays were longer.

We found that days of low staffing were also associated with increased risk of the potentially nurse-sensitive adverse events of deep vein thrombosis (DVT), pneumonia and pressure ulcers in surgical admissions. Low RN staffing had a larger effect than low assistant staffing. For example, a patient who experienced low RN staffing on all the first 5 days of their stay had a 59% increased risk of DVT. For low assistant staffing, the risk was increased by 33%. Analysis of data from incident reporting systems gave counterintuitive associations consistent with ascertainment and reporting bias.

We used the ward mean as a threshold to define low staffing. Although our previous research has shown that mean staffing correlates strongly with planned staffing on a ward, it is an arbitrary threshold and may not reflect a desirable staffing level or any tolerance to lower than planned staffing. Estimated effects of low staffing were largely unchanged if the threshold to define low staffing was set at lower levels. For assistant staffing, this is also the case for higher thresholds, above the mean. However, for RN staffing the effects of days of low staffing increase when a higher threshold is used. For example, if low staffing were defined as when staffing falls below 110% of the mean, the HR associated with a day of low staffing is 1.12. We also explored continuous (net hours) effects, nonlinear and interaction effects. There was some evidence of nonlinearity and interaction, but the effects were mostly subtle, although there was some indication that the effect of additional days of low RN staffing was greater as the cumulative number of days exposure increased and that the marginal effect of change in net hours was slightly greater at very high as well as very low staffing levels.

We considered the mix of staff in terms of grades of RN and assistant staff and proportions of bank and agency staff. A higher proportion of band 6+ RNs and band 4 nursing assistants were associated with reduced hazard of death (aHR 0.98), although the result was not statistically significant for nursing assistants nor when all staff mix variables were considered simultaneously. Higher proportions of bank and agency staff (both RNs and nursing assistants) were associated with increased hazard of death, with the strongest adverse effect associated with agency assistant staff.

Economic analysis

The estimated cost of providing care for our analysis cohort of adult inpatients was £4173 per admission. The mean cost of avoiding low staffing was £197 per admission. The mean discounted QALY lost among patients who died was estimated to be 6.82 and we used this as our base-case assumption for the QALY gained when modelling the effects of reducing understaffing. Eliminating low staffing cost £2778 per QALY gained. Savings from avoided sickness absence and re-admissions made cost-effectiveness estimates more favourable but did not have a major impact. If costs of avoided hospital stays are included, avoiding low staffing generates a net cost saving under all scenarios modelled except reducing understaffing by nursing assistants for highly acute patients, which led to a net cost increase and worse outcomes.

Avoidance of RN rather than assistant understaffing for highly acute patients gave more benefits and was more cost-effective. Avoiding low staffing using temporary agency staff reduced mortality but was less cost-effective than using substantive staff because benefits were reduced, and costs increased. Assuming agency staff cost 50% more than the cost of substantive staff, reducing low staffing with agency staff cost £10,980 per QALY. If agency staff were assumed to cost the same as substantive staff, this was still less cost-effective (£7320 per QALY) because the benefit of reduced mortality was weakened by the adverse effect from a higher proportion of temporary staff.

Conclusions

The NHS faces multiple competing demands for scarce resources. The evidence presented here suggests that investment in nurse staffing in acute hospitals could be cost-effective, on a par with many public health interventions. If the benefits of reduced length of stay are considered and realised, for example through freeing capacity to improve flow through emergency departments or for elective surgery, then there could be net gains. The relative increase in costs is modest, although the supply of staff to meet demand remains challenging. It is important that this scarcity does not obscure the need and demonstrated value for money. While decision-makers may, of necessity, need to experiment with novel approaches to addressing staffing shortages, this needs to be done in the context of a full understanding of what is already known. The safety and cost-effectiveness of alternatives should not be assumed.

Several priorities for future research emerge from this work.

- More research is needed into methods to determine nurse staffing requirements in hospital wards, for planning, real-time monitoring and for use in research. The requirements of service should inform decisions about the required timeliness of data, acceptable data gathering load and the necessary precision.
- Our findings, combined with the results of previous research, leave uncertainty about the trade-offs between staff shortages and temporary staffing levels, including the relative (adverse) effects of temporary staff at different levels and from different sources. Both qualitative and quantitative research would be of value.
- There remains uncertainty about the interaction between RN and assistant staffing levels which should be addressed through both qualitative and quantitative research.
- Research is required to better understand whether the observed variation in AHP staffing is based on variation in service and patient need. The observed association with mortality rates in this study suggests it may not be, and if that is the case, evidence-based methods for determining appropriate staffing need to be developed.

Our results not only show the adverse effects of low nurse staffing but also show that medical and AHP staffing are important considerations for patient safety. Eliminating low RN staffing gives more benefits than eliminating assistant staffing but both interventions are cost-effective in terms of QALYs gained relative to many public health interventions. Using agency staff to reduce staffing shortages is also less cost-effective than using substantive staff, because of higher costs but also reduced benefits. However, these findings suggest that while relatively less cost-effective, the use of agency staff to avoid staff shortages is still cost-effective relative to many public health interventions. These findings lend support to policy initiatives aimed at increasing the supply of RNs.

Study registration

This study is registered as Current Controlled Trials ClinicalTrials.gov NCT04374812.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme (NIHR award ref: NIHR128056) and is published in full in *Health and Social Care Delivery Research*; Vol. 13, No. 25. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Adequately staffing hospital wards is challenging for England's NHS. As of March 2022, approximately 10% of nursing posts in acute settings were vacant with 38,972 vacancies recorded.[1] In the aftermath of the Francis inquiry and the publication of guidance for safe staffing in adult wards in acute hospitals by the National Institute for Health and Care Excellence (NICE) in 2014, the number of registered nurses (RNs) employed in acute hospitals increased, but this followed a period where absolute numbers fell, and the increases have not matched activity growth.[2] Although there appeared to be an immediate uplift in the number of applicants for nursing courses following the start of the COVID-19 pandemic in 2020, this has not been sustained and steps taken to increase supply will not resolve shortages for some time, assuming they are successful.[3] In the face of such scarcity and the need to manage expenditure on staff, care providers and policy-makers face difficult decisions as they plan how to provide adequate nurse staffing levels.

The consequences of staff shortages are potentially serious. Several inquiries, NHS guidance and an extensive body of research indicate that lower RN staffing levels are associated with adverse patient and staff outcomes.[4–8] Patient and care outcomes associated with lower nurse staffing include increased risk of death,[9] hospital-acquired infections,[10] falls,[11] poor patient experience[12] and omissions in nursing care.[13] Adverse nurse outcomes include burnout, reduced job satisfaction and intention to leave.[14]

Temporary and flexible staffing

One consequence of high levels of vacancies and increased demand for staff has been a reliance on temporary staffing solutions to maintain safe staffing levels. On one hand, use of temporary staffing provides a potentially efficient way of managing a scarce labour resource, with staff only deployed when required. On the other hand, temporary staff are potentially expensive to employ, especially when agencies can inflate prices in the face of scarcity. Concerns about the high cost of using agency staff have led to the implementation of price caps and other measures designed to limit expenditure on temporary staff.

In a modelling study, we found that staffing models that rely heavily on flexible staffing to make up staffing shortfalls when demand exceeds expected levels are not cost-effective.[15] Risks to patients are increased when high use is made of temporary staff,[16] partly because temporary staff are inevitably less efficient than fully orientated permanent staff and partly because some shifts are likely to remain unfilled. Cost savings of this flexible staffing model are relatively small and arise largely because the required staffing level is not achieved.[15] Where there is unlimited availability of temporary staff, the risk to patients is still increased and there is little, if any, cost saving, even in the face of salary caps for agency staff. The balance of risks between short-staffing and use of staff who are unfamiliar with the ward remains uncertain.

Skill-mix

Unsurprisingly, as RNs are in short supply, alternative approaches to meeting patient needs are being considered. Hospitals already deploy significant numbers of unregistered healthcare assistants to deliver much hands-on care to patients in hospital wards.[17] There is currently substantial variation in both absolute staffing levels and the composition of the nursing team (generally referred to as 'skill-mix') both between and within hospitals,[17–19] and the support workforce has grown at a faster rate than the RN workforce in recent years.[2]

While the evidence indicates the potential adverse effects arising from shortages of RNs, there is no clear evidence that increasing the number of support workers can mitigate the safety risks from low nurse staffing. Results from our pan-European RN4CAST study found that benefits from increases in team size associated with adding support staff were offset by reduced skill-mix, which was associated with worse outcomes.[20,21] Our cross-sectional study in England found some evidence that hospitals employing more nursing support staff per patient had increased mortality, after controlling for the number of RNs and medically qualified doctors.[17] In contrast, our patient-level longitudinal study found a U-shaped relationship between nursing assistant (NA) staffing and mortality, with higher levels of staff reducing the risk of death up to a tipping point, after which risk of death began to increase rapidly.[19,22] These and other findings (e.g. Bridges *et al.*[23]) point towards complex interactions between the increased capacity provided by support staff, the appropriate delegation of care and the capacity of registered staff to properly supervise.[19] Consequently, the extent to which assistant staff or new support roles can be used to support or substitute for RNs in NHS hospital wards remains uncertain. Likewise, the consequences of current variations in staffing, both planned and unplanned, are unclear.

Despite this, for some time, workforce policy has advocated an expansion of the workforce at 'band 4', the grade below a RN, and in 2017 the NHS in England introduced a new role, regulated by the Nursing and Midwifery Council, the Nursing Associate. Two thousand trainees commenced in 2017 and have started entering the register since early 2019 with numbers set to increase rapidly. However, as of March 2022, the number of Nursing Associates on the Nursing and Midwifery Council register was small (6874) compared to the number of RNs (711,264) and is likely to remain so for some time.[24]

Cost-effectiveness

Decisions about the size and composition of the nursing workforce in hospitals have major implications for both the costs and quality of care delivered to patients. In 2018, there were approximately 9 million admissions to hospitals in England.[25] Although the risk of death is low, nearly 300,000 people (3.3%) die while in hospital or shortly after discharge. Investment in nurse staffing, specifically in RNs, has long been advocated by professional bodies and others as a key strategy to reduce avoidable harm.[26,27] Commenting on a recent history of the patient safety movement, Schiff and Shojania point out:

> *Most safety interventions focus on educating staff or bombarding them with alerts, alarms and policies. We have done little to reduce the production pressures that drive care at the frontlines, nor improve nurse staffing ratios.*[28]
>
> *p. 151*

While the associations reported between nurse staffing and rates of death are typically small, many people are affected and so the potential benefits are great. Taking an estimate derived from our recent National Institute for Health and Care Research (NIHR)-funded study, each additional hour of RN care per patient day available over the first 5 days of a patient's stay was associated with a 3% reduction in the hazard of death.[22] If the relationship observed in this research was a causal one and generalised across the English hospital population, increasing RN care available by 1 hour per patient per day could be associated with approximately 24,000 fewer deaths per year.

On the other hand, changes in nursing workforce have substantial cost implications and RNs are currently in short supply. The workforce accounts for 65% of all NHS spending.[29] In 2017, there were about 175,000 RNs employed in English hospitals, with an additional 90,000 nursing support staff.[30] Training a RN takes 3 years and is estimated to cost £75,000.[31] There are limited data available to determine the cost-effectiveness of changes in nurse staffing. Most data come from hospital-level cross-sectional studies, with significant risk of bias in estimates of effect, and few studies have given estimates based on NHS costs.[4] In our recent longitudinal patient-level study, we estimated that providing 1 additional hour of RN care per patient per day would increase labour costs for a large (800-bed) general hospital by about £10.1M per year, consisting of £219 per patient and amounting to a 2% increase relative to the hospital's annual budget.[19] We estimated the cost per life saved from this increase in staffing to be £47,376.

Whether increasing nurse staffing levels would be considered a 'cost-effective' intervention is dependent on many factors including uncertainties about life expectancy and quality, and the appropriate threshold for cost-effectiveness, considering that in a resource-constrained system, there are opportunity costs, and the key decision is about where scarce resources are best invested. While cost-effectiveness thresholds approximating to per capita gross domestic product (GDP) are often quoted, the opportunity cost arising from other benefits forgone means that appropriate thresholds are likely to be much lower for the NHS,[32] especially in view of the large costs involved in investing in nurse staffing. Furthermore, this cost-effectiveness estimate was based on parameters derived from a single-site study, albeit using national reference costs.

Causal inferences and estimation of staffing effects

Much previous research in this area has been limited because of the widespread use of cross-sectional designs, often estimating associations between annual staffing averages and outcomes. While there seems little room for doubt that there is some causal element in the resulting associations, for some, doubt nonetheless remains. In 2012, the chief executive of the Council for Healthcare Regulatory Excellence, a body that oversaw the work of professional regulators including the Nursing and Midwifery Council, was quoted thus:

> *There is no direct correlation between number of staff and good or bad care.*[33]

In complex human interactions, few causes are singular and many sources of bias make estimates of the causal effects challenging.[4] A singular limitation of much of the cross-sectional research has been the omission of consideration of staff groups other than nursing, even though the potential contribution to patient safety seems obvious and has been demonstrated in a number of studies. This is problematic not only because staffing levels from different groups tend to be correlated and thus the effect of other staff groups can be wrongly attributed to nursing[17] but also because an exclusive focus on nursing potentially masks the contribution of other staff groups in producing quality and safety.

Recent developments in e-rostering software have created opportunities to move beyond cross-sectional research to explore longitudinal relationships between daily staffing experienced by individual patients on wards and outcomes. Because these longitudinal studies measure patient-level variation in staffing exposures, nurse staffing associations are not intrinsically confounded by staffing from other groups other than where day-to-day variation is directly correlated. However, compared to the large body of evidence from cross-sectional studies, the number of studies using longitudinal designs is limited.

Building from previous research

While our previous study went some way to address the lack of economic evidence, it is limited to a single site.[19] Furthermore, while the staffing/outcome associations we found are consistent with other evidence, studies showing longitudinal relationships at the patient level remain rare and have largely been limited to single hospital studies.[34] Our findings on support staff provided a more nuanced picture of the contribution of support staff compared to previous research, identifying important benefits from having sufficient support but harm associated with having too many, but so far, this is the first study to find such a relationship.

In our economic analysis we considered only staffing and excess bed-day costs, and limited range of outcomes. We did not consider, for example, re-admission as an outcome or cost. Additionally, low nurse staffing has been associated with adverse staff outcomes including burnout and intention to leave.[14] While important in themselves, these factors also have implications for the costs and/or effectiveness of the ward team as sickness absence and staff turnover may arise as a result. Finally, while we estimated the cost per life saved, we were unable to consider the likely length or quality of lives saved, making comparison with the value of other investments in health care difficult. As the challenge to maintain a safe nursing workforce increases, it is essential to further our understanding of the costs and consequences of nurse shortages and strategies to remedy them, such as substituting assistant staff.

As part of the dissemination of our previous study, we organised a gathering of interested NHS leaders, researchers, and members of the public. We took the opportunity to discuss future research priorities with this group. In addition to addressing the gaps outlined above, they identified the importance of up-to-date data in the light of changing profiles within the nursing care workforce and the importance of considering the contributions of a wider range of professional and nonprofessional staff. This is particularly important, as shortages of RNs may not be easily remedied in the short or even medium term. Evidence that can suggest alternative potential ways of addressing safe staffing is vital and, furthermore, it is important to ensure that focusing on nurse staffing does not have unintended consequences by implicitly encouraging disinvestment in other staff groups who may make an equally important contribution to safe and effective care on hospital wards.

Aims and objectives

This study therefore aims to provide evidence to support decision-making about cost-effective deployment of nursing and other care staff on hospital wards in England and strategic decisions about policy to address current nursing workforce shortages. We address the question of which combinations of care staff employed by hospitals and deployed on hospital wards provide the most cost-effective care in terms of patient safety, experience and efficient use of resources.

Staffing inputs by some important groups (e.g. doctors and allied health professionals) are not readily tracked at a ward/patient level in routine data. Therefore, we initially explore the impact of different workforce configurations (including medical and therapy staff) on risk-adjusted mortality rates and patient and staff experience at a hospital level, using nationally reported hospital-level staffing and outcome data (study 1).

In study 2 (longitudinal, patient level), we use hospital patient records and electronic roster data from four hospitals to model how variation in daily care staffing levels and staff mix are associated with:

- death in hospital and within 30 days of admission
- length of hospital stay
- re-admission within 30 days of discharge
- other adverse outcomes [e.g. pressure ulcers (PUs), hospital-acquired infection, falls]
- staffing costs
- staff absence due to sickness
- incremental cost per life saved and per quality-adjusted life-year (QALY)

Chapter 2 Literature review

Background and overview of relevant literature

There are several hundred studies of associations between nurse staffing and patient outcomes in acute hospitals, including data from thousands of hospitals and millions of patients over many years and countries. Evidence has been synthesised in numerous reviews.[4,10,13,35–45] Reviews generally note the extensive body of evidence (both in terms of number of studies and the size of individual studies) but also highlight limitations, in particular the fact that studies are nearly all observational and mostly cross-sectional. Most studies focus on general medical and surgical patients in general wards, but research has also been undertaken in settings such as intensive care units (ICUs)[35,45] and paediatrics.[41] Studies have also focused on patient subgroups and specialties including older people and those with cognitive impairment,[46] oncology,[47] aortic aneurysm[48] and stroke.[49]

The most widely studied outcome by far is mortality. Not all studies have shown statistically significant associations, but the overall pattern of results remains consistent. Lower nurse staffing in general medical/surgical units and intensive care is associated with a higher risk of patients dying.[4,35,39,42] Typical effect sizes associated with changes in nurse staffing are low, although variations in how staffing is measured make direct comparison between studies difficult.

An early meta-analysis provided a pooled estimate of effect from North American studies. An additional 8 hours of RN time per patient day was associated with a 16% reduction in the odds of dying for surgical patients (eight studies), a 6% reduction for medical patients (six studies) and a 9% reduction for intensive care patients (five studies).[39] While the effects estimated in the Kane *et al.*'s review appear to be large, the scale of the implied staffing change is also large. Staffing levels in North America are higher than those typically observed in Europe but even in the USA, an 8-hour increase in RN time would be a large proportional increase in staffing. It would represent more than doubling of the California mandatory minimum staffing level in general medical/surgical units, which equates to approximately 5 RN hours per patient day (HPPD; 1 : 5 nurse-to-patient ratio). Typical staffing levels in equivalent units in the UK and many other European countries are lower,[14] with < 3 RN HPPD observed in some general wards.[19]

The growing volume of studies, the diversity of health systems covered in these studies and the range of different measures of staffing input used mean that more recent reviews have avoided meta-analytic synthesis. An indication of likely effects from real-world change and variability is the 2% reduction in the odds of death associated with 1 additional hour of RN care per patient per day, estimated in our study of 138,000 patients admitted to one large English hospital over 3 years.[19] While the effects may be small in absolute terms, the number of patients who are affected means that at a population level such small effects are important. We estimated that an increase of 1 RN hour per patient day would be associated with 219 avoided deaths. There was, however, substantial uncertainty in this estimate, associated with the precision of the estimated effect of staffing on death.[19]

A recent 'umbrella review' of 15 reviews concluded that there was strong or moderate evidence for associations between nurse staffing levels and a range of other outcomes including length of stay (LOS), patient satisfaction, quality of care processes, re-admissions and hospital-acquired pneumonia.[36] 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.[3,50] However, past reviews of this literature have noted the preponderance of cross-sectional studies. Such studies are unable to establish that the observed variation in staffing levels between hospitals or units corresponds to variation experienced by the patients whose outcomes were measured. This is because both exposures and outcomes are aggregated, frequently representing annual averages at the hospital level over a year. As noted above, this leads some to reject a causal inference entirely. Although the application of epidemiological principles supports a causal inference[4,39] and makes it difficult to sustain an argument that there is no causal influence between nurse staffing and patient outcomes, the potential for bias in existing research remains high. This means that estimates of the consequences of low staffing or the costs and benefits of investment in nurse staffing from existing research have a high chance of being systematically incorrect. Longitudinal studies, which directly link patient outcomes to the staffing experienced by the individual patient, or which otherwise prospectively assess the impact of change, have the potential to reduce or remove many sources of bias, but previous reviews have noted that such studies are rare.

Some outcomes that have been attributed to nurse staffing levels could also be attributed to the input of other professions, most notably doctors.[36] While some previous studies have also controlled for or included medical (physician) staffing when assessing the association between nurse staffing and outcomes (e.g. Griffiths *et al.*[17]), medical and other multidisciplinary team staffing is a notable omission and potential source of bias in most studies estimating the effect of nurse staffing.[4] When undertaking scoping searches to prepare this background literature review, we found no papers giving an overview of studies exploring the associations between staffing and outcomes for other professional groups and so the extent to which cross-sectional associations might be biased by the omission of other professional groups is hard to assess, as is the relative effect from investment in nurse staffing as opposed to other groups.

There have been previous reviews focusing specifically on economic studies, but studies were limited, and the reviews are now dated.[4,51] of cost-effectiveness/benefit vary widely and are hugely contingent on estimates of underlying effect (which are, in turn, subject to bias in the underlying cross-sectional studies), the cost perspective taken, as well as health costs and earnings in the country of study. Reviews revealed no studies that would allow a comparison of investment in nurse staffing with other staff groups or other changes in health provision using cost–utility. Our cost-effectiveness analysis for one English hospital identified a cost per life saved from increasing average staffing by 1 RN hour per patient day of £69,097. Caution needs to be applied to using any estimate derived from a single site study, but in any case, the extent to which this estimate represents a good investment from a health system perspective depends not simply on the 'lives saved' but also on the length and quality of life that results.

To enhance the broad overview of the literature presented above, we undertook more in-depth searching to establish the extent of existing literature in the following areas:

1. In order to add to and update existing reviews of economic studies, we searched for and reviewed economic evaluations of alternative nurse staffing configurations (PROSPERO registration CRD42021281202).[52]
2. To identify studies that are potentially less prone to bias, we reviewed studies that explored the association of nurse staffing with patient outcomes using longitudinal designs (PROSPERO registration CRD42020191798).[53]
3. To consider the influence of staffing levels from other professional groups in maintaining safety, and the consequence of their omission on the estimated effects of nurse staffing, we reviewed studies that considered the wider multidisciplinary team, focusing on mortality as a common measure of patient safety and in order to help focus our searches (PROSPERO registration CRD42020219869).[54]

Here we focus on reporting the first of these reviews in detail and provide brief summaries of the others which have been published in full elsewhere.

Review methods

We based our search strategies on those developed for the systematic reviews of nurse staffing literature for NICE.[55] We searched PubMed/MEDLINE, EMBASE, Cumulative Index to Nursing and Allied Health Literature and the Cochrane Library. Additionally, for economic evaluations, we searched the NHS Economic Evaluation Database and the Health Technology Assessment database maintained by the International Network of Agencies for Health Technology Assessment. For detailed search strategies, see *Appendix 1*, *Table 27*. Searches were undertaken and finalised between April 2021 and October 2022 (economics review). Although we cannot guarantee the searches are comprehensive beyond these dates, members of the study team monitored content alerts for new studies that would significantly alter the overall conclusions.

We selected material against criteria specified in the individual study protocols with a view to identifying (1) economic evaluations of the consequences of nurse staffing variations (including cost-minimisation, cost–benefit, cost–consequences, cost-effectiveness and cost–utility studies) in secondary modelling studies, observational studies or prospective intervention studies, (2) all studies that explored nurse staffing and patient outcomes relationships using a prospective or other longitudinal design and (3) studies that explored associations between mortality rates and at least one other non-nursing staff group alongside nursing. We undertook a structured assessment of risk of bias appropriate to the study designs. We used the Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) tool[56] to assess the risk of bias for our review of longitudinal studies because we wanted to determine the strength of each study to estimate an intervention effect from changes in nurse staffing. For elements of the review, where we anticipated mainly cross-sectional studies, we used a framework based on that used for NICE public health guidance, adapted and tailored to studies of staffing and outcome associations.[55,57] For economic studies, we additionally used Henrikson and Skelly's framework, which brings together common domains from three economic reporting checklists.[58]

Economic evidence

Parts of this section is reproduced with permission from Griffiths *et al.*[59] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Volume and quality of economic evidence on nurse staffing

From 6783 records identified in searches, we found 23 studies (reported in 24 papers) that included some form of economic analysis of change/differences in nurse staffing (see *Appendix 1*, *Figure 6*). Fourteen were published in the decade from 2013 and most were not available when previous reviews were undertaken. Most studies (17/22) addressed staffing on general medical and/or surgical units while the rest addressed specific surgical specialties or procedures[60–64] or maternity care.[65,66] Most studies (16/23) were conducted in the USA,[61,63–65,67–77] 3 (reported in four papers) in the UK,[15,19,66,78] and 1 each in Australia,[79] Belgium,[80] China[62] and South Korea.[60] In total, data came from 5900 hospitals and over 41 million patients. All studies were observational or sourced parameters and data for modelling from observational studies. Most studies used estimates of effects based on natural variation in RN or midwife staffing expressed as a staff-to-patient ratio (or vice versa), using outcome associations to model the effect of various changes in staffing levels. Of these, three used parameters from natural variation to model the effects of planned change across health systems[67,73,80] (typically increasing staffing to the 75th centile). One study used a mathematical simulation model to explore the effects of different approaches to determining staffing levels on achieved staffing.[15,78] In two studies, the observed variation in staffing was associated with implementing a method to determine staff requirements which led to increased staffing levels,[79] or a minimum nurse-to-patient ratio staffing policy[81] (see *Table 1* and *Appendix 1*, *Table 28*).

Overall, nine studies were rated as having a moderate risk of bias with the remaining rated as having a high risk of bias (see *Appendix 1*, *Table 29*). However, because of large sample sizes and use of risk adjustment, most studies were assessed as having a low risk of bias related to power and control of confounders, but the intrinsic design limitations of cross-sectional studies meant that only two studies (reported in three papers) were rated as strong for internal validity,[15,19,78] with a further seven rated as moderate.[69–71,75,77,79,81] Most studies (18) used estimates for the effect of nurse staffing that were cross-sectional in the sense that staffing levels are aggregated over a large unit (typically a hospital) over time (typically a year) and linked with outcomes of patients admitted over that period. Of these, four use potentially stronger matched cohort designs.[69–72] Four studies provide (or use) estimates of staffing/outcome associations that directly link patients to staffing at a day or shift level (or equivalent)[15,64,77,78] and one compared outcomes before and after a planned change in staffing[79] and one derived estimates of effect with changes associated with implementing minimum staffing legislation.[81] Despite the preponderance of large multi-hospital studies, only seven studies were rated as potentially strong for external validity[61,66,67,72,73,75,80] with a number of large studies downgraded to moderate risk of bias because there was a mismatch between the patient subgroup providing outcomes and the patient population served by the staff included.

The economic analysis in most studies was a disaggregated cost–consequences analysis with a range of consequences reported – typically restricted to some or all of mortality rates, adverse incidents, LOS and re-admissions (see *Appendix 1*, *Table 30*). The economic perspective was that of the hospital in most studies and in five studies only staffing costs were considered.[61,62,65,66,80] The remainder considered at least some consequential costs ranging from costs of extended stays, treatments of adverse events, re-admissions and, in two cases, societal costs in terms of lost earning or productive capacity.[67,75] While most studies took an immediate perspective on both outcomes and cost, two took a lifetime perspective on outcomes,[79,80] estimating life expectancies, while two[67,75] considered lifetime future earnings/productivity. We were able to extract or calculate a cost-effectiveness ratio related to death as an outcome from eight studies[15,19,67,69,70,73,79,80] and one study provided a cost–benefit analysis in terms of a ratio of staff costs to financial benefits arising from care cost savings and future productivity.[75] While many studies undertook some form of sensitivity analysis, estimates of economic parameters did not always reflect underlying uncertainty [e.g. 95% confidence intervals (CIs) of staffing-outcome effects].

Costs and cost-effectiveness

For details of the main economic results, see *Table 1*. In all cases, increases in staff led to increased staffing costs, as did increases in skill-mix. Seventeen studies provided estimates of net costs associated with staffing increases, considering other costs/savings that might result from staff changes. Of these, five found that increases in RN staffing levels in general medical/surgical or other surgical specialties led to reduced costs overall.[60,64,75,76,81] All but two of these studies were rated as being at high risk of bias. Of the studies with moderate risk of bias, one found that economic benefits to society, including losses to productivity avoided, exceeded costs with a benefit-to-cost ratio for each additional RN between 1.27 and 2.51.[75] In this study, 1 additional RN per 1000 surgical patients in US hospitals cost $923,832 but yielded a benefit of $1,646,190. For medical/intensive care, costs of an additional RN per 1000 patients ($982,800/$589,680) were also less than benefits ($1,244,061/$1,479,933).[75] A second study, rated as moderate risk of bias, estimated that implementing mandatory minimum staffing levels in Australia also yielded net financial benefits with cost savings from reduced hospital stays and re-admissions exceeding the costs of the increased staffing required to meet the mandatory minimums by a factor of 2.[81]

Two studies in US general medical/surgical patients found no statistically significant differences in net cost from staff increases,[70,72] while the remaining nine found net cost increases.[15,19,61,63,67,69,71,73,74,78] In all but one study, there was evidence of improved health outcomes associated with increased staffing. Clark *et al.* (2014) found increased costs but no evidence of reduced complications from increased nurse staffing to achieve one-to-one staffing during induction of labour in US maternity settings.[65]

In total, five studies considered aspects of skill-mix. Of these studies, two were assessed as moderate risk of bias with one rated as low risk of bias based on internal validity, although both were single-site studies.[19,77] Four studies found improved health outcomes from increasing the proportion of RNs in the nursing team in medical/surgical settings (three general and one neurology/neurosurgery), while the fifth found that increasing the proportion of bachelors-educated RNs was associated with improved outcomes.[77] All studies that considered net costs found that a skill-mix that was richer in RNs was associated with reduced net costs overall.[19,72,73]

Figure 1 summarises results in a hierarchical matrix.[82] In total, six studies provided results that clearly supported increased RN staffing when using net costs, with a combination of no cost change but improved outcomes[70,72] or reduced costs and improved outcomes[64,68,75,76] in medical and/or surgical wards. Of these, two studies were rated as moderate for the underlying risk of bias.[70,72] However, most studies showed both increased costs and improvements in health outcomes, where incremental (cost-effectiveness) analysis is required to inform the economic decision. A single study gave results that clearly rejected RN staffing increases, but this used limited cost data and was at high risk of bias.[65] Of the studies that evaluated changes in skill-mix, all three studies that considered net costs supported a decision to increase skill-mix,[19,72,73,77] including one study with moderate (rather than high) risk of bias. However, all four studies using staff costs alone showed improved outcomes and increased costs, meaning that incremental analysis is required for decision-making.

We were able to extract or calculate mortality-based cost-effectiveness estimates from eight studies, and these are summarised, alongside the 2021 per capita GDP of the country providing the estimate in *Table 2*. Twigg *et al.*[79] and Van den Heede *et al.*[80] provide cost per life-year. In both cases, the cost per life-year is far below per capita GDP and adjustment for loss of utility (quality) is unlikely to substantively alter the conclusions that the staffing increases (in Australian general medical/surgical units and Belgium cardiac units, respectively) are likely to be cost-effective at a GDP-based threshold. McHugh's finding that staff increases lead to improved outcomes and net cost reductions means that the decision to implement a mandatory minimum staffing policy (and the resulting staff increases) dominates. For

TABLE 1 Summary of results of economic studies of nurse staffing

	Country	Patient group	Main results
General medical/surgical			
Dall *et al.* 2009[67]	USA	Med/surg	Employment costs of each additional RN, $83,000, yield economic benefit (through reduced treatment costs and increased productivity) of $60,000. Net cost $23,000 × 133,000 to save 5900 lives by increasing all hospitals to 75th centile.
Griffiths *et al.* 2018[19]	UK	Med/surg	Staff cost £65,092 (net cost £47,376) per life saved (1 RN HPPD increase). Staff cost £26,351 (net saving −£486) from skill-mix change + 0.3 RN and −0.3 assistant nurse HPPD.
Griffiths *et al.* 2020[15,78]	UK	Med/surg	Standard staffing (achieved 3.6 RN HPPD) vs low (achieved staffing 3.2 RN HPPD) £19,437/£13,117 staff/net cost per life saved. High staffing (achieved staffing 3.9 RN HPPD) vs standard £21,766/£8653 staff/net cost per life saved.
Lasater *et al.* 2021[68]	USA	Med/surg (selected patient groups)	Moving all hospitals to a 4 : 1 average patient-to-RN ratio (current mean 6.3) leads to 4370 lives saved (ARR 1%) and $720 million saved in shorter lengths of stay (0.5 days per patient) and avoided re-admissions (ARR 1.4%) over the 2-year study period. Costs of increased staffing not included.
Lasater *et al.* 2021[69]	USA	Surg (select)	Better-resourced hospitals (mean 4.3 patients per nurse, 85% RN skill-mix, 68% BSN nurse, PES 3.01) cost $203,500 per life saved vs worse-resourced (mean 5.8 patients per nurse, 78% RN skill-mix, 43% BSN nurse, PES 2.68).
Lasater *et al.* 2021[70]	USA	Med (select)	Better-resourced hospitals (mean 4.3 patient per nurse, 85% RN skill-mix, 68% BSN nurse, PES 3.01) had lower 30-day mortality (16.1 vs 17.1%), shorter stays (5.38 vs 5.66 days), more intensive care admissions (26.92% vs 25.97%), fewer re-admissions (32.3 vs 33.6%) vs worse-resourced (mean 5.8 patient per nurse, 78% RN skill-mix, 43% BSN nurse, PES 2.68). Net costs were similar ($18,848 vs $18,671 NS).
Li *et al.* 2011[61]	USA	Med/surg	For surgical admissions: + 1 nursing staff (RN, Licensed Practical/Vocational and assistant nurses) HPPD cost $261.45 (NS, $p = 0.095$) and + 1% skill-mix cost $27.54 (NS $p = 0.253$) per admission. For medical admissions: + 1 nursing staff HPPD cost $164.49 ($p < 0.001$) and + 1% skill-mix cost saved −$2.73 (NS $p = 0.704$).
Martsolf *et al.* 2014[72]	USA	Med/surg	One additional licensed nurse per 1000 inpatient days is associated with a −0.25% reduction in adverse events, a −0.033 reduction in LOS and a $166.5 increase in cost (NS, −$35 to $368.1 95% CI). Increasing the percentage of licensed nurses that are RNs by 1% is associated with −$87 reduction in cost.
McHugh *et al.* 2021[81]	Australia	Med/surg	167 FTE needed to meet ratio requirements for RN/enrolled nurses at a cost of AU$33,000,000 would prevent 145 deaths, avoid 29,222 days of stay and 255 re-admissions at a net cost of AU$33,000,000, saving AU$67,561,264 from LOS and AU$1,589,594 from re-admissions
Needleman *et al.* 2006[73]	USA	Med/surg	The staff/net cost of raising the number of licensed hours nationally to the 75th centile (10.23 HPPD) is $7538/$5819 million, avoiding 1801 deaths, 10,813 adverse outcomes and 2,598,339 hospital days. The staff/net cost of raising the proportion of RNs in licensed hours to the 75th percentile (0.94 HPPD) is $811/−$242 million, avoiding 4997 deaths, 59,938 adverse outcomes and 1,507,493 hospital days. The staff/net cost of both raising the proportion of RNs and the number of licensed hours to the 75th percentile is $8488/$5716 million nationally and avoids 6754 deaths, 70,416 adverse outcomes and 4,106,315 hospital days.
Rothberg *et al.* 2005[74]	USA	Med/surg	The incremental staff/net cost of moving from a patient-to-RN ratio of 8 : 1 to 7 : 1 is $45,900/$24,900 per life saved. The incremental staff/net cost of moving from a 5 : 1 to 4 : 1 ratio is $142,100/$70,700 per life saved.
Shamliyan *et al.* 2009[75]	USA	Med/surg	For intensive care admissions: + 1 RN per 1000 admissions net cost/societal benefit/benefit-to-cost ratio are: $589,680/$1,479,933/2.51. For surgical admissions, corresponding figures are: $923,832/$1,646,190/1.79. For medical admissions, figures are: $982,800/$1,244,061/1.27.

continued

TABLE 1 Summary of results of economic studies of nurse staffing *(continued)*

	Country	Patient group	Main results
Twigg *et al.* 2013[79]	Australia	Med/surg	Comparing pre–post net 12% increase in RN hours: Staff/net cost AU$13575/AUD$8907 per life-year.
Weiss *et al.* 2011[76]	USA	Med/surg	Increasing RN (non-overtime) staffing by one standard deviation (0.75 HPPD) led to a staffing cost $145.74 with a net −$409.59 saving (due to reduced re-admissions). Reducing RN overtime staffing by 1 standard deviation (0.07 HPPD) led to reduced staff cost of −$8.18, net saving −$19.16 per patient.
Yakusheva *et al.* 2014[77]	USA	Med/surg	Staff cost of increasing the BSN-educated staff to 80–100% was between $1,843,266 and $3,446,106 with $5,653,022.97 cost savings from shorter stays (−0.03 days) and re-admission rate (−1.7%).
Maternity			
Clark *et al.* 2014[65]	USA	Induction of labour	Staff cost of universal (in the USA) one-to-one midwifery staffing $97,000,000 (1618 FTE staff) annually and no evidence of benefits in terms of complications.
Cookson *et al.* 2014[66]	UK	General labour	Incremental cost-effectiveness ratio £85,560 per ‘healthy mother’ (staff) and £193,426 per delivery with bodily integrity from one additional midwife per 100 births.
Other			
Behner et al. 1990[64]	USA	Surg (back and neck)	For one unit for a 6-month period, days of low staffing (more than 20% below standard) reduce staff cost (−$13,600) and are associated with a 34% absolute increase in risk of complications and net cost $17,200 ($130 per patient). Nurse staffing undefined.
Kim *et al.* 2016[60]	South Korea	Surg (hip and knee)	Patients in high-staffed hospitals (bed-to-nurse ratio ≤ 2.0) are charged US$1142.2 less than those with the lowest nurse staffing level (bed-to-nurse ratio ≥ 6.0) and have shorter stays (13 vs 25 days).
Li *et al.* 2016[71]	USA	Surg (cardiac)	Hospitals with higher RN staffing (above median) had higher mean costs $2123, a 10–25% reduction in different ‘Healthcare-Acquired Infections’, a 6% reduction in mortality and a 0.3-day reduction in mean LOS.
Pang *et al.* 2019[62]	China	Med/surg (neuro)	Compared with 100% RNs, 75% RNs are associated with a decrease in staff costs of −CN¥573 (22%), an increase in urinary tract infection (OR 1.503, 95% CI 1.189 to 1.900; $p = 0.001$), fewer medication errors (OR 0.684, 95% CI 0.499 to 0.936; $p = 0.018$) and more successful ventilator weaning (OR 0.677, 95% CI 0.592 to 0.775; $p < 0.001$). Other outcomes NS.
Ross *et al.* 2021[63]	USA	Surg (pulmonary lobectomy)	Compared to low-staffed hospitals (≤ 3.5 RN FTEs per 1000 patient days) hospitals with high staffing (≥ 5.6 RN FTEs per 1000 patient days) had $4388 increased costs per patient, 0.37-day shorter stays and 36% lower odds of mortality (OR = 0.64, $p = 0.014$).
Van den Heede *et al.* 2010[80]	Belgium	Surg (cardiac)	On average, increasing RN staffing to the 75th percentile, an additional 0.8 FTE per unit costing a total €1,211,022 for all Belgian general cardiac postoperative units, €26,372 per life saved and €2639 per life-year gained.

ARR, annual rate of return; BSN, Bachelors of Science in Nursing; FTE, full-time equivalent; med, medical; neuro, neurology; NS, not significant; OR, odds ratio; PES, Practice Environment Scale; RR, relative risk; surg, surgical.

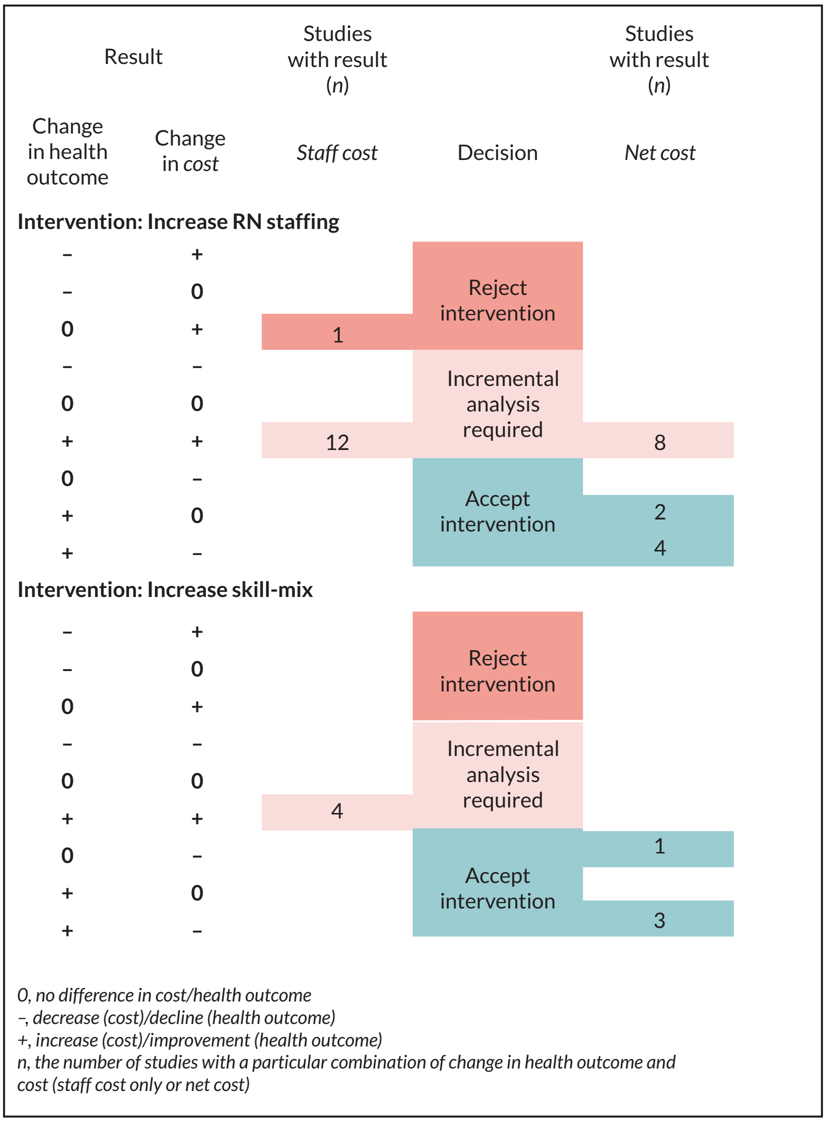


FIGURE 1 Hierarchical matrix to summarise findings and economic conclusions from economic studies of nurse staffing/skill-mix increase.

four studies in US and UK general medicine/surgery,[15,19,69,70,78] the ratio between per capita GDP/cost per life saved ranged from 0.3[70] to 3.2.[69] However, both the US studies provided estimates for a 'combined' intervention, implying both increased RN staffing and additional changes in skill-mix beyond that which would result from the staff increases. Even the higher end of this range is potentially cost-effective if each 'life saved' gains 3.2 QALYs. Other US studies require that each life saved yield more than 12 QALYs[67] or, in the case of Needleman *et al.*,[73] nearly 70 QALYs to achieve the GDP-based threshold. Therefore, in this context, simple staffing increases are unlikely to be cost-effective from a provider perspective, although Dall found net economic benefit arising from increased productivity (of survivors),[67] and Needleman *et al.* found that increased RN skill-mix was an economically dominant strategy, resulting in improved outcomes at reduced cost.[73]

Longitudinal studies

As the major source of bias in most economic studies arose from their cross-sectional design (or the cross-sectional design of underlying studies), we undertook a systematic review of longitudinal studies of nurse staffing and outcomes. These potentially offer confirmation of findings from cross-sectional studies, with lower risk of bias and stronger ability for causal inference. This review was published under a CC BY licence as '*Nurse staffing levels and patient outcomes: A systematic review of longitudinal studies*',[34] where it is reported in full.

TABLE 2 Costs per life saved from studies of increased staffing

Paper	Country	Patient group	Main results	Cost per life (life-year) saved (2021 equivalent)	Cost per life (life-year) saved (2021 US$ PPPE)	2021 per capita GDP (in US$)	Cost per life/per capita GDP ratio
Dall *et al.* 2009[67]	USA	General med/surg	Increasing RN staffing in all hospitals to the 75th centile	US$839,930	US$839,930	US$69,287	12.1
Griffiths *et al.* 2018[19]	UK	General med/surg	Increase of 1 RN HPPD	£54,009	US$77,957	US$47,334	1.6
Griffiths *et al.* 2020[15,78]	UK	General med/surg	Standard staffing policy (achieved RN HPPD 3.6) vs low staffing (achieved staffing 3.2 RN HPPD)	£14,560	US$21,016	US$47,334	0.4
Lasater *et al.* 2021[69]	USA	General surg (select)	Better-resourced hospitals (mean 4.3 patients per nurse, 85% RN skill-mix, 68% BSN) vs worse-resourced (mean 5.8 patients per nurse, 78% RN skill-mix, 43% BSN)	US$221,815	US$221,815	US$69,287	3.2
Lasater *et al.* 2021[70]	USA	General med (select)	Better-resourced hospitals (mean 4.3 patients per nurse, 85% RN skill-mix, 68% BSN) vs worse-resourced (mean 5.8 patients per nurse, 78% RN skill-mix, 43% BSN)	US$18,127	US$18,127	US$69,287	0.3
McHugh *et al.* 2022[81]	Australia	General med/surg	Increase staffing to meet specified ratio policy	AU$0 (AU$227,586)	(US$158,158)	US$59,934	0.0
Needleman *et al.* 2006[73]	USA	General med/surg	Raising the number of licensed hours nationally to the 75th centile	US$4,840,377	US$ 4,840,377	US$69,287	69.9
Twigg *et al.* 2013[79]	Australia	General med/surg	Implementation of RN HPPD staffing model – net 12% increase in RN hours	AU$12,114[a]	US$ 8418[a]	US$59,934	0.1
Van den Heede *et al.* 2010[80]	Belgium	Cardiac surgery	Increasing staffing to the 75th percentile (additional 0.8 FTE per unit)	€3510[a]	US$ 4726[a]	US$51,768	0.1

a cost per life-year.
BSN, Bachelor of Science in Nursing; FTE, full-time equivalent; Med, medical; PPPE, purchasing power parity equivalent; Surg, surgical.

Despite the widespread conclusion from previous reviews that most such studies were cross-sectional, we found 27 papers reporting longitudinal studies published since 2003, with a majority (17) published since 2016. More studies (seven) came from the USA than any other country,[83–89] with four other studies from North America, all from Canada.[90–93] Other studies came from the UK,[16,19,22,46,94] other European countries[95–100] and Australia.[81,101,102] Early studies mostly (six out of nine) addressed staffing in ICUs (including neonatal intensive care), while the more recent studies, from 2016 on, have mostly addressed general adult inpatient populations (13/17) in medical and/or surgical units. Most were single hospital studies (15/27), but there were some notable exceptions, with 3 studies including over 50 hospitals, but none of these were from the UK.[81,89,94] Patient samples ranged between 85[99] and 489,155.[81] Most post-2016 studies used electronic sources such as e-rosters for staffing data (11/17). All studies were observational, quantifying the consequences of 'naturally' occurring variation, but two could be construed as quasi-experimental[81,102] because they evaluated deliberate changes in staffing policy. Most studies measured individual patient exposures to low staffing or variation in staffing, with the 'exposure window' (i.e. the period of staffing measured prior to the outcome) ranging between 6 hours[90] and 30 days.[89] Staffing measures were operationalised in a variety of different ways, making direct comparisons between studies difficult and pooled estimates impossible.

We used the ROBINS-I tool[56] to assess the risk of bias relative to the results of a well-conducted randomised controlled trial with staffing change as the intervention. Despite the potential strength of longitudinal design, most studies were assessed as either at serious (12 studies) or critical (5 studies) risk of bias. Eight had moderate risk of bias and only three[22,91,97] were judged to be at low risk of bias.

Findings are consistent with an overall picture of a beneficial effect from higher RN staffing on preventing patient death, the most studied outcome. Focusing on results from studies with low risk of bias, among 138,133 adult patients admitted to general wards in a UK general hospital, the hazard of death was increased by 3% for every day a patient experienced low RN staffing. Low and high NA staffing was also associated with increases in mortality.[22] Among 146,349 adult medical, surgical and intensive care patients admitted to a Canadian university health centre, every 5% increase in the cumulative proportion of understaffed shifts was associated with a 1% increase in mortality. Neither RN experience nor skill-mix was significantly associated with mortality, but every 5% increase in the proportion of hours worked by bachelor's degree-prepared RNs was associated with a 2% reduction in the risk of death.[91] Among 79,893 adult inpatients in general and ICUs of a Swiss university hospital, exposure to shifts with low staffing was associated with 10% higher odds of mortality. The associations between mortality and staffing by other groups were less clear, with both high and low staffing of unlicensed and administrative personnel associated with higher mortality.[97]

While all three of these studies used deviations from expected staffing as a criterion, different definitions and thresholds were applied. On the one hand, Griffiths *et al.*[22] used the ward mean hours to define expected staffing and this was also the threshold for low staffing. On the other hand, Rochefort *et al.*[91] used the shift mean to define expected staffing, with 8 hours below expected as a threshold. Conversely, Musy *et al.*[97] used a regression-based approach to define expected staffing, incorporating patient numbers and turnover, using the median observed-to-expected ratio ± 0.5 to define high/low staffing. Griffiths *et al.*[22] and Rochefort *et al.*[91] used time-varying exposures in a survival framework, with Griffiths *et al.*[22] considering different exposure windows; the primary analysis reported exposure during the first 5 days. Thus, while these effect estimates are (broadly) similar in magnitude, they are impossible to compare directly and arise from single hospitals in three different health systems.

The evidence is less clear for other patient outcomes such as infections and these results also had a higher risk of bias, but in general, the proposition that higher RN staffing is likely to lead to better patient outcomes is supported.

Multidisciplinary staffing

We found no existing reviews of studies reporting associations between outcomes and healthcare staff levels for professional groups other than nursing. As a significant potential confounding factor is the omission of staffing by other professional groups, in our review we included only studies of staffing by nursing plus one or more other groups, as the association with nurse staffing is established. This also had the advantage of allowing us to assess if nurse staffing effects remained when other staffing effects were included in multivariable analyses. We focused on mortality as a critical safety outcome potentially relevant to all groups. This review was published under a CC BY licence as *'The association between multi-disciplinary staffing levels and mortality in acute hospitals: a systematic review'*,[103] where it is reported in full.

We found only 12 studies that explored the association between multidisciplinary staffing levels and mortality where nurse staffing was also included. Studies were published between 1999 and 2020 and included data from the USA (five studies), the UK (two) or other European countries (three) and South Korea (two). Studies were all multicentred, including data from between 4[104] and 3763 hospitals.[105] Patient sample sizes varied, ranging from 1864[106] to 23,879,998.[105] Studies with smaller samples focused on specific patient populations, for example patients who had a gastrectomy[106] or patients from intensive care settings only,[104,107] whereas the larger studies included less specific populations of general medical and/or surgical patients. All but one study used bed- or patient-to-staff ratios (or vice versa) to measure staffing levels. The sole exception used a staff-per-admission ratio.[108]

The risk of bias was assessed as high in most studies because of the underlying weakness of the cross-sectional designs. Cross-sectional associations with high levels of aggregation such as annual unit- or hospital-level staffing averages and outcome rates were typical. The sole exception explored the effect of shift-by-shift staffing levels on risk of death in ICUs.[104]

There was a statistically significant association between higher levels of RN staffing and lower mortality rates in 7 studies out of 12.[104,105,109–113] Where associations with RN staffing levels were not statistically significant, point estimates were still in the direction of a beneficial effect from higher levels of staffing.[17,106–108,114] The effect sizes were typically small. For example, an increase of 1 RN hour per patient day in Danish medical units reduced odds of death by < 1%[109] and an additional nurse per bed in US hospitals reduced the absolute death rate by 0.26% (mean death rate was 9.6%).[105] Estimated nurse staffing effects, while generally statistically significant, did tend to be lower in multivariable models controlling for other staff groups than in models including nurse staffing only. For example, in Griffiths *et al.*'s study of English NHS hospitals, a decrease in the mean RN workload from 10 or more patients to 6 or fewer was associated with a 20% reduction in the risk of death in the single staff group model, which reduced to 11% in the model including medical staffing levels.[17] Studies examining NA staffing and other grades of nurses found mixed results with some studies showing harm and some benefit from higher staffing levels by these groups.

The most frequently studied non-nursing staff group was medically qualified (physician/surgeon) staff, with 11 studies reporting associations. Of these, seven found that higher levels of staffing were statistically significantly associated with lower hospital mortality rates.[17,104,105,107,109,112,113] Other associations were not statistically significant but, where reported, coefficients favoured higher staffing.[110,111,113] Only two studies reported on staff groups other than medical and nursing staff, both from the USA. Although both these studies were very large, both are very old. Bond *et al.* analysed 1992 data from 3763 US hospitals and found that hospitals with more pharmacists and medical technologist staff per bed had lower mortality rates. Hospitals with more administrative staff per bed had higher hospital mortality. Other associations with multidisciplinary staff were not statistically significant.[105] Robertson and colleagues, analysing data from 1791 US hospitals (1989–91), found that higher levels of respiratory therapists and respiratory therapy technicians employed per 100 adjusted admissions were associated with lower mortality rates from chronic obstructive pulmonary disease. Other associations were not statistically significant.[108]

Review summary and conclusions

We found a considerable body of economic evidence about the effects of changes in nurse staffing and skill-mix that has been published since the last known reviews. Many studies use a limited range of costs, and few provide estimates in terms of cost per (quality-adjusted) life-year saved. Nonetheless, the results are consistent with increasing the level of RN staffing or the proportion of RNs in the nursing team being cost-effective interventions, potentially costing substantially less than per capita GDP per QALY. However, the studies come from diverse contexts and evidence may not generalise. Furthermore, in most cases, the underlying study designs used to estimate staffing outcome associations are at risk of bias.

The relationship between nurse staffing levels and patient outcomes, in particular mortality, is confirmed in longitudinal studies and a causal interpretation of the relationship seems beyond doubt, but few of these studies are able to provide unbiased estimates of effect sizes. Many of the sources of bias are likely to reduce the estimated effects of nurse staffing, because more acutely ill patients are likely to be judged to require a higher staffing level, which creates an endogenous relationship whereby high risk leads to higher staffing levels. Studies with lower bias are single-site studies and variation in contexts and the staffing and exposure measures used means that estimates cannot be pooled or generalised. Considering a major potential source of bias in cross-sectional studies, we found evidence that estimates of nurse staffing effects from such studies may be overestimated when medical staff are omitted from the analysis. Our review of evidence of multidisciplinary staffing levels also revealed that other than nursing and medicine there is remarkably little evidence about the association between other staff groups such as allied health professions and patient safety, although the limited available evidence does indicate potentially important effects from such staff.

Chapter 3 Methods

We undertook a series of analyses to explore the association between healthcare staffing levels and outcomes using several data sources. Broadly, the study comprises two parts. In one, we use national routine data to explore associations at a hospital Trust level. The analyses are cross-sectional and are predominantly hospital Trust level with annual cross-sections, although we were able to analyse monthly associations between ward staffing and some nurse-sensitive quality indicators. In the second part, we use detailed patient and nurse roster data to explore longitudinal associations between staffing levels and outcomes in a sample of hospital Trusts. The full protocol for this study is available at www.fundingawards.nihr.ac.uk/award/NIHR128056.

Cross-sectional study (national, hospital Trust level)

This was a retrospective observational study using routinely available data on clinical healthcare staffing and quality. We included all 138 NHS hospital Trusts providing general acute adult inpatient services in England between April 2015 and March 2019. Hospital Trusts are defined as organisational units within the NHS that service a defined geographical area or that provide a specialised function. One Trust can therefore encompass several hospitals. This level of analysis was necessary because staffing data and most of the potential outcomes were only reported at a Trust level. We linked Trust-level data sets for workforce data, other Trust characteristics and outcomes for patients and staff from NHS digital and NHS England data collections using the unique hospital Trust ID. These data sets are openly available and accessible, along with the data dictionary.[115–117]

The unit of analysis was Trust per year. Trusts that reported at least 1 year of complete data were included. The number of hospital Trusts that were included varied by outcomes and also varied by year as organisations merged. In general, missing data were low. For example, for the mortality analysis, there were 540 observations from a total of 138 Trusts, of which 519 (96%) contained data on all variables and were therefore included in our analyses. Observations with missing data were removed from the analyses.

Trust and staffing data

Staffing variables are published monthly. We used full-time equivalent (FTE) values to calculate the annual average available staff at each hospital Trust from April to March the following year (the NHS financial year and the reporting period used for annual reports). One FTE is based on the standard contractual hours for the relevant staff group. Clinical staff are classified according to occupation codes. We grouped categories of clinical professional staff, merging smaller groups with related groups to create eight groups: medically qualified staff (subdivided by specialty: medical, surgical, other), nurses, support to nurses, allied healthcare professionals (AHPs), support to AHP, and scientific, therapeutic and technical (ST&T) staff. The nurses' group was composed of nurses in adult services only. The nurse support group included support to adult and general nurses, and nurses in training which would include trainee nursing associates and apprentices but not student nurses in placement on wards. The AHP group included occupational therapy, physiotherapy, speech and language therapy and other professionally qualified therapy staff. The ST&T group included pharmacists, and operating theatre staff among others. Because our core clinical outcome and the patient experience survey related to adult general hospital care, we excluded psychiatrists and paediatricians. See *Appendix 2* for more details and the full list of included groups. As these data represent employed staff as opposed to hours deployed, overtime and other hours worked outside contracted hours are not recorded in these FTE.

We calculated the average number of occupied (general) beds per Trust per year by averaging across the four quarters reported each year. Beds assigned to maternity, mental illness and learning disabilities services were excluded (as were staff in these areas). These general beds represented 96.2% of the total beds available across all hospital Trusts included in the study. The number of acute beds was used to calculate an occupied bed-per-staff ratio for each staff group per hospital Trust per year. We also used the number of beds to describe Trust size. Hospital Trusts in the upper tertile (i.e. the upper third of Trusts, ranked by size) were classified as large, while those in the lower tertile were deemed small Trusts, with the remaining classified as medium-sized Trusts. Trust teaching status was derived from the Estates Returns Information Collection,[118] which contains variables on the Trust profile.

Outcomes

For risk-adjusted mortality, we used the Summary Hospital-level Mortality Indicator (SHMI), which compares observed and expected deaths.[119] The SHMI includes all in-hospital deaths and those that occurred within 30 days of discharge from patients admitted to non-specialist acute Trusts. SHMI data are derived from the Hospital Episode Statistics (HES) at the level of provider spells (i.e. total continuous stay of a patient using a hospital bed at an NHS organisation, which can comprise several episodes under the care of different specialists), and the HES-Office of National Statistics linked mortality data. The latter captures deaths that occur outside of a hospital. Data used to calculate the SHMI are submitted by each hospital Trust.

Summary Hospital-level Mortality Indicator expected deaths are calculated based on individual patient characteristics that can affect the risk of mortality, including the patient's condition for hospitalisation, comorbidities,[120] age, gender and method of admission to hospital. Logistic regression models estimate the risk for each provider spell, with binary variable as outcome (i.e. died or survived). The SHMI model performs well with early validation demonstrating that it accounted for 81% of between-hospital variations with a *c*-statistic (area under the receiver operator curve) of 0.9.[121] The models are based on the preceding 3 years' data, with the last year's data used to calculate the current SHMI.[119] We used the data set that contains mortality data from April to March the following year to report on annual mortality rate for each hospital Trust included in this study to match the corresponding annual staffing data.

To explore patient experiences, we used results from the adult inpatient survey, undertaken every year in a sample of NHS patients. Sample sizes vary but are generally large with a fair response rate (e.g. 77,850 responses/44% response rate in 2016 and 76,915 responses/45% response rate in 2019). The survey was originally developed in consultation with NHS service users and others by the Picker Institute and covers a broad range of experiences. To reflect the contributions of all staff along the patient's pathway during admission, we used the single summary evaluation 'Overall, how was your experience while you were in the hospital?' rated from 0 (very poor) to 10 (very good), averaged at the hospital Trust level.

To explore staff experiences we used results from the NHS staff survey, which is conducted annually.[122] The survey is sent to all NHS staff with achieved samples in acute hospital Trusts in recent years ranging from 128,198 staff in 2015 (response rate 41%) to 258,011 in 2019 (response rate 44%). The survey was developed by the Picker Institute with items and themes based on existing measures and constructs with adaptations made by a steering group followed by cognitive testing of revised items.[123] Because of slight changes in items from year to year and because results for individual items are not weighted by staff groups, we focused our analysis on domain scores calculated for a range of themes:

1. equality, diversity and inclusion
2. health and well-being
3. immediate managers
4. morale
5. quality of care
6. safe environment – bullying and harassment
7. safe environment – violence
8. safety culture
9. staff engagement
10. team working

We scrutinised the contents of each theme and consulted with clinical experts, academics with an interest in workforce and patient/public contributors in our advisory group to select the themes that were likely to be directly influenced by and/or represent a possible staff outcome resulting from staffing levels. As a result of this consultation, consensus emerged that health and well-being, morale and quality of care were the outcomes most likely to be influenced by staffing levels (see *Patient, public and service engagement/involvement*). For the detailed content of each theme, see *Appendix 2*, *Table 31*. The scores are assigned based on positive outcome, so the most favourable response will be scored 10, while the worst will be scored 0 with intermediate items assigned scores with equal intervals. The theme score is the mean of response in that theme (subject to at least three items being scored). To make each Trust's scores comparable with other Trusts of the same type, individuals' scores within each Trust are weighted so that the occupational group profile of the organisation reflects that of a typical Trust of its type.

Analysis

We performed data linkage, cleaning, coding and analyses in R statistical software, version 4.0.2 (R Core Team. R: A Language and Environment for Statistical Computing. v4.0.2 ed. Vienna, Austria: R Foundation for Statistical Computing 2020), using the lme4 and glmer packages v1.1–27.1[124] and the plm package v2.4–3 [The plm Package – Linear Models and Tests for Panel Data (V2.4–3) (program). V2.4–3 version, 2021]. Continuous variables were described as mean and standard deviation (SD) or median and interquartile range (IQR), depending on the distribution of each variable. Categorical variables were described as frequencies and proportions. Crude annual mortality rates were calculated by dividing the number of observed deaths by the number of patient spells, for each hospital Trust each year.

To explore the relationship between staffing levels and outcomes, statistical modelling was conducted at the hospital Trust level using mixed-effects models on up to 4 years of data (2016–9). Outcomes were regressed onto hospital Trust staffing levels (expressed as the number of occupied beds per FTE staff). To account for variations in mortality risk related to case mix, we modelled observed deaths as a rate using expected deaths (as predicted from the SHMI model) as an offset in a multilevel negative binomial model.

Models were adjusted for Trust size and teaching hospital status with Trust included as a random effect to adjust for clustering. Trust size was included as a control variable as there may be economies of scale that provide efficiency savings for staff or additional clinical facilities which could impact outcomes. Teaching status was included because there may be systematic variations in care quality or staffing requirements linked to teaching hospitals that are otherwise unmeasured. For patient and staff experience models, we assessed whether including each Trust's crude mortality rate from the corresponding SHMI data set as an additional covariate improved model fit. For patient experience, we also considered the percentage of older patients (aged 66 years old and over) in the sample, to account for possible age-related differences in response patterns.

As well as simple random-effects models, we tested within-between random-effects models (WBRE), with Trusts as random effect, and staffing and hospital characteristics as fixed effects. These models can differentiate between effects arising from staffing differences between hospital Trusts and those that are associated with year-on-year changes of staffing levels within each hospital.[125,126] Where we considered a range of potential models and/or variables to include in models, model selection was based on minimising the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the likelihood ratios.

We assessed potential collinearity between covariates using correlation plots and Spearman's correlation coefficient, and the overall multicollinearity for all predictors in the models using generalised variance inflation factors (GVIFs), with a GVIF of more than 10 taken as a criterion to indicate substantial collinearity. As a sensitivity analysis, we reran the models excluding any variables that had a GVIF above 5[127] and noted any resulting substantive variation when reporting results.

Cross-sectional study (national, ward level)

While most national data available to us are reported at the hospital Trust level, there are additionally some nationally reported ward-level data for nurse staffing levels and quality. The care HPPD (CHPPD) data set for NHS hospital wards over the period of May 2016–March 2020 was supplied by NHS England. This data set reports an average value for each month (the sum of care hours for the month divided by the sum of midnight census patient counts) by staff group (RN and Unregistered Health Care Support Worker) for each ward in NHS Trusts. The data supplied for this analysis differ from that made publicly available (www.england.nhs.uk/care-hours-patient-day-chppd-data/) by including planned as well as actual CHPPD.

Quality measures were derived from harms identified in the NHS Safety Thermometer which is based on a monthly point prevalence survey of a range of potentially nurse-sensitive patient harms.[128] The data set from January 2016 to March 2019 was publicly available to be downloaded from NHS Digital. The Safety Thermometer data include patient age, location (Trust, ward and specialty), the presence/absence and type of harm. Harms included in the Safety Thermometer are PUs, falls, urinary catheterisation and urinary tract infections in catheterised patients and venous thromboembolism (see *Appendix 2* for details of harms and coding).

Data cleaning and linkage

The data were cleaned prior to linkage using basic rules – observations with zero values for planned RN staffing and those with missing identifiers or linkage variables were dropped.

The CHPPD data set contained data from 215 NHS Trusts with 257,159 observations over the date range from May 2016 to March 2020. Excluding observations with missing, invalid or unfeasible values for planned/actual CHPPD resulted in 256,775 observations, of which 207,780 were in acute hospital Trusts. From the Safety Thermometer data, we compiled monthly counts of harms, patients and patients over 70 years old (the only demographic data contained in the data). The resulting Safety Thermometer data set contains 316,068 observations over the date range from January 2016 to March 2019, of which 207,229 were in acute hospital Trusts. We then created aligned data sets in terms of common dates and Trusts. Trusts and dates were linked exactly, using the organisation code.

We then linked data by ward within each Trust. As there was no standardisation of ward naming between the two data sets, wards were linked probabilistically using the R fedmatch package,[129] using weighted Jaccard distance based on matching words in ward names and how frequently those words appear within the Trust (e.g. in the case of ‘Pembroke Ward’ the word ‘Ward’ is not very informative, whereas the name ‘Pembroke’ is). When matching was less than exact, we used matching on specialty as a verification criterion. The merged data set contains 96,050 observations, of which 40,064 were exact matches on ward name. Of the remaining 55,986 fuzzy matches, 32,181 matched on specialty. As a result, the final data set contained 72,245 observations. While most of the data show credible CHPPD, patient counts and harms for a single day, extreme values of 842 (for patient count) or 161 (for harms) suggest some sites may have reported cumulated figures of the full month. A pragmatic decision was taken to exclude rows where the patient count was greater than the 90th percentile (33 patients). Similarly, we excluded records where CHPPD exceeded 48 (i.e. a 2 : 1 staff-to-patient ratio).

Analysis

The relationship between staffing levels and incidents was assessed using a linear mixed model (to account for repeated observations), with the count of incidents (monthly) as the dependent variable, using the R lme4 package. Log of the number of patients was used as an offset. Covariates in the model were ward specialty, the proportion of patients over 70 years old, Trust type (specialist, teaching or other), Trust size (available beds) and region (London, South West, South East, Midlands, East of England, North West, North East and Yorkshire). Staffing was measured as the deviation between planned and actual staffing (shortfall), calculated as planned minus actual HPPD.

Longitudinal study (selected Trusts, patient level)

We recruited four NHS Trusts to participate in the retrospective longitudinal study where we linked daily ward-level staffing data to individual patient records and ward incident reports, as well as exploring the impacts of staffing levels experienced by staff members on subsequent sickness absence. The individual patient data came from the electronic care records of the four participating Trusts and was provided in an anonymised form with patient pseudonyms replacing identifiers. Similarly, roster data were provided with staff pseudonyms.

The Trusts were diverse in many respects including RN, medical and allied health professional staffing levels, size, teaching status and region (*Table 3*). Trust 1 was located in an urban area with a relatively high proportion of people born outside the UK (25% vs 17% national average) with 15% identifying as Asian ethnicity (the largest non-white ethnic group) in the 2021 census. A higher-than-average proportion was economically inactive (46% vs 39%), and a lower-than-average proportion was in higher managerial, administrative and professional occupations (9% vs 13%). Trust 2 was also in an urban area with a proportion born outside the UK of about the national average (17%) with 7% identifying as Asian ethnicity (the largest non-white ethnic group). Slightly fewer people than the national average were

TABLE 3 Contextual data for participating Trusts (2019)

	Teaching	Region	Beds (occupancy)[a]	Admissions[b]	RN/occupied bed[c]	Medics per occupied bed	AHP	Mortality rate(%)	SHMI (%)	Staff satisfaction (%)[d]	Patient satisfaction[e]
Trust 1	Y	Mid	1746 (87%)	139,775	2.02	1.11	0.44	3.2	103	56	7.91
Trust 2	N	SE	1056 (96%)	83,595	1.67	0.88	0.31	3.8	105	64	8.01
Trust 3	N	SW	523 (94%)	38,355	1.38	0.72	0.34	2.7	101	68	8.33
Trust 4	Y	Lon	993 (79%)	117,990	2.74	1.71	0.68	1.1	69	63	8.34
National median			753 (90%)[f]	66,440	1.64	0.87	0.44	3.3	101	63	7.91

a Source: Average daily number of available and occupied general acute beds open overnight April 2018–March 2019 NHS England: SDCS data collection – KH03.
b Spells contributing to SHMI March 2019 (April 2018–March 2019).
c Staff per bed based on annual workforce reports 2018–9 using groups defined in Cross-sectional study (national, hospital Trust level) above.
d Staff satisfaction survey 2019 – percentage of staff selecting Agree or Strongly Agree for q21c – I would recommend my organisation as a place to work. Benchmark group is acute and acute and community.
e 2019 Patient survey – average score overall rating/10.
f Median calculated from all Trusts classified as acute (small, medium, large) or acute teaching in NHS benchmark groups.

economically inactive (38% vs 39%) but a lower-than-average proportion were in higher managerial, administrative and professional occupations (10% vs 13%). Trust 3 was in a small town in a primarily affluent rural area. There were a lower-than-average proportion of people born outside the UK (11% vs 17%) and the largest non-white ethnic group (Asian) was only 3% of the population. In contrast with the previous two Trusts, the age profile included higher proportions of older people, a lower-than-average proportion were economically inactive (37.5% vs 39%), and a higher-than-average proportion were in higher managerial, administrative and professional occupations (15% vs 13%). Trust 4 was in Greater London, with the local area having an ethnically diverse population with a high proportion born outside the UK (45% vs 17%) and 18% identifying as of an Asian ethnic background. A lower-than-average proportion were economically inactive (37% vs 39%) and a much higher-than-average proportion were in higher managerial, administrative and professional occupations (24% vs 13%), although the area also included a higher-than-average proportion of those who had never worked or were long-term unemployed (10% vs 8.5%).

Sampling and data sources

Patient data were sourced from electronic care records and ward-based incident reports. Staffing data were sourced from electronic roster systems and electronic temporary (bank and agency) staffing records where these were held separately. We gathered data over the period April 2015–August 2020 for patients who met the inclusion criteria of having at least part of their admission as an overnight stay in an inpatient area, although discontinuities and differences in implementation in roster systems meant that we were unable to access data for the entire period from all Trusts. We also took a decision to discard data after March 2020 because the COVID-19 pandemic meant that patient flows and staffing patterns were highly atypical and emergency reorganisations of care provision meant that we could not be confident that nursing rosters accurately reflected staff deployment. See *Appendix 2*, data set preparation and cleaning (longitudinal study) for more detail.

Patient data

Patient data consisted of patient demographics, method of admission, patient movements between wards, clinical *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision (ICD-10) diagnoses codes and Office of Population Censuses and Surveys' Classification of Surgical Operations version-4 procedure codes. A code for the Health Care Resource Group (HRG version 4) was also supplied.[130] This is a summary code for clinical treatments provided during the admission. Nominal costs (tariffs) are associated with these. Three of the four Trusts also provided the first National Early Warning Score (NEWS)[131] for each admission from electronic care records. The demographics data provided included gender, age group (5-year grouping) and method of admission.

To account for patient-level risk factors, we used details in the patient record to calculate a predicted risk of mortality using the published SHMI model[132] based on diagnostic group, age and comorbidity according to the definitions in the SHMI specification version 1.30 issued 30 April 2019.[25] We used coefficients from this April 2019 model, which had been developed from national data for the previous 3 years (April 2016–March 2019), approximately coinciding with the study period. This approach allowed us to apply robust estimates of risk associated with a wide range of diagnostic groups (including the effect of interaction between age and comorbidity) derived from a national population, whereas directly adjusting for mortality risk using the same factors using only local data would inevitably mean that risks associated with some diagnoses could not be accurately estimated. As the risk of adverse outcomes has associated with particular times of year and days of the week, we included admission month in the data and derived indicator variables to identify weekend admissions and weekend stays.

The primary outcome for the study was death from all causes within 30 days of admission, where day 1 was the admission day and day 30 is included. This was determined either from the date of discharge when a discharge method of 'death' was recorded or from a date of death on patient records in the case of those discharged alive. Post-discharge dates of death were retrieved from a national patient demographic service and automatically incorporated into the patient records. The precision was by day rather than by exact time.

Length of stay was calculated from admission and discharge dates and times. Re-admission within 30 days of discharge was inferred by identifying a non-elective admission for the same patient to the same Trust by day 30 after discharge, where day 1 is the day following discharge and day 30 is included in the interval.

After considering nursing-sensitive outcomes from an umbrella review of research,[36] we identified a small number of adverse events and incidents that were identified as sensitive to nursing, and which could feasibly be identified in the available data (see *Appendix 2*, *Table 32*). We identified PUs, pneumonia and venous thromboembolisms from ICD-10 diagnosis coding in the patient record. The data set did not identify whether conditions were present on admission, so a subcohort of surgical patients for analysis was determined by admission specialty as the conditions were less likely to be present on admission in this group.

Records of patient safety incidents were extracted from the Trusts' risk management systems. These data were supplied in an anonymous form whereby a coded incident was recorded along with the ward and date, but with no details about the patient. From these incident reports we identified falls, medication errors and PUs. Incidents were also categorised by severity (no harm, harm, significant harm and severe harm).

Staffing data

Staffing data consisted of worked shifts and absences. These included details of the occupational group and pay banding along with the shift fulfilment source (substantive contract, bank or agency). Staff demographics were unavailable. Some data on other occupational groups such as physiotherapists, dietitians and speech therapists were provided but the electronic records were not complete and did not record time worked by ward and so could not be linked directly to patient admissions. Nursing staff were categorised as RN or NA according to their pay band, with RNs being a band 5 or above and those on bands 2–4 categorised as NAs.[133] We defined senior nurses as those on band 6 and above and senior NAs as those on band 4. Skill-mix was defined as the proportion of all nursing staff hours that were worked by RNs.

The primary variables of interest for the study are staffing levels and workload for the nursing team, measured in HPPD. Hours of care per study day in a ward was calculated by job type (RN or NA) from the worked shifts recorded for rostered staff (including temporary assignments to the ward) and from bank and agency records. Worked shifts exclude absences such as sickness and leave. To correspond as closely as possible with staff working patterns based on shifts, we defined two study periods corresponding to the day period of 7 a.m.–6:59 p.m. and the night period of 7 p.m.–6:59 a.m. and allocated hours to these periods. Each study day consisted of the concatenation of 1 day period with 1 night period; thus, a study day was a 24-hour period starting at 7 a.m. Worked shift hours minus break hours were apportioned to study days according to start and end times, with an assumption that all the break time was taken in the middle of the shift.

Patient occupancy per study day in a ward was calculated as patient-days from patient movement records, with one patient-day equivalent to one patient occupying one bed for 24 hours. HPPD for each job group was calculated per ward study-day (staff hours/patient days). Additionally, we derived a measure of workload associated with patient admissions to and discharges from a ward. We added the number of these movements together and divided by the staffing care hours to get a measure of ward throughput per RN hour and per NA hour.

As different wards have different staffing requirements, we calculated expected staffing levels to identify a threshold for low staffing. We had no consistent measure of staffing requirements, so we estimated the established (planned) staffing levels from the ward means. This approach has been used elsewhere in the absence of direct measurement of planned staffing requirements,[87] and our previous study found that ward means corresponded closely with the ward establishments (i.e. the planned number of staff per shift).[19] Where we saw changes in ward use or case mix for the ward over time, we divided the time series and calculated the expected staffing level for each period, treating the ward as (effectively) different wards with different expected staffing levels and workload associated with turnover (see *Appendix 2*).

We linked successive shifts for a staff member using a pseudonymised staff identifier. We identified planned shifts which were rostered but not worked and used the absence reason to identify sickness. We defined a sickness episode as starting on the first day a staff member was absent from work and ending when they next worked a shift. Variation in practices across Trusts meant that subsequent planned workdays missed through sickness could not be reliably identified, as standard weekday shifts were often substituted to record sickness for payroll purposes, so a 'return to work' date was inferred from the next worked shift after the start of the sickness episode. We estimated the sickness duration from these dates but limited the duration to a maximum of 365 days, as most employees are not eligible to receive any pay after 12 months' sickness absence. We excluded worked shifts and absence episodes not preceded by any worked shifts in the previous 7 days since they had no prior exposure.

Both worked shifts and sickness episodes for substantive staff were linked to staffing levels for shifts worked by the staff member in preceding week. The shift staffing levels were calculated in HPPD by linking each shift to ward occupancy records by location and date.

We explored the possibility of identifying important demographic factors related to staff for use in modelling associations between staffing levels and sickness absence (e.g. age) but data governance constraints proved insurmountable as personal data would be linked to working patterns and there is no clear framework to permit the use of these data in this manner without consent. Similarly, we were unable to reliably identify staff turnover as we did not have access to records of when staff joined or left the hospital.

Analysis

Patient outcomes

Staffing levels for each day of each patient's stay were linked to the patient records. The explanatory variables of interest were exposure to variation in staffing levels during admission. Based on previous research, we focused analysis on exposures that occurred during the first 5 days of the hospital stay, accounting for most of the stay for most patients, although we assessed the sensitivity of results to staffing experienced over longer and shorter time periods.

Since detrimental understaffing effects may accumulate over time, we modelled staffing levels with cumulative time-dependent covariates in survival models for mortality and re-admission. We constructed repeated observations, one per day, on the same patient over a period of time from the admission (i.e. onset of risk) until death or censoring at 30 days. We varied this approach where appropriate for other outcomes (see *Modelling approach* below). We modelled patient staffing exposures both as days of low staffing and net deviations from norms over time. Separate variables were used for RN and NA staffing levels.

Understaffed days were identified as an observed HPPD which fell short of the expected HPPD. We used the mean as the threshold for our main analyses, but we also tested alternative thresholds both above and below the mean. Cumulative shortfalls were calculated in several different ways. We accumulated counts of days where staffing was low (below the chosen threshold). Net deviation was calculated as the difference between the cumulative sum of expected staffing levels in HPPD and the cumulative sum of observed HPPD. A positive value indicated net understaffing. An accumulated one-sided difference between the expected and observed HPPD was also tested, so a patient day where the observed level was at or above the expected level contributed zero to the sum. If the patient spent time on two or more different wards on 1 day, then a simple mean, or weighted mean by occupancy time, of the 2 part-days of the ward staffing or workload measure was used for that day.

In addition to staffing levels, we considered other variables to account for patient turnover, seniority within each job type, and proportions of bank and agency care hours of total hours. These were added to models as time-varying covariates to determine the contribution of these additional factors. Workload from high patient turnover was also considered as a cumulative time-varying covariate. All models included ward as a random effect to account for the hierarchical nature of the data and nonspecific ward characteristics. If a patient day included a ward transfer, the ward at the start of the day was used as the clustering ward. Other variables were non-time varying (patient risk, admission month, weekend admission/stay). Patient case-mix adjustment was undertaken using the SHMI risk score. As we had an admission NEWS score for a subsample, we tested the impact of this direct measure of acuity on admission in a sensitivity analysis.

We explored the data to determine evidence of nonlinear effects by adding cubic and quadratic terms for staffing variables and explored whether the effect of one staff group is contingent on the level of another – specifically exploring the hypothesis that adding assistant staff will improve the relative effectiveness of RNs – by adding interaction terms.

Staff outcomes

We calculated the following summary exposures to staffing levels and configurations for each staff member from their shifts in the 7 days prior to the start of each sickness absence episode, and prior to each worked shift as a control.

- The mean shortfall (one-sided) of observed to expected HPPD staffing as a proportion of the expected staffing.
- The proportion of RN/NA hours that were provided by bank staff.
- The proportion of RN/NA hours that were provided by agency staff.
- Mean skill-mix, calculated as the proportion of RN hours over all nursing hours.
- The proportion of long shifts (i.e. 12-hours or more), which the subject had worked in the week prior.

Shift pattern variables were included in our analyses as they have been found to affect sickness absence in previous studies.[134] We also inferred whether the staff member was working part-time by checking if their median weekly worked hours in the previous 13 weeks was ≤ 26 hours excluding breaks. We used worked shifts for substantive staff as controls for sickness episodes.

Modelling approach

We used Cox mixed-effect survival models extended to handle time-varying covariates for mortality and re-admission outcomes.[135] Hazard ratios (HRs) and their 95% CIs were produced along with p-values from two-sided Wald tests. The assumption of proportional hazards (PH) was examined with graphs of scaled Schoenfeld residuals and was seen to be reasonable for the staffing measures. We explored several alternative parametric modelling frameworks including exponential and Weibull distributions, but models ran considerably slower than the Cox models, some taking days to run and often failing to converge, though where we were able to obtain estimates these models generally gave similar results. We considered alternative frameworks including pooled logistic and discrete time survival models which also took into account time-varying covariates, and logistic models with summary measures of staffing (e.g. proportion of days understaffed) – neither improved on the precision of the Cox models, nor produced significantly different results. As expected, the results closely approximated those of the Cox PH models, but this confirmed that our conclusions were unlikely to be altered by any particular assumptions about the survival function.[136,137]

For survival models, we set the time origin to be the day after admission, requiring an overnight stay before being 'recruited' to the analysis. Patients discharged on the same day as their admission were thus excluded but for remaining patients, exposure levels on the admission day were retained as part of the analysis. This method served to exclude day attenders at low risk and leads to mortality models with better discrimination and was thought to be appropriate when investigating the effect of staffing level exposures on inpatients. Although it did have the unintended effect of excluding those who died early in their hospital stay, emergency department staffing, a likely major influence on adverse events happening on the admission day, was outside the scope of the study.

Re-admissions were modelled in a very similar fashion to mortality. The main analysis was carried out for admissions with discharges but excluding admissions which ended with death. A sensitivity analysis was done with all admissions, ending with discharge or death.

Length of stay measured in hours is a continuous outcome which exhibits a right-skewed distribution with mode near zero and heavy tails. Therefore, to better represent LOS data features, the gamma distribution was used in a generalised linear regression model using summary measures of staffing and the discharge ward as a random effect. The staffing measure was the proportion of understaffed days for RNs and NAs during the exposure period which was the first 5 days of the hospital stay (or the entire stay if < 5).

Potentially nurse-sensitive adverse events were identified from diagnosis codes, but precise dates and times were not available. Therefore, survival analysis was not appropriate as we did not have a time for the event. We used logistic regression with a random effect for the admission ward. As with the LOS analysis, a summary of the staffing exposure for each job type as the proportion of days during the 5-day exposure period which were understaffed was used.

As incident reports were linked to ward days as opposed to individual patients, we modelled counts of incidents as the outcome using Poisson regression with daily ward occupancy as the exposure offset. We assessed the association between nurse staffing levels and the number of incidents occurring on that day. Because tissue damage can take time to become apparent, we also looked at lagged associations, using staffing on the previous 3 days. All incident models included control for the overall acuity of patients on the ward by including the mean predicted SHMI mortality risk for patients on the ward on that day.

We explored the association between staffing and sickness absence episodes with logistic mixed-effect models, where the event was the start of a sickness episode (first unworked shift), and the controls were worked shifts. Ward and staff member were included as random effects to account for clustering of observations within individuals and within groups of staff in wards. Because sickness absence levels differ substantially for RNs and NAs,[134] we modelled them separately.

Missing values

As the data sources were critical operational systems, there were few missing values recorded, although there were small numbers of admissions without coded diagnoses. These were omitted from regression models since diagnoses were needed for the patient adjustment. Missingness also arose from failures to match or align staffing data by location with patient data. This situation occurs when, for example, wards as organisational units change physical locations, and a mismatch arises between codes used in patient and roster systems. Patient days without any matched RN staffing were excluded, although patient days without any NA staffing were retained as it is plausible that a ward shift might operate without NAs. The previous day's staffing values were carried forward in the case of missing values.

Model fit

When considering alternative models in the context of survival and generalised linear mixed models, the approach we adopted was to balance model fit and parsimony based on information criteria obtained through maximum log-likelihood. As we built from our core model (patient risk and staffing exposure), we considered additional factors and judged the extent to which the model was improved. We used the AIC and the BIC to assess model fit, preferring models that minimise the values of both.[138] We considered nonlinear effects before interactions as mis-specified linear main effects can generate spurious interactions.[139]

Because there are structural factors that largely operate at a hospital level, including levels of medical and other therapy staff, we initially modelled each Trust separately. Results were broadly similar. We found that models with effects for both Trust and ward had negligible difference in AIC (0.4) and BIC (1.2) from models with just ward effects, while providing almost exactly the same coefficients in our core models. We concluded that any hospital-level effects were properly accounted for by the ward random effect and so we omitted Trust from the models to reduce the computational intensity.

Subgroup analyses

We undertook subgroup analyses considering particular patient groups (e.g. older patients, groups with high mortality risk) and groups exposed to staffing in particular types of wards such as critical care and high care. We analysed paediatric admissions as a separate cohort comprising those who stayed in a paediatric ward and whose 5-year age band was no higher than the 15-to-19 band.

Analysis software

Data set construction and analysis were undertaken in R 4.3.0 (R Core Team. R: A Language and Environment for Statistical Computing. v4.3.0 ed. Vienna, Austria: R Foundation for Statistical Computing 2023) using the Rstudio development environment (Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software 2023). The tidyverse 2.0.0 package was used for data wrangling.[140] The survival analysis data set used to model mortality and re-admission was constructed with the survival 3.5–5 package[141] and analysed using the coxme 2.2–18.1 package.[142] The LOS gamma regression modelling, and patient conditions and patient incident logistic modelling was undertaken with the lme4 1.1–33 generalised linear model package.[124] Descriptive statistics used the finalfit 1.0.6 package wrapper functions.[143] The rmarkdown 2.21 package was used to format output in the preparation of this report.[144]

Economic analysis

We used the results of these analyses to model the costs and consequences of changes in staffing levels required to avoid short staffing, estimating the incremental cost-effectiveness associated with change. Specifically, we used the observed staffing shortfall averaged over the study period and applied relative risks estimated in the statistical models to observed baseline data to estimate additional staffing requirements, the expected impact on patient outcomes, nurse staff sickness absence, re-admission and other relevant effects. The cost perspective is that of the hospital and so costs relate to the index admission and re-admissions arising shortly after (30 days post discharge). Although our mortality outcome is measured up to 30 days after admission, we take a lifetime perspective by extrapolating quality-adjusted life expectancies.

The quantity of staff hours required to eliminate staffing shortfalls was estimated from the count of patient days in the study, the proportion of days where staffing was below the mean and the average shortfall (in HPPD for RNs and for support staff). The costs of staffing to avoid these shortfalls were estimated, as were the costs associated with any differences in LOS and re-admissions. We adopted an NHS hospital provider perspective for costs, using national (NHS) reference costs for hospital stay[145] and nationally representative unit costs for substantive staff.[146] Additional charges were included to account for costs associated with employing agency staff. A summary of staffing unit costs used in the analysis, including upper and lower estimates used for sensitivity analysis, are shown in *Appendix 2*, *Table 33*. Costs of re-admissions were based on the distribution of HRG codes[130] for re-admissions observed in study patients in combination with the national reference cost. Similarly, baseline cost for the initial admission was estimated using the distribution of HRG codes in the original cohort. Costs of change in LOS associated with staffing change were based on the estimated change in LOS and valued using the NHS reference costs/tariff for excess bed days.

We extrapolated the effects of reducing/avoiding shortfalls on patient mortality and re-admission using HRs estimated in the statistical models. We assumed that the HR approximated to a risk ratio, an assumption that is generally valid in the short run and when probabilities of events are small, which is the case with mortality in this study.[147,148] We applied the risks over the average patient exposure to days of low staffing levels during the first 5 days observed during the study period. For each staffing group, the mortality effect was estimated as: $1 - [(|1 - \mathrm{HR}|) \times \text{Mean Exposure}]$. The combined (multiplicative) effect for each staff group was applied to the observed mortality rate. The difference between the observed and modelled death rate was applied to the total number of deaths observed in the study to estimate the number of deaths averted by avoiding staffing shortfalls. A similar approach was adopted for modelling the effect on re-admissions.

To facilitate comparison with other investments, we estimated the quality-adjusted life-years that might be expected for each life saved/gained. Our approach was based on the discounted and quality-adjusted life expectancy (DANQALE) approach adopted in previous evaluations of policy interventions.[149] This approach combines age/sex-specific life expectancy estimates with quality-of-life valuations on the utility scale (anchor points being zero for death and one for perfect health) and discounts future life-years by 3.5% per year (as specified in UK government guidance for policy evaluation),[150] so that years of life in the distant future have less value than current years. This is more conservative than the widely recognised ‘Gates reference case’ which recommends 3% for global health evaluations,[151] in the sense that larger discount rates lead to reduced estimates of future quality-of-life gains.

We used Office for National Statistics life tables to estimate age- and sex-specific life expectancy.[152] Since the data supplied reported patient age groups, rather than a specific age (due to issues of confidentiality and identifiability), we calculated a sex-specific weighted average mortality rate for each age band. This was based on the population distribution in each age within each age band using the LifeTable() function from the R library MortalityLaws to create sex-specific abridged life tables. As those who are admitted to hospital are likely to be less healthy than the general population, we applied age-specific health state valuations for people with at least one existing health condition,[153] derived from the Health Survey for England (see *Appendix 2*, *Table 34*).

We compared the observed data with the estimated costs and consequences of eliminating low staffing, testing a range of scenarios focusing on full or partial elimination of low staffing and targeting specific patient groups. We calculated incremental cost-effectiveness ratios for each scenario, using cost per life saved and cost per QALY gained.

In the base-case analysis, we estimated QALY expectations for the population of patients who die. In the base-case scenario, we assumed that if those patients were to survive, the age-group/sex-specific life expectancy would be the same as that of a similar individual with an existing condition in the general population – we tested this assumption by assuming a higher mortality expectation (hence shorter life expectancy) for hospitalised patients. Additional sensitivity analyses were conducted to determine the robustness of results with respect to other sources of uncertainty in the QALY calculation including statistical uncertainty in the life expectancy estimates, selection of quality-of-life weights and choice of discount rate.

Ethics and approvals

The analyses of national routine data involved aggregate data that were publicly available and so no specific ethical approval was required. We used the screening questions for the Health Research Authority integrated research application service to determine the required approvals and determined that the study was exempt from the requirement for approval by the National Research Ethics Service. Approval was sought and granted to undertake the research in NHS Trusts (IRAS ref 273185). Additionally, as a requirement of the University of Southampton, we sought ethical review from the University's system and were granted approval (ERGO 52957).

Although this is a non-interventional study using anonymous data with no realistic opportunity to reidentify individuals or to link to other data sources to do so, we are aware that there are matters of significant public concern over the use of data sourced from health records. These extend beyond the remit of data protection regulation, which cover personal data where a living person could be identified. Consequently, we used opportunities for patient and public engagement to share openly and in detail the nature of the data we had and to explore issues of concern that arose (see *Patient, public and service engagement/involvement* and *Report Supplementary Material 1*).

Patient, public and service engagement/involvement

Francesca Lambert (FL) is a lay co-investigator who contributed to the development of this project from the outset. Francesca as a parent of four, including adult twins who have learning disabilities and health problems, has engaged as a long-standing collaborator both as a someone who accesses health services (including acute care) and as an expert in linking patient and public involvement to research. Francesca has worked in both NHS and university research delivering patient and public involvement and engagement support and activities with many different patient and public groups and translating research for the lay audience. Her work informed the priority setting and engagement that preceded the development of this project and the initial engagement and dissemination strategy. Findings from these engagements included recognition of how important proper nurse staffing levels are for people's confidence and trust in the NHS, the support for research to determine the right level and mix of staff but also people's shock that this is a 'safety' issue.

> *…. patients don't want to think that you're just about 'safe', they want more than that … I don't think that that term is right because it doesn't mean anything to the public … their expectation is that of course care is safe … I think patients feel that that's a given.*[2]
>
> *p. 86*

This project was developed with this concern in mind and was shaped by conversations and meetings with members of the patient and public involvement group attached to our Workforce Theme at the University of Southampton. At a joint workshop held with staff, patients and members of the public (23/39 patient and public participants), nurse staffing were voted the top ranking priority.[154] and further workshops helped to refine questions and informed our decision to include patient satisfaction as an outcome in our panel study.

Because of the nature of this study – focusing on workforce requirements and provision, and in line with advice received from our patient and public co-applicant – we took a wide view of the potential beneficiaries of this research. Applying the widely quoted maxim '*nothing about me without me*',[155] we included NHS staff as potential beneficiaries of research but we also sought engagement from a public perspective as well as specifically those with patient experience. One result of this was the appointment of a local councillor as a lay member of our study steering group.

The study began concurrently with the beginning of the COVID-19 pandemic, and this impacted on many of our plans for ongoing involvement and engagement. Planned events involving NHS staff had to be cancelled as COVID-19-related work had to be prioritised and early public events had to be conducted online. Despite this, we have received significant ongoing engagement from the Trusts that volunteered to provide data for the study and many members of staff worked hard to progress approvals for a study that would not otherwise have been prioritised because it was recognised as an important for them. Further details of activities undertaken during the project are given in *Patient, public and service engagement/involvement* and *Report Supplementary Material 1*.

Chapter 4 National cross-sectional study results

We included a total of 138 hospital Trusts providing general acute adult inpatient services in England between April 2015 and March 2019. Of these, 41 (29.7%) of them were classified as teaching Trusts. The median number of acute general occupied beds was 674 (IQR 486–912). There was considerable variation in staffing levels between Trusts. Beds per RN (mean 0.61 beds per FTE) showed least relative variation (SD 18.0% of the mean). AHPs and AHP support showed most variation (SD 37.7% and 44.4% of the mean). Most variation was between Trusts and within-Trust variation was relatively small, with SD ranging from 4.9% of the mean for RN staffing to 10.9% for AHP support (*Table 4*).

Registered nurse staffing levels were strongly correlated with staffing levels in all doctors' groups ($\rho > 0.71$). Nurse support staff levels were positively correlated to RN ($\rho = 0.30$), as were AHP and AHP support ($\rho = 0.47$).

Hospital mortality

Our final sample consisted of 519 observations of staffing, bed occupancy and mortality variables for the 4-year study period, clustered within 138 hospital Trusts. Annual mortality rates across all hospital Trusts were 3.44% for 2015–6, 3.46% for 2016–7, 3.23% for 2017–8 and 3.31% for 2018–9. We first estimated models including a single staff group only (controlling for Trust factors). In these models, hospital Trusts with more occupied beds per FTE staff (i.e. lower staffing levels) were associated with a higher rate of death relative to the expected rate for all professionally qualified staff groups. The opposite effect was observed for support staff, where hospital Trusts with more beds per nurse support and AHP support staff staffing had lower mortality (*Table 5*).

In most cases, estimated effects were reduced and some were no longer statistically significant when we included all staff groups in the model simultaneously. In the multistaff group model, hospitals with more beds per FTE medical doctor (1.04, 95% CI 1.02 to 1.06), other medical specialties (1.03, 95% CI 1.00 to 1.06) and AHP staff (1.04, 95% CI 1.02 to 1.06) were associated with higher death rates ($p < 0.05$). The associations with beds per surgical doctor and RN were not statistically significant, although the observed effect for RN remained relatively large (1.07). More beds per FTE nurse support (0.85, 95% CI 0.79 to 0.91), and AHP support (1.00, 95% CI 0.99 to 1.00), that is lower staffing levels from these groups, were associated with lower death rates ($p < 0.05$) (see *Table 5*). Despite potentially providing additional information by separately estimating effects due to differences between Trusts and staffing changes within Trust, the WBRE models provided considerably worse fit (e.g. AIC for random-effects model was 250 vs 6418 for the WBRE model). Between-Trust effect estimates from these models were generally the same as in the simple random-effect models while within-Trust effects were not statistically significant.

Generalised variance inflation factor for all variables was < 10 which is generally accepted as indicative of no evidence of multicollinearity. However, the GVIF for beds per other medical specialties was 6.1, so we ran the model omitting this variable as a sensitivity analysis and found similar estimates for the remaining coefficients, with no changes to substantive conclusions based on statistical significance.

Patient experience

Across the 4 years we linked a total of 583 responses from 149 Trusts. The mean response to the rating of overall experience was 8.15/10 with a SD of 0.32. Trust average responses ranged from 7.35 to 9.23. While more beds per AHP and scientific and technical staff were associated with lower patient experience scores in the single staff group models, only beds per RN was associated were a statistically significant reduction in experience scores in the multistaff group models (*Table 6*). Other associations with staffing were small and not statistically significant.

Adding the proportion of older people worsened the model fit (AIC and BIC increased, coefficient = 0). Including the Trusts' mortality rates in the model improved fit and so this variable was retained, although the relationship was not statistically significant in the multistaff group model. WBRE models provided considerably worse model fit.

TABLE 4 Occupied bed-per-staff across 138 hospital Trusts (2016–9)

	Beds per FTE staff							
Staff group	**Q1**	**Median**	**Q3**	**Mean**	**SD**	**SD as % of mean**	**Within-trust SD (mean)**	**Within-trust SD as % of mean**
Doctors (medical)	3.05	3.66	4.30	3.70	0.96	25.9	0.28	7.6
Doctors (surgical)	3.10	3.79	4.28	3.73	0.81	21.7	0.22	5.9
Doctors (other medical specialties)	3.26	3.96	4.69	3.94	1.01	25.6	0.27	6.9
RN	0.54	0.61	0.68	0.61	0.11	18.0	0.03	4.9
Nurse support	0.80	0.93	1.08	0.93	0.21	22.6	0.07	7.5
AHP	1.88	2.25	2.74	2.36	0.89	37.7	0.18	7.6
AHP support	7.91	10.03	13.01	11.07	4.91	44.4	1.21	10.9
ST&T	1.80	2.24	2.82	2.31	0.75	32.5	0.16	6.9

Q1, lower quartile; Q3, upper quartile.

TABLE 5 Associations between beds per FTE staff and mortality from the negative binomial random-effect models, adjusted for hospital Trust characteristics

Predictors	**Estimates for single staff group models[a] (rate ratio)**	**95% CI**	**Estimates for multiple staff group model (rate ratio)**	**95% CI**
Bed per medical doctor	1.05***	1.04 to 1.07	1.04***	1.02 to 1.06
Bed per other medical specialties	1.05***	1.03 to 1.06	1.03*	1.00 to 1.06
Bed per surgical doctor	1.04***	1.02 to 1.06	0.98	0.96 to 1.01
Bed per RN	1.33***	1.15 to 1.54	1.07	0.88 to 1.31
Bed per nurse support	0.92*	0.86 to 0.98	0.85***	0.79 to 0.91
Bed per AHP	1.02*	1.00 to 1.04	1.04***	1.02 to 1.06
Bed per AHP support	0.99*	0.99 to 0.99	1.00**	0.99 to 1.00
Bed per ST&T	1.02**	1.00 to 1.04	0.99	0.98 to 1.05
Teaching Trust	0.96*	0.92 to 0.99	1.01	0.98 to 1.05
Small Trust (ref)				
Medium Trust	1.01	0.97 to 1.05	1.01	0.98 to 1.04
Large Trust	1	0.96 to 1.04	1.01	1.98 to 1.05

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

[a] Single staff group models include one staff variable + Trust characteristics + random effect for Trust. Coefficients reported for Trust characteristics are from models with no staff groups.

Staff experience

One hundred and twenty-one Trusts in total provided returns for the staff satisfaction survey with 468 responses across the 4 years (241 for morale, as this was first introduced in 2018). The mean score (out of 10) for health and well-being was 6.0 (SD 0.28), for morale was 6.1 (SD 0.25) and for quality of care was 6.5 (SD 0.21).

Models exploring the relationship between single staff groups and staff experience outcomes showed several statistically significant associations with more beds per staff member associated with lower staff experience scores, especially in relation to quality of care. However, most of these associations were reduced and were not statistically significant in models that included all staff groups simultaneously. Hospitals with more beds per RN scored lower in terms of staff health and well-being and quality of care while hospitals with more beds per nursing support scored lower in terms of staff morale. Hospitals with more beds per surgical doctor (i.e. relatively lower staffing by surgeons) scored higher in terms of staff morale. Other associations with staffing were small and not statistically significant (*Table 7*). There was no evidence of collinearity in our models (GVIF < 5).

For morale and quality of care variables, the simple random-effects mixed models were clearly preferable, giving the smallest AIC and BIC. For staff health and well-being, the WBRE was marginally preferable according to AIC (−199.5 vs −195.1) but not the BIC (−104.7 vs −133.3). The statistically significant association between health and well-being scores and bed per RN was identified as a within-hospital effect, arising from changes in beds per nurse within hospitals (*Table 8*). This model also identified a significant association between hospitals with more beds per surgeon and higher health and well-being scores.

Ward-level associations between nurse staffing and harms (national data)

Figure 2 plots planned (*x*-axis) versus actual (*y*-axis) CHPPD at ward level for acute hospitals in the CHPPD data set. The figure also shows the 45° line where actual equals planned – the plot suggests that actual staffing is below the plan for the majority of observations.

The median number of patients was 21, and the median number of harms identified was 1 per ward month. In the actual staffing model, after controlling for the ward specialty, the proportion of patients over 70 years old and Trust characteristics, there was a small but statistically significant reduction in the rate of harms recorded on wards with more registered nursing hours per patient [incidence rate ratio (IRR) 0.99, 95% CI 0.99 to 1.00]. Most specialties were associated with significant differences (from the mean) incident rate with paediatric specialties generally having lower incident rates. Larger and more specialist Trusts had lower rates of harm. Models using the shortfall in hours resulted in a model with worse model fit (increased AIC/BIC). In these models, coefficients for most variables were largely unaltered, but there was no statistically significant effect from RN staffing and an adverse effect from NA staffing, that is the larger the assistant shortfall, the lower the incident rate (IRR 0.98, 95% CI 0.97 to 1.00). See *Table 9* for details.

Summary of main results

- We included 138 hospital Trusts. The number of beds per FTE staff member varied considerably between Trusts.
- The largest variation was observed in AHP staff (mean 2.4 beds per FTE) and support to AHP staff (mean 11.1 beds per FTE), where the SD was 38% and 44% of the mean, respectively.
- RN staffing levels were strongly correlated with staffing levels by doctors ($\rho > 0.71$).
- Although the number of beds per RN had the largest effect on mortality in models including single staff groups [rate ratio (RR) 1.33, 95% CI 1.15 to 1.54], this was greatly reduced and no longer statistically significant when all staff groups were included in the model, although it remained the largest effect size (RR 1.07, 95% CI 0.88 to 1.31).
- In multiprofessional models, more occupied beds per AHP (RR 1.04, 95% CI 1.02 to 1.06) and per medical doctor (RR 1.04, 95% CI 1.02 to 1.06) were associated with increased risk of death.
- More beds per nurse support (RR 0.85, 95% CI 0.79 to 0.91), and AHP support (1.00, 95% CI 0.99 to 1.00) were associated with lower death rates.

TABLE 6 Associations between beds per FTE staff and patient experience, from random-effect models, adjusted for hospital Trust characteristics

Predictors	Estimates for single staff group models[a]	95% CI	Estimates for multiple staff group model	95% CI
Bed per medical doctor	0	−0.03 to 0.03	0	−0.04 to 0.05
Bed per other medical specialties	0.01	−0.02 to 0.04	0.03	−0.03 to 0.08
Bed per surgical doctor	0.02	−0.01 to 0.05	0.07	0.01 to 0.14
Bed per RN	−0.11	−0.33 to 0.10	−0.39*	−0.76 to −0.02
Bed per nurse support	−0.09	−0.22 to 0.03	−0.02	−0.20 to 0.16
Bed per AHP	−0.04*	−0.08 to −0.01	−0.04	−0.09 to 0.01
Bed per AHP support	0	−0.01 to 0.00	0	−0.01 to 0.01
Bed per ST&T	−0.02*	−0.05 to −0.00	−0.02	−0.05 to 0.01
Mortality rate	4.71*	0.52 to 8.89	3.34	−1.34 to 8.02
Teaching Trust	0.06	−0.02 to 0.14	0.09	−0.00 to 0.18
Small Trust (ref)				
Medium trust	0.02	−0.05 to 0.09	−0.01	−0.09 to 0.06
Large trust	0.01	−0.06 to 0.09	−0.04	−0.13 to 0.05

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

[a] Single staff group models include one staff variable + % respondents over 65 years old (not reported effect size = 0) + mortality rate + Trust characteristics + random effect for Trust. Coefficients reported for Trust characteristics are from models with no staff groups.

TABLE 7 Associations between beds per FTE staff and staff experience from the random-effect mixed models, adjusted for hospital characteristics (single and multivariable multiple staff group models)

	Health and well-being		Morale		Quality of care	
	Estimate	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**
Single staff group models[a]						
Bed per medical doctor	−0.03	−0.06 to 0.01	0.01	−0.04 to 0.06	−0.06***	−0.09 to −0.03
Bed per other medical specialties	−0.04*	−0.08 to −0.00	−0.01	−0.06 to 0.04	−0.07***	−0.10 to −0.04
Bed per surgical doctor	−0.05*	−0.09 to −0.00	0.02	−0.04 to 0.08	−0.08***	−0.11 to −0.05
Bed per RN	−0.68***	−0.94 to −0.43	−0.19	−0.60 to 0.23	−0.75***	−0.94 to −0.57
Bed per nurse support	−0.21**	−0.37 to −0.06	−0.5***	−0.70 to −0.31	−0.28***	−0.40 to −0.16

TABLE 7 Associations between beds per FTE staff and staff experience from the random-effect mixed models, adjusted for hospital characteristics (single and multivariable multiple staff group models) (*continued*)

	Health and well-being		Morale		Quality of care	
	Estimate	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**
Bed per AHP	−0.06*	−0.10 to −0.01	−0.09***	−0.14 to −0.04	−0.08***	−0.11 to −0.04
Bed per AHP support	0	−0.01 to 0.00	−0.01***	−0.02 to −0.01	0	−0.01 to 0.00
Bed per ST&T	−0.01	−0.04 to 0.01	−0.07***	−0.10 to −0.03	−0.04***	−0.06 to −0.02
Mortality rate	8.49***	3.28 to 13.69	1.22	−4.36 to 6.80	−4.72*	−8.58 to −0.85
Teaching Trust	−0.07	−0.17 to 0.03	−0.04	−0.13 to 0.06	0.01	−0.06 to 0.08
Small Trust (ref)						
Medium trust	0	−0.09 to 0.09	0.01	−0.09 to 0.10	0.02	−0.05 to 0.09
Large trust	−0.06	−0.15 to 0.04	−0.07	−0.17 to 0.03	0	−0.07 to 0.08
Multiple staff group models						
Bed per medical doctor	0.02	−0.04 to 0.08	0.02	−0.04 to 0.09	0.01	−0.03 to 0.06
Bed per other medical specialties	0.03	−0.04 to 0.10	−0.05	−0.14 to 0.04	0.02	−0.04 to 0.07
Bed per surgical doctor	0.08	−0.00 to 0.16	0.10*	0.02 to 0.19	0.02	−0.04 to 0.08
Bed per RN	−1.30***	−1.76 to −0.84	0.10	−0.50 to 0.69	−0.92***	−1.26 to −0.58
Bed per nurse support	−0.06	−0.28 to 0.17	−0.45***	−0.68 to −0.22	−0.02	−0.18 to 0.15
Bed per AHP	−0.02	−0.08 to 0.05	−0.01	−0.08 to 0.06	−0.01	−0.06 to 0.04
Bed per AHP support	−0.00	−0.01 to 0.01	−0.01	−0.02 to 0.00	0.00	−0.00 to 0.01
Bed per ST&T	0.03	−0.01 to 0.07	−0.03	−0.07 to 0.02	−0.00	−0.03 to 0.02
Mortality rate	4.74	−1.14 to 10.62	−3.50	−9.86 to 2.86	−5.46*	−9.78 to −1.15
Teaching Trust	−0.04	−0.16 to 0.09	−0.00	−0.12 to 0.11	−0.08	−0.17 to 0.01
Small Trust (ref)						
Medium trust	−0.04	−0.13 to 0.06	0.04	−0.05 to 0.14	0.00	−0.07 to 0.07
Large trust	−0.11	−0.22 to 0.01	−0.05	−0.16 to 0.06	−0.04	−0.12 to 0.04

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

a Single staff group models include one staff variable + Trust characteristics + random effect for Trust. Coefficients reported for Trust characteristics are from models with no staff groups.

TABLE 8 Associations between beds per FTE staff and staff-reported health and well-being theme score, within between random-effects model (adjusted for hospital characteristics[a])

	Effect	Estimate	95% CI	*p*
Bed per medical doctor	w	0.02	−0.05 to 0.10	0.536
	b	0.01	−0.09 to 0.10	0.877
Bed per other medical specialties	w	0.07	−0.02 to 0.17	0.120
	b	−0.07	−0.19 to 0.05	0.247
Bed per surgical doctor	w	0.01	−0.11 to 0.14	0.815
	b	0.13*	0.02 to 0.24	0.017
Bed per RN	w	−1.64**	−2.19 to −1.10	< 0.001
	b	−0.01	−0.87 to 0.86	0.986
Bed per nurse support	w	0.15	−0.18 to 0.47	0.384
	b	−0.22	−0.53 to 0.10	0.177
Bed per AHP	w	−0.03	−0.14 to 0.07	0.506
	b	−0.02	−0.10 to 0.06	0.640
Bed per AHP support	w	0.00	−0.01 to 0.01	0.428
	b	−0.00	−0.01 to 0.01	0.704
Bed per ST&T	w	0.02	−0.03 to 0.07	0.344
	b	0.02	−0.04 to 0.08	0.488

*$p < 0.05$, **$p < 0.001$.
b, between-effect; w, within-effect.
a Model includes control for Trust mortality rate, teaching and Trust size.

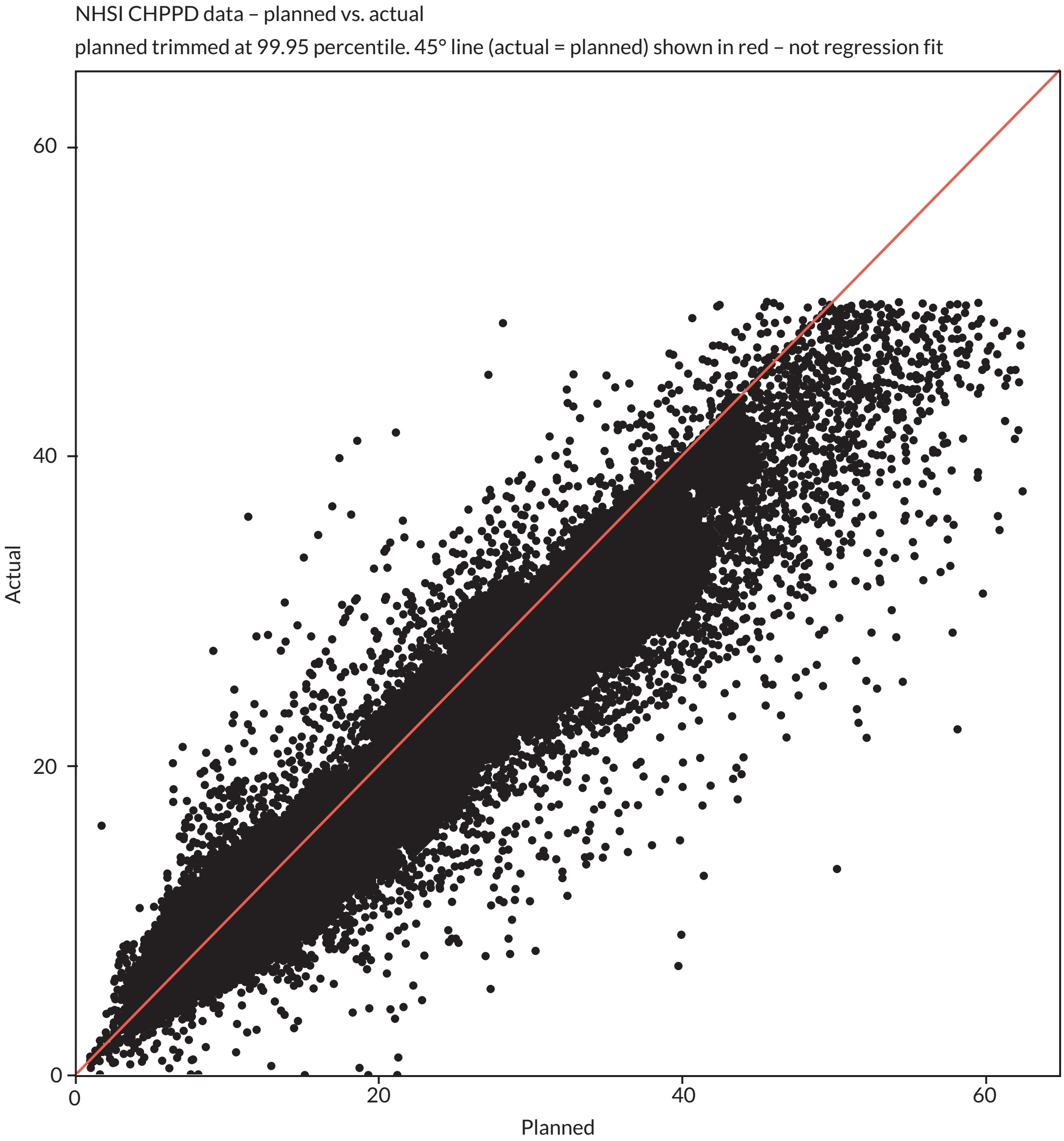


FIGURE 2 Scatterplot of planned vs actual CHPPD for acute hospital wards (monthly, May 2016–March 2020).

- In multiprofessional models, having more beds per RN was associated with lower scores for patient experience, staff health and well-being and staff reports of quality of care.
- More beds per nurse support were associated with lower morale scores but more beds per surgical doctor were associated with higher morale scores.
- Using ward-level reports of nurse staffing we found that wards with more RN HPPD reported fewer harms on the national 'safety thermometer' but calculating a staffing shortfall, relative to the Trusts' reported staffing plans, did not strengthen the observed relationship and, for NA staff, shortfalls were associated with reduced harms.

TABLE 9 Association between ward-level HPPD and harms recorded in the safety thermometer

	Actual staffing model				Shortfall (planned minus actual) model			
Staff groups	**IRR**	**UCL**	**LCL**		**IRR**	**UCL**	**LCL**	
RN HPPD	0.993	0.999	0.987	*	0.999	1.013	0.986	
NA HPPD	1.003	1.013	0.993		0.983	0.998	0.969	*
Proportion over 70 years old	2.084	2.215	1.960	***	2.097	2.228	1.973	***
Ward specialty								
Critical care medicine	1.945	2.312	1.637	***	1.733	1.976	1.519	***
Oncology/radiology/haematology	1.791	2.205	1.455	***	1.798	2.214	1.461	***
Misc other medicine and rehabilitation	1.700	1.911	1.513	***	1.704	1.914	1.517	***
Medicine for older people	1.579	1.717	1.452	***	1.584	1.722	1.457	***
Medicine	1.551	1.660	1.450	***	1.557	1.665	1.455	***
Neurosurgery	1.427	1.802	1.129	**	1.432	1.809	1.134	**
Accident and emergency	1.047	1.667	0.658		1.082	1.721	0.681	
Trauma and orthopaedics	0.942	1.032	0.860		0.949	1.039	0.866	
Cardiology	0.905	1.015	0.807		0.899	1.008	0.801	
Psychiatry, mental health and learning disability	0.813	1.127	0.586		0.819	1.134	0.591	
Cardiothoracic surgery	0.750	0.948	0.593	*	0.725	0.916	0.575	**
Obstetrics and midwifery	0.300	0.341	0.264	***	0.297	0.337	0.261	***
Paediatrics	0.091	0.118	0.070	***	0.089	0.115	0.069	***
Paediatric intensive care	0.050	0.115	0.021	***	0.048	0.111	0.021	***
Paediatric surgery	0.047	0.154	0.014	***	0.042	0.144	0.012	***
Trust type								
Specialist	0.607	0.749	0.492	***	0.604	0.744	0.490	***
Teaching	1.004	1.062	0.949		1.003	1.060	0.948	
Size (beds)								
2nd tertile	0.914	0.989	0.844	*	0.911	0.987	0.842	*
3rd tertile	0.832	0.905	0.764	***	0.829	0.903	0.762	***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
LCL, lower 95% confidence limit; misc, miscellaneous; UCL, upper 95% confidence limit.

Chapter 5 Trust data results

We received data from four hospital Trusts. For three of these, acute inpatient services were largely provided from a single site. For brevity, we refer to these hospital Trusts as 'hospitals'. Across the 4 hospitals staffing and patient data were linked using 192 ward identifiers. After we identified discontinuities in the data, indicating significant changes in case mix or other organisational changes, we had a total of 216 inpatient units which were treated as separate wards for analysis purposes. Of these, 185 were adult acute with the remainder being wards for children and/or women and children (maternity). The largest hospital included 100 units while the smallest included 29. The majority of units (185) were adult acute units, of which 106 were medical specialties. The majority of beds (51%) were in 65 larger units (31%) with 18 or more beds – the median size of these larger units was 22 beds. Forty-five units had a majority (50% or more) of elective admissions but only 12 had more than 75%. A small number of wards (31) had an estimated mortality risk of close to 0 (< 0.05%) but these were mostly paediatric and or maternity services. Most (122) had an appreciable risk (1% or more).

Adult inpatient ward admissions

We extracted data on 1,706,388 admissions. Of these, 714,526 were admitted to an eligible adult inpatient ward and had an overnight stay. Of these, we were able to link staffing data for 626,313 (88%). Reasons for not matching staffing to patients were due to stays in ineligible units, lack of standardisation between ward names in e-rosters and the Patient Administration System (PAS) system and, most frequently, periods of ward reorganisation and transition including ward moves where it appeared that different labels were applied in the e-roster and PAS (so, e.g., patients appeared to be cared for on wards with 0 staff). The median LOS was 3.6 days (mean 7.9). Eighty per cent of admissions were emergencies. The median age group was 65–69 with a mode of 80–84. Most admissions had at least one comorbidity contributing to the Charlson Comorbidity Index (CMI), with 45% having a CMI of > 5. Ninety-five per cent were alive 30 days after admission. The risk of death was higher for males, patients who presented as emergency admissions, who were older, and those with a CMI > 5. Fifteen per cent were re-admitted within 30 days (*Table 10*).

Wards and staffing

Registered nurse HPPD varied between wards within trusts ranging from an average of 2 (in some medical and surgical units) to over 24 (generally ICUs or small highly specialised units). The (unweighted) ward mean was 7.1 (median 5) ranging from 6.1 in Trust C to 9.6 in Trust D. Smaller units tended to be more specialised and had higher staffing levels ($r = -0.46$).

Across the analysis sample for adult admissions the mean staffing over the first 5 days of a patient stay was 5.3 RN HPPD (4.7 when excluding ICU stays) with 2.9 NA HPPD. Staffing levels varied considerably with an IQR from 3.1 to 5.7 RN HPPD. During the first 5 days of hospital stay patients were exposed for a mean of 3.32 (average LOS up to 5 days) and low staffing (relative to the ward mean) occurred on 45% of days. On days when RN staffing was low, the mean shortfall was 0.87 RN hours and 25% of patient days had deficits of 0.53 RN hours or more (13% of the mean RN HPPD). For NAs, the mean shortfall on days of low staffing was 0.49 hours. Days with a higher-than-average number of admissions relative to the number of RNs occurred less often than days of low staffing (37% of days), although this was moderately correlated with days of low HPPD for patients ($r = 0.37$).

Mean skill-mix was 61% RN. Twenty-five per cent of RN hours were provided by senior (band 6 and above) permanent members of staff and 10% were provided by temporary staff (5% each bank and agency). For NAs, only 2% of hours were provided by senior (band 4) staff (including nursing associates)/14% of hours were provided by temporary staff, mostly bank (12%).

On days of low staffing, the composition of the workforce was, overall, very similar to the overall mean with small absolute differences. RN skill-mix was slightly lower (−1.4%), the proportion of senior staff slightly higher with slightly

TABLE 10 Adult inpatient admissions

	All		**Alive**[a]		**Dead**[a]	
	n	**(%)**	***n***	**(row%)**	***n***	**(row%)**
Total *N*	626,313	(100.0%)	594,428	(94.9%)	31,885	(5.1%)
Female	348,464	(55.6%)	333,411	(56.1%)	15,053	(47.2%)
Emergency admission	502,717	(80.3%)	471,895	(79.4%)	30,822	(96.7%)
Age group						
15–19	13,699	(2.2%)	13,665	(2.3%)	34	(0.1%)
20–24	25,980	(4.1%)	25,927	(4.4%)	53	(0.2%)
25–29	33,779	(5.4%)	33,702	(5.7%)	77	(0.2%)
30–34	36,544	(5.8%)	36,413	(6.1%)	131	(0.4%)
35–39	29,740	(4.7%)	29,543	(5.0%)	197	(0.6%)
40–44	23,420	(3.7%)	23,131	(3.9%)	289	(0.9%)
45–49	28,092	(4.5%)	27,589	(4.6%)	503	(1.6%)
50–54	34,560	(5.5%)	33,756	(5.7%)	804	(2.5%)
55–59	38,843	(6.2%)	37,595	(6.3%)	1248	(3.9%)
60–64	42,138	(6.7%)	40,401	(6.8%)	1737	(5.4%)
65–69[b]	49,617	(7.9%)	47,071	(7.9%)	2546	(8.0%)
70–74	60,212	(9.6%)	56,532	(9.5%)	3680	(11.5%)
75–79	59,095	(9.4%)	54,974	(9.2%)	4121	(12.9%)
80–84	60,436	(9.6%)	55,156	(9.3%)	5280	(16.6%)
85–89	51,579	(8.2%)	45,983	(7.7%)	5596	(17.6%)
90–120	38,579	(6.2%)	32,990	(5.5%)	5589	(17.5%)
Charlson Index						
0	245,515	(39.2%)	242,986	(40.9%)	2529	(7.9%)
1–5	100,037	(16.0%)	97,842	(16.5%)	2195	(6.9%)
> 5	279,415	(44.6%)	252,275	(42.4%)	27,140	(85.1%)
Re-admitted[c]	94,821	(15.1%)				
LOS						
Mean (SD)	7.85	(13.88)	7.79	(14.15)	9.0	(7.00)
Median (IQR)	3.63	(1.77–8.28)	3.46	1.72–8.00	7.02	3.25–13.02

a Within 30 days of admission.
b Median age group.
c Within 30 days of discharge.

lower proportions of bank and agency staff (generally 1% or less difference). Although the overall staffing level was slightly lower, the composition of the workforce was generally very similar after excluding ICU stays (*Table 11*).

Effects of days of low hours per patient day

Patients who were exposed to days of low RN hours during the first 5 days of their stay were more likely to die (5.3% vs 4.0%) and be re-admitted (15.3% vs 14.3%). They had a longer overall LOS (8.4 days vs 5.3 days). Results were found similar for patients exposed to days of low NA staffing (*Table 12*).

In the multivariable survival model, each day a patient was exposed to understaffing (below the mean) in the first 5 days of their stay increased the hazard of death, re-admission and the LOS. Each day of RN understaffing increased the hazard of death by 8% (aHR 1.079, 95% CI 1.070 to 1.089) and the hazard of re-admission increased by 1% (aHR 1.01, 95% CI 1.005 to 1.016). When all days of a patient stay within 5 days of admission were understaffed, the LOS was increased by 65%. Days of low NA staffing were associated with similar but slightly smaller differences in mortality and LOS, but the hazard of re-admission was slightly lower (< 1%) when patients were exposed to days of NA understaffing. For full results, see *Table 13*.

Sensitivity analyses

There are a number of structural factors that largely operate at a hospital level, including levels of medical and other therapy staff, so we initially modelled each hospital trust separately. Results were broadly similar across the four. We then combined and tested models with Trust and ward-level effects. We found that models with effects for both Trust and ward had negligible difference in AIC (0.4) and BIC (1.2) from models with just ward effects, while providing almost precisely the same coefficients in our core models. Thus, we concluded that any hospital-level effects were properly accounted for by the ward random effect and so we omitted trust from the models to reduce the computational intensity, which for many models was considerable (some taking many hours to run with trust and ward effects included).

We assessed the effects of our decision to focus on the first 5 days of understaffing by computing models for the impact of low staffing on mortality using 3 and 10 days' exposure windows. Results of these analyses, available in *Appendix 3* (*Table 35*), were very similar to those from the 5-day exposure models. For example, 1 day of low RN staffing was estimated to increase the hazard of death by 6.8% using a 3-day exposure window and by 8.0% using a 10-day window.

In our previous study, we included an acuity risk score based on vital signs on admission, derived from the NEWS, in our mortality analyses. In this study, we could only derive a NEWS score for approximately 75% of patient days with no data from one Trust. We assessed the impact of this change by estimating models on this sub sample with and without the NEWS variable. Inclusion of NEWS improved model fit but made no appreciable difference to the estimated effect of low RN staffing (HR 1.087 vs 1.086) and a small reduction in the estimated effect of assistant staffing (HR 1.071 vs 1.059). As only those who were discharged alive could be re-admitted, we omitted patients who died from our re-admission models (reported above). Including all admissions in the model made little difference to the results, with a marginally smaller aHR for days of RN understaffing (HR 1.009).

As we were unable to directly assess variation in staffing by other groups, we tested models that assessed effects that might be associated with variation in staffing by other groups, specifically time of year (month) and weekends (both weekend admission and weekend stay). There was some evidence of a seasonal effect, with higher risk of death in January and February, but the overall model fit was worse. Weekend admissions were associated with increased risk (HR 1.092, 95% CI 1.065 to 1.120; $p < 0.001$), but there was no statistically significant association with weekend stays. However, in all cases, the estimated effects of low nurse staffing were unaltered.

TABLE 11 Patient day staffing characteristics for all days and days of low staffing (*n* = 626,313)

	All days			Days of low staffing			
All	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**	**Mean difference**
RN staffing							
(% days below mean)				(45%)			
HPPD	5.29	(4.22)	4.17	4.67	(3.88)	3.57	−0.62
Hours low staffing (one way)				0.87	(0.93)	0.62	0.38
Days with high transfers per RN (%)	37%	(0.43)	10%	43%	(0.44)	33%	6.2%
Permanent RN skill-mix (%)	61%	(0.12)	61%	60%	(0.13)	60%	−1.4%
Senior RN (band 6 +) (%)	25%	(0.18)	21%	26%	(0.20)	22%	0.9%
Bank %	5.2%	(0.10)	0.0%	4.8%	(0.10)	0.0%	−0.4%
Agency %	4.9%	(0.10)	0.0%	4.1%	(0.10)	0.0%	−0.8%
Assistant staffing							
(% days below mean)				(45%)			
HPPD	2.93	(1.37)	2.74	2.72	(1.19)	2.58	−0.21
Hours low staffing (one way)				0.49	(0.61)	0.29	0.10
Days with high transfers per NA (%)	33%	(0.42)	0%	33%	(0.42)	0%	0.0%
Senior NA (band 4) %	2.0%	(0.06)	0.0%	2.2%	(0.07)	0.0%	0.2%
Bank %	12.3%	(0.19)	0.0%	11.2%	(0.19)	0.0%	−1.1%
Agency %	1.5%	(0.06)	0.0%	1.2%	(0.05)	0.0%	−0.3%
Excluding ICU staffing							
RN staffing							
(% days below mean)				(45%)			
HPPD	4.76	(2.73)	4.09	4.13	(2.33)	3.48	−0.63
Hours low staffing (one way)				0.80	(0.78)	0.60	0.35
Days with high transfers per RN (%)	37%	(0.43)	8%	43%	(0.44)	33%	6.1%
Permanent RN skill-mix (%)	61%	(0.12)	61%	59%	(0.12)	59%	−1.5%
Senior RN (band 6 +) (%)	26%	(0.19)	22%	26.9%	(0.20)	22.4%	0.9%
Bank %	5.0%	(0.09)	0.0%	4.6%	(0.10)	0.0%	−0.4%
Agency %	5.0%	(0.10)	0.0%	4.1%	(0.10)	0.0%	−0.8%

TABLE 11 Patient day staffing characteristics for all days and days of low staffing (n = 626,313) (*continued*)

All	All days Mean	SD	Median	Days of low staffing Mean	SD	Median	Mean difference
Assistant staffing							
(% days below mean)				(45%)			
HPPD	2.92	(1.36)	2.74	2.70	(1.18)	2.58	−0.22
Hours low staffing (one way)				0.49	(0.60)	0.29	0.10
Days with high transfers per NA (%)	33%	(0.42)	0%	33%	(0.42)	0%	0.1%
Senior NA (band 4) %	1.8%	(0.05)	0.0%	2.0%	(0.06)	0.0%	0.1%
Bank %	11.9%	(0.18)	0.0%	10.9%	(0.19)	0.0%	−1.1%
Agency %	1.5%	(0.06)	0.0%	1.1%	(0.05)	0.0%	−0.4%

TABLE 12 Exposures and outcomes for those exposed vs not exposed to low staffing

		No exposure to low staffing		Exposed to low staffing		*p*
Low RN HPPD						
Days exposed	Mean (SD)	2.49	(1.45)	3.32	(1.40)	< 0.001
Net low staffing (cumulative hours)	Mean (SD)	−3.36	(4.04)	1.15	(3.51)	< 0.001
LOS	Mean (SD)	5.32	(10.67)	8.39	(14.43)	< 0.001
Died	*n* (%)	4488	(4.0%)	27,397	(5.3%)	< 0.001
Re-admitted	*n* (%)	15,924	(14.3%)	78,897	(15.3%)	< 0.001
Total	*N* (%)	111,414	(17.8%)	514,899	(82.2%)	
Low NA per patient day						
Days exposed	Mean (SD)	2.54	(1.47)	3.32	(1.40)	< 0.001
Net low staffing (cumulative hours)	Mean (SD)	−3.05	(3.07)	0.88	(2.54)	< 0.001
LOS	Mean (SD)	5.56	(11.50)	8.36	(14.31)	< 0.001
Died	*n* (%)	4405	(3.9%)	27,440	(5.4%)	< 0.001
Re-admitted	*n* (%)	16,247	(14.2%)	78,342	(15.4%)	< 0.001
Total	*N* (%)	114,363	(18.3%)	510,517	(81.7%)	

TABLE 13 Association between days of understaffing and outcomes, adjusted for risk

All staffing	Mortality				Re-admission				LOS			
Term	HR	*p*	LCL	UCL	HR	*p*	LCL	UCL	Ratio[a]	*p*	LCL	UCL
SHMI (risk)	1.063	0.000	1.062	1.064	1.006	0.000	1.005	1.007	1.032	0.000	1.031	1.032
Days of low RN HPPD	1.079	0.000	1.070	1.089	1.010	0.000	1.005	1.016	1.687	0.000	1.666	1.707
Days of low NA HPPD	1.072	0.000	1.062	1.081	0.994	0.020	0.988	0.999	1.608	0.000	1.589	1.627
Ward random effect (SD)	1.205		AIC	BIC	0.657		AIC	BIC	0.902		AIC	BIC
Model fit			788,843	790,453			2,424,881	2,426,790			7,025,587	7,025,655

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

[a] Gamma regression exponentiated coefficient is a ratio and the exposure is the proportion of days in the patient stay with understaffing.

Patient and staffing subgroups

We explored effects in subgroups. These included staffing in ICU and general (non-ICU) wards; acutely ill patients and those who were less acute (patients with NEWS score of 5+ or < 5 on admission); and older people (75+) and wards for older people (75% of patients 75+). Results are summarised in *Figure 3* and *Appendix 3*, *Table 39*.

Results from the lower acuity patient subgroup (NEWS < 5 on admission) and when considering general ward staffing only were similar to the full sample analysis, with marginally greater effects on mortality. When considering ICU staffing only there were no statistically significant effects other than a reduction in LOS associated with fewer days of low assistant staffing. For patients who were more acutely ill on admission (NEWS 5+), days of low RN were associated with an increased hazard of death (4%) and longer hospital stays. Low assistant staffing was not associated with increased risk of death, but it was associated with increased LOSs. Looking at staffing on wards for older people low staffing by RNs and NAs was associated with an increased hazard of death (4%) and longer LOSs, but there was no statistically significant association with re-admissions. Looking at older people (75+) only effects were generally small and nonsignificant, although days of low staffing were associated with increased LOS and days of low assistant staffing were associated with reduced re-admissions.

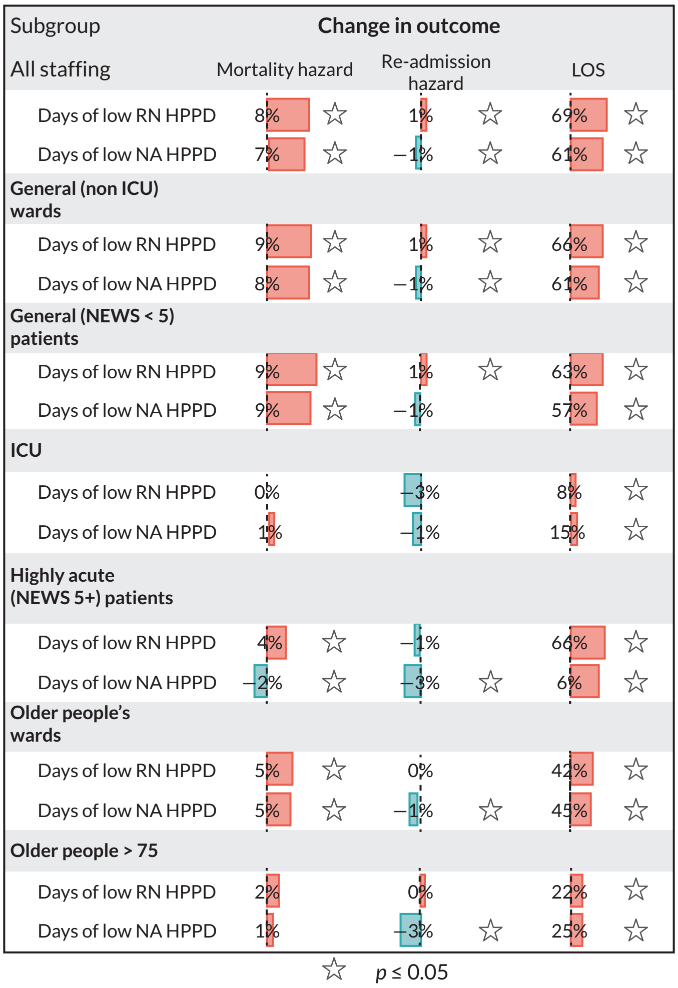


FIGURE 3 Summary of subgroup analyses.

Effects of staff-mix

We assessed the effect of the mix of staff on mortality, initially by adding variables to the models for each of the following: RN skill-mix (RN hours as a proportion of all hours), grade mix (proportion of permanent senior RNs, band 6 and above and proportion of permanent band 4 staff), bank staff (proportion of RN and NA staff who were employed through the bank) and agency staff (proportion of RN and NA staff who were employed through an agency). Full results of these survival models are reported in *Appendix 3*, *Table 36*. In all cases, adding these staff mix variables clearly improved model fit (reduced both AIC/BIC) relative to the model with low staffing effects only, except RN skill-mix, which led to improvement in one criterion but worsening of another. The association between RN skill-mix and mortality was not statistically significant. We therefore combined all statistically significant factors (grade mix, bank staffing and agency staffing) into a single model, which further improved model fit.

In the combined models the effect of low staffing was marginally higher than in the low staffing model only. For each 10% increase in the proportion of temporary RNs (bank and agency), the hazard of death was increased by 2.3%. Increases in bank assistants had a slightly smaller effect but every 10% increase in the proportion of agency assistants increased the hazard of death by 4%. The effect of grade mix was reduced compared to the models of low staffing and grade mix only and the effect was no longer statistically significant (*Table 14*). We tested for interactions between low staffing and temporary staffing but none of the effects were statistically significant.

Low staffing thresholds

We used the mean staffing for each ward as an indication of the ward norm. In previous research, we found that this threshold was highly correlated with, but slightly lower than the planned RN staffing level and slightly higher than the planned assistant staffing level. Because of the large number of wards in this study and the limited availability of acuity dependency data to estimate staffing requirements, we tested the effect of different thresholds to define low staffing. Estimated effects of low staffing were largely unchanged if thresholds were set below 100% of the mean. For assistant staffing this is also the case for thresholds above the mean. However, for RN staffing the effects of days of low staffing increase when a higher threshold is used. For example, if low staffing were defined as below 110% of the mean the HR associated with a day of low staffing is 1.12 (*Figure 4*).

Interactions and non-linear effects

We repeated our models using cumulative net hours of understaffing (relative to the mean) as a continuous measure of understaffing. Results were qualitatively consistent (in terms of direction of effect) with the estimated effects of days of understaffing, although associations between net hours and outcomes were small and not always statistically significant (see *Appendix 3*, *Table 37*). Therefore, we moved to explore nonlinear and interaction effects for both continuous (net hours) and cumulative sum of days of understaffing. Adding nonlinear (up to third-order polynomials) and interaction terms independently either improved or did not worsen model fit (as judged by at least one of AIC/BIC being lower) compared to the linear model with no interactions. Interaction terms and some nonlinear terms were statistically significant (see *Appendix 3*, *Table 38*). There is an improvement in model fit when nonlinear terms are included, although the nonlinear effects are subtle and do not diverge dramatically from purely linear estimates.

In *Figure 5* staffing levels decrease from left to right (as understaffing increases) expressed as deciles of hours or days of low staffing. Considering net hours (*Figure 5a*), there is some evidence of larger incremental effects at extreme values (both high and low) but no evidence of a plateau where benefits of higher staffing cease. Focusing on the measures of days of low staffing (*Figure 5b*), a small increase in the rate of adverse effect is seen at higher levels (more days) of low staffing. Similarly, although the interaction was statistically significant, there is little appreciable difference in the effect of change in RN hours at different levels of assistant staffing. Days with low HCA staffing seem to make more difference when RN staffing is adequate (low levels of understaffing), but the effect is modest.

TABLE 14 Staff mix and mortality

	HR	*p*	LCL	UCL
SHMI	1.063	0.000	1.062	1.064
Days of RN understaffing	1.083	0.000	1.074	1.093
Days of NA understaffing	1.081	0.000	1.071	1.091
Proportion of senior RNs	0.993	0.201	0.982	1.004
Proportion of senior NAs	0.983	0.173	0.958	1.008
Proportion of bank RNs	1.023	0.005	1.007	1.039
Proportion of bank NAs	1.019	0.000	1.010	1.028
Proportion of agency RNs	1.023	0.000	1.011	1.037
Proportion of agency NAs	1.040	0.000	1.021	1.060
Ward random effect (SD)	1.206			
Model fit			AIC	BIC
			788,748	790,409

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

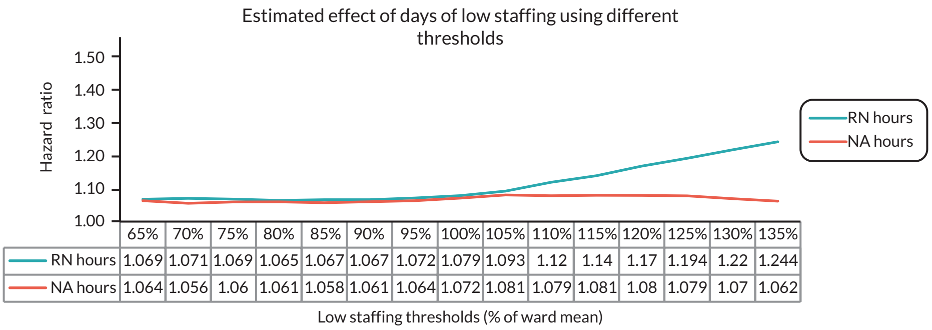
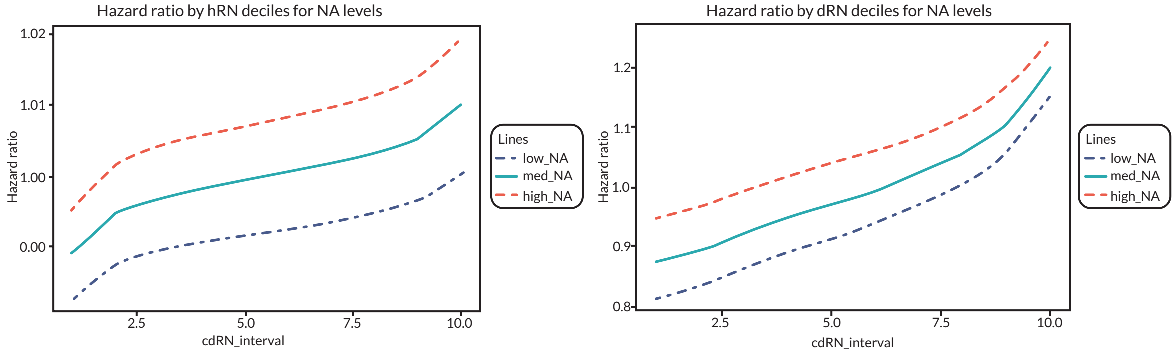


FIGURE 4 Hazard ratio for days of low staffing for different thresholds.



(a) Hours of low staffing (b) Days of low staffing

FIGURE 5 Nonlinear interactions: low RN staffing conditional on NA staffing.

Patient transfers

To assess the additional impact of workload associated with patient transfers, we added variables for days of with high numbers of patient movements (above the mean) in or out of the ward (transfers = admissions + discharges), relative to the number of staff. The survival models for the relationship with mortality are presented in *Table 15*. In general, adding terms for workload associated with patient transfer slightly reduced the estimated effects of days of low HPPD compared to models without patient transfers (see *Table 15*, model a). When high transfers per RN and NA staff were included (see *Table 15*, model b), the effect of 1 day with high transfer per RN was to increase the hazard of death by 6% (aHR 1.056, 95% CI 1.040 to 1.072). While high transfers per NA was also associated with increased hazard of death, the effect was smaller and not statistically significant (aHR 1.016, 95% CI 1.000 to 1.033).

Dropping the term for transfers per NA staff (model c) resulted in similar model fit (small increase in AIC, lower BIC) and broadly similar estimates for RN staffing effects (small decrease in the estimate of the effect of understaffing and an increased estimate of the effect of high transfers).

Adverse events (patient)

Nurse-sensitive adverse events were extracted from the patient record (PAS) for 220,741 adult surgical admissions. Medical and other admissions were excluded as there is no 'present on admission flag' and all three conditions (deep vein thrombosis – DVT, pneumonia and PU) are possible causes of admission or likely to be prevalent in some patient groups admitted to medical wards.

As the timing of the event within the admission is not recorded, a survival approach to analysis, which can accommodate time-varying covariates, was not feasible or meaningful. Therefore, we undertook logistic regression using the proportion of days with low staffing, cumulative (net) hours of low staffing and low staffing as a proportion of the expected staffing. In all cases, staffing measures were based on the first 5 days of the patient stay and all staffing measures were relative to the ward means. All analyses included the SHMI risk score (a composite measure of risk including age, main diagnostic group and comorbidities) as control and a random effect for ward.

The recorded prevalence of all three conditions was low, ranging from 0.3% for DVT to 3.2% for pneumonia (*Table 16*). There was evidence of associations between low RN staffing and increased odds of adverse events across all three adverse events. Where all of the first 5 days of a patient's stay had low staffing (HPPD below ward mean), the risk of PU increased by 85% (OR 1.851, 95% CI 1.516 to 2.261), the risk of pneumonia increased by 73% and DVT by 58%. Low assistant staffing was also associated with an increased risk of adverse events, although the effect size was generally smaller.

We assessed the impact of high patient turnover (above mean turnover per RN) by estimating models that included this variable. Adding this variable to the models did not clearly improve model fit except for pneumonia. If all days in a patient's stay had high turnover, the risk of pneumonia increased by 60% (OR 1.600, 95% CI 1.443 to 1.775). In these models, the coefficients for days of low HPPD were found slightly reduced.

Incidents (Datix)

We identified incident reports recorded as patient-related occurring on included wards. We grouped incidents using the major descriptors and reported tissue damage, falls, medication incidents and others. Incidents were also categorised by severity (no harm, harm, significant harm and severe harm). The incident reports were linked to ward days as opposed to individual patients; therefore, we calculated Poisson regression models using daily ward occupancy as an offset to assess the association between nurse staffing levels and the number of incidents occurring on that day. Because tissue damage can take time to become apparent, we also looked at lagged associations, using staffing on the previous 3 days. All models included control for the overall acuity of patients on the ward by including the mean SHMI risk score of patients on the ward on that day.

In total, there were 128,620 incidents recorded over 229,382 ward days relating to 4,931,659 patient days. The overall incident rate was 26.1 per 1000 patient days. Tissue damage was the most reported incident (8.4 per 1000 patient days). Overall, most incidents (68.5%) were recorded as leading to no harm, but for tissue damage most incidents were recorded as causing harm (54%). Few incidents were recorded as causing significant or severe harm (2% overall). See *Appendix 3*, *Table 40*.

Results are summarised in *Table 17*. There was a counterintuitive association between the rate of incidents reported and RN HPPD. The rate of all incident reporting was 4% higher (RR 1.041; $p < 0.001$) for each additional RN hour per patient day. Similar associations were observed for the specific incident types of medication errors (RR 1.041; $p < 0.001$) and tissue damage (RR 1.062; $p < 0.001$). In contrast, for falls, higher RN staffing was associated with fewer reported incidents. Each additional RN hour per patient day was associated with a 4% reduction in the rate of falls reported (RR 0.936; $p < 0.001$).

Where higher staffing was associated with higher rates of reported incidents, the magnitude of the effect of staffing was smaller for incidents causing greater harm and, in the case of medication errors, the relationship was reversed, with fewer incidents causing severe harm reported when staffing is higher (not statistically significant, $p = 0.276$). For falls, the magnitude of the beneficial effect of higher staffing was greater for more severe incidents, with the reported rate of falls with severe harm reducing by 15% for each additional RN hour per patient day (RR 0.855; $p = 0.001$).

For assistant staffing, there is a broadly similar pattern of results, albeit with fewer statistically significant associations, less consistent patterns and, in most cases, smaller effect sizes. Additional hours of HCA staffing were associated with increased reporting of total incidents falls, medication errors and tissue damage (all statistically significant, $p < 0.05$).

For tissue damage however, increased assistant staffing was associated with a reduced rate of reported incidents with harm. Reports of PUs with severe harm are reduced by 14% for each additional NA hour per patient (OR

TABLE 15 Association between days of understaffing, patient transfers and mortality, adjusted for risk (all admissions)

All staffing	Model a				Model b				Model c			
Term	HR	*p*	LCL	UCL	HR	*p*	LCL	UCL	HR	*p*	LCL	UCL
SHMI (risk)	1.063	0.000	1.062	1.064	1.063	0.000	1.062	1.064	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.079	0.000	1.070	1.089	1.060	0.000	1.049	1.070	1.056	0.000	1.046	1.067
Days of high transfers per RN					1.056	0.000	1.040	1.072	1.066	0.000	1.054	1.079
Days of low NA HPPD	1.072	0.000	1.062	1.081	1.060	0.000	1.048	1.071	1.067	0.000	1.057	1.076
Days of high transfers per NA					1.016	0.051	1.000	1.033				
Ward random effect (SD)	1.205		AIC	BIC	1.196		AIC	BIC	1.197		AIC	BIC
Model fit			788,843	790,453			788,725.3	790,350.7			788,726.9	790,344

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

TABLE 16 Frequency of nurse-sensitive adverse events and associations with staffing (logistic regression)

	DVT				Pneumonia				PU			
Events *n* (%)	654	(0.3%)			7102	(3.2%)			1174	(0.5%)		
Term	OR	*p*	LCL	UCL	OR	*p*	LCL	UCL	OR	*p*	LCL	UCL
SHMI (risk)	1.016	0.002	1.006	1.026	1.065	0.000	1.062	1.067	1.052	0.000	1.046	1.057
Days of low RN HPPD	1.588	0.001	1.216	2.075	1.737	0.000	1.595	1.891	1.851	0.000	1.516	2.261
Days of NA HPPD	1.332	0.032	1.025	1.731	1.643	0.000	1.512	1.785	1.331	0.004	1.096	1.616
Ward random effect (SD)	0.764		AIC	BIC	1.030		AIC	BIC	0.895		AIC	BIC
Model fit			8307	8359			53,096	53,147			13,384	13,435

LCL, lower 95% confidence limit; OR, odds ratio; UCL, upper 95% confidence limit.

TABLE 17 Association between reported incident rates and staffing levels for different levels of harm

	RN HPPD		NA HPPD	
Incident type	**RR**	***p***	**RR**	***p***
All incidents	1.041	0.000	1.010	0.000
Incidents with any harm	1.041	0.000	0.997	0.489
Incidents with significant or severe harm	1.025	0.000	1.021	0.147
Incidents with severe harm	1.024	0.000	0.969	0.145
Specific incident types				
Falls				
All falls	0.936	0.000	1.052	0.000
With any harm	0.936	0.000	1.044	0.000
With significant or severe harm	0.888	0.000	1.059	0.163
With severe harm	0.855	0.001	1.042	0.544
Medication error				
All medication errors	1.040	0.000	1.014	0.003
With any harm	1.018	0.000	1.017	0.227
With significant or severe harm	1.013	0.510	1.006	0.928
With severe harm	0.824	0.276	1.059	0.811
Tissue damage				
All tissue damage	1.062	0.000	1.015	0.000
With any harm	1.059	0.000	0.979	0.000
With significant or severe harm	1.031	0.000	0.975	0.304
With severe harm	1.043	0.025	0.857	0.013

0.857, $p = 0.013$). However, for other incident types, there was no clear pattern of changes in effect size as incident severity increased.

We explored lagged effects of staffing on tissue damage because damage and its consequences might take time to become apparent. Adverse effects of higher RN staffing were generally smaller with longer lags, although the differences were small and overall results and patterns of statistical significance were very similar. For example, 1 additional RN hour per patient day was associated with a 38% increase in reports of PUs with severe harm 3 days later (compared to a 43% increase for staffing on the day). A 2-day lag appeared to be associated with a larger beneficial effect from higher HCA staffing in relation to incidents with harm, although the magnitude of difference was still relatively small (e.g. 22% reduction in severe harm for a 2-day lag vs 14% for 0 lag).

Paediatrics

We calculated separate models for 57,375 paediatric admissions (< 18) in paediatric units. Most (53%) were under 5 years old. Fourteen per cent had one or more comorbidities and the estimated risk of death was low (mean SHMI risk 0.15). Mean LOS was 3.6 days (median 1.7), 11.3% were re-admitted within 30 days and 0.25% died.

Patients who were exposed to days of low RN hours during the first 5 days of their stay were more likely to die (0.26% vs 0.21%) and be re-admitted (11.6% vs 9.5%). They had a longer overall LOS (3.89 days vs 2.18 days). Results were similar for patients exposed to days of low NA staffing (*Table 18*).

In the multivariable survival model, each day a patient was exposed to understaffing (below the mean) in the first 5 days of their stay increased the hazard of death, re-admission and the LOS, although the associations with hazard of death were not statistically significant. Each day of RN understaffing increased the hazard of re-admission by 8% (aHR 1.077, 95% CI 1.050 to 1.105). The effect of assistant understaffing was similar. When all days of a patient's stay were understaffed, the LOS was more than doubled (RN/NA understaffing 2.46/2.36). For full results, see *Table 19*.

Staff sickness absence

We estimated the effects of staffing levels and skill-mix by summarising the staffing, staff mix and working patterns in the week preceding every shift worked and the first shift that was missed through sickness.

In total, there were 43,097 sickness absence episodes and 2,646,983 worked shifts included in the analysis. The mean length of a sickness episode was 11.23 days (SD = 32.12 days), and the median sickness length was 4.47 days. *Table 20* reports the mean distribution of exposure to understaffing, skill-mix, bank and agency, proportion of long shifts in the 7 days by sickness cohort.

Effects of exposure to registered nurse and nursing assistant understaffing

Exposure to 10% RN understaffing in the prior 7 days was associated with a 2% increase in the odds of experiencing sickness absence for RNs (OR 1.018, 95% CI 1.004 to 1.032) and a 3% increase in the odds of experiencing sickness absence for NAs (OR 1.025, 95% CI 1.011 to 1.039). Exposure to NA understaffing was not associated with a statistically significant change in the odds of experiencing sickness absence for either RNs or NAs (*Table 21*). We tested for interactions between RN and NA understaffing, but the relationship was not statistically significant and model fit was worse.

Staff mix and working patterns

We assessed the effect of the mix of staff and other aspects of working patterns on sickness, initially by adding variables to the models for each of the following: RN skill-mix (RN hours as a proportion of all hours), bank staff (proportion of RN and NA staff who were employed through the bank), agency staff (proportion of RN and NA staff who were employed through an agency) and part-time work and working long shifts. In all cases, the variables were calculated from proportions of hours over the previous 7 days, except long shifts (proportion of shifts of 12 or more hours) and

TABLE 18 Exposures and outcomes for those exposed vs not exposed to low staffing (paediatrics)

		No exposure to low staffing		Exposed to low staffing		*p*
Low RN HPPD						
Days exposed	Mean (SD)	1.543	(1.09)	2.267	(1.43)	< 0.001
Net low staffing (cumulative hours)	Mean (SD)	−6.648	(8.26)	3.064	(6.70)	< 0.001
LOS	Mean (SD)	2.18	(6.36)	3.89	(10.99)	< 0.001
Died	*n* (%)	20	(0.21%)	123	(0.26%)	0.431
Re-admitted	*n* (%)	916	(9.51%)	5553	(11.63%)	< 0.001
Total	*N* (%)	9636	(16.79%)	47,739	(83.21%)	
Low NA per patient day						
Days exposed	Mean (SD)	1.527	(1.06)	2.307	(1.44)	< 0.001
Net low staffing (cumulative hours)	Mean (SD)	−3.6	(3.87)	1.304	(3.05)	< 0.001
LOS	Mean (SD)	2.132	(7.39)	3.984	(11.00)	< 0.001
Died	*n* (%)	18	(0.16%)	122	(0.27%)	0.05
Re-admitted	*n* (%)	1087	(9.65%)	5336	(11.70%)	< 0.001
Total	*N* (%)	11,263	(19.80%)	45,620	(80.20%)	

TABLE 19 Association between days of understaffing and outcomes, adjusted for risk (paediatrics)

All staffing	Mortality				Re-admission				LOS			
Term	**HR**	***p***	**LCL**	**UCL**	**HR**	***p***	**LCL**	**UCL**	**Ratio**[a]	***p***	**LCL**	**UCL**
SHMI (risk)	1.064	0.000	1.047	1.081	1.011	0.242	0.992	1.031	1.319	0.000	1.298	1.341
Days of low RN HPPD	1.099	0.195	0.952	1.269	1.077	0.000	1.050	1.105	2.461	0.000	2.323	2.607
Days of low NA HPPD	1.158	0.069	0.989	1.357	1.078	0.000	1.049	1.108	2.358	0.000	2.226	2.498
Ward random effect (SD)	1.522		AIC	BIC	0.536		AIC	BIC	1.162		AIC	BIC
Model fit			2614	2662			136,237	136,386			579,579	579,633

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.
a Gamma regression exponentiated coefficient is a ratio and the exposure is the proportion of days in the patient stay with understaffing.

TABLE 20 Staffing configurations in the previous week by sickness cohort

	All		RNs		NAs	
	Median (%)	Mean (%)	Median (%)	Mean (%)	Median (%)	Mean (%)
RN understaffing[a]						
Sick	6	9	4.8	8.2	10	15
Worked	5	9	5	8	10	15
NA understaffing[a]						
Sick	8	13	6.8	10.4	5.7	10.9
Worked	8	14	7	10	6	11
Agency staffing[b]						
Sick	0	4	0	3	0	4
Worked	0	3	0	2	0	3
Bank staffing[b]						
Sick	0	7	0	7.1	0	7.8
Worked	0	7	0	6	0	8
Skill-mix[c]						
Sick	61	63	66	67	55	56
Worked	63	64	68	68	56	56
Proportion of long shifts worked						
Sick	100	79	100	81.2	100	76.5
Worked	100	78	100	78.2	100	78

a RN and NA understaffing were expressed as a fraction of the mean [1 – (observed/expected)] where observed HPPD are smaller than expected HPPD in the week prior.
b Bank and agency staffing were expressed as proportion of total bank and agency hours for simultaneous ward shifts over total nursing hours during week prior.
c Calculated as proportion of total ward RN hours over total ward nursing hours for simultaneous ward shifts.

TABLE 21 Association between exposure to understaffing and sickness absence

	RNs				NAs			
Term	OR	*p*	LCL	UCL	OR	*p*	LCL	UCL
Exposure to + 10% RN understaffing per shift in prior 7 days	1.018	0.007	1.004	1.032	1.025	< 0.0001	1.011	1.039
Exposure to + 10% NA understaffing per shift in prior 7 days	0.998	0.677	0.990	1.006	1.002	0.710	0.990	1.013
Model fit			AIC	BIC			AIC	BIC
AIC			255,068	255,129			170,349	170,407

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

Note
All models include staff ID and ward as random effects.

part-time work, which was identified if the median weekly worked hours in the previous 13 weeks was ≤ 26. Higher exposure to agency shifts resulted in odds of sickness absence increasing for both and RNs (OR 1.026, 95% CI 0.996 to 1.057) and NAs (OR 1.025, 95% CI 0.995 to 1.056), although these relationships were not statistically significant. Higher exposure to bank hours (each 10% additional increment) was associated with a 2% reduction in the odds of a sickness absence episode for RNs (OR 0.983, 95% CI 0.966 to 0.999). Results were similar for NAs' sickness absence (OR 0.954, 95% CI 0.935 to 0.973). A skill-mix richer in RNs was associated with a reduced likelihood of experiencing sickness absence for RNs, although the association was not statistically significant, and it was associated with a higher likelihood of experiencing sickness absence for NAs (OR 1.034, 95% CI 1.008 to 1.06).

Working part-time was associated with higher odds of sickness absence for both RNs (OR 1.033, 95% CI 0.989 to 1.079) and NAs (OR 1.041, 95% CI 0.990 to 1.094), although the association was not statistically significant. Working higher proportion of long shifts in the prior 7 days was associated with higher sickness absence for RNs (each 10% increase in the proportion of long shifts: OR 1.026, 95% CI 1.020 to 1.031) but not for NAs (each 10% increase in the proportion of long shifts: OR 1.003, 95% CI 0.997 to 1.009) (*Table 22*).

Our final model included, in addition to RN and NA understaffing variables, all variables that were statistically significant associated with sickness absence. In this, the only statistically significant predictor of RN sickness absence was the percentage of long shifts worked. For RNs, each 10% increase in understaffed shifts in the previous 7 days was associated with a 1.2% increase in the odds of sickness, but this relationship was not statistically significant. Each 10% increase in the proportion of long shifts increased the odds of sickness by 2.6%, so a shift scheduled for a RN working all long shifts is 29% more likely to result in sickness absence than a shift for a RN who does no long shifts. A higher proportion of bank hours in the previous 7 days was associated with a 2.6% reduction in the odds of sickness absence.

Registered nurse understaffing remained a significant predictor of NA sickness, with every 10% increase in the proportion of RN understaffing associated with a 3% increase in the odds of sickness absence. A higher proportion of bank hours in the previous 7 days was associated with a 2.4% reduction in the odds of sickness absence. The full models can be found in *Table 23*.

Cost consequences and cost-effectiveness

Based on the weighted average HRG cost, the estimated total cost of providing care for the 626,313 admissions in our analysis cohort for adult inpatients was £2,613,385,125, or £4173 per admission. The mean estimated discounted and QALY among patients who died was 6.82. For our base-case scenario, we assumed that this reflected the quality-adjusted life expectancy of any lives saved from avoided deaths resulting from reduced understaffing.

Six thousand five hundred and twenty-seven of 31,885 deaths could have been avoided if all episodes of low staffing were eliminated, assuming the estimated effects represent the effect that would be achieved through intervening to reduce low staffing. The cost of additional staffing would be £197 per admission, £18,930 per life saved and £2788 per QALY. Cost savings associated with reduced sickness absence (1.4% of additional staff costs) and re-admission (2%) were modest, but the estimated savings resulting from a shorter LOS exceeded the additional staff costs (267% of staff costs). The estimated cost per QALY was lower if savings from sickness absence and re-admissions were considered, but the difference from an estimate based on staff costs only was not great (£2685 vs £2778 per QALY). When costs associated with LOS were considered, the cost per QALY was negative because the intervention improved outcomes and saved money (*Table 24*). Therefore, in the rest of this section, unless explicitly noted, the figures quoted are for simple staffing costs per QALY gained, providing a conservative estimate. For other estimates, refer to the relevant tables.

We tested the sensitivity of the analysis to the discount rate applied, assumptions about the quality and length of life that would be expected for lives saved, statistical uncertainties about consequences and cost assumptions (see *Table 24*). If a higher discount rate of 5% was applied, as recommended for economic evaluations in lower and lower middle-income countries,[151] cost per QALY was increased by 11% to £3081. If the lower CI was taken for all effect estimates, the estimated cost per QALY was 14% higher (£3150) and under the most pessimistic cost assumptions we

TABLE 22 Association between understaffing and sickness absence, adjusted for skill-mix, temporary staffing, shift work

	RNs				NAs			
Term	**OR**	***p***	**LCL**	**UCL**	**OR**	***p***	**LCL**	**UCL**
Bank staffing								
+ 10% RN understaffing per shift	1.018	0.007	1.004	1.032	1.025	< 0.0001	1.011	1.039
+ 10% NA understaffing per shift	0.998	0.677	0.99	1.006	1.002	0.710	0.99	1.013
+ 10% bank hours	0.983	0.044	0.966	0.999	0.954	0	0.935	0.973
Model fit			AIC	BIC			AIC	BIC
			255,068	255,129			170,349	170,407
Agency staffing								
+ 10% RN understaffing per shift	1.020	0.005	1.006	1.034	1.027	0.0001	1.013	1.042
+ 10% NA understaffing per shift	0.999	0.825	0.991	1.007	1.003	0.579	0.992	1.015
+ 10% agency hours on ward	1.026	0.094	0.996	1.057	1.025	0.099	0.995	1.056
Model fit			AIC	BIC			AIC	BIC
			255,067	255,141			170,291	170,362
RN skill -mix								
+ 10% RN understaffing per shift	1.016	0.028	1.002	1.031	1.036	0	1.011	1.052
+ 10% NA understaffing per shift	1.001	0.916	0.991	1.01	0.994	0.344	0.981	1.007
+ 10% RN skill-mix	0.991	0.369	0.971	1.011	1.034	0.010	1.008	1.06
Model fit			AIC	BIC			AIC	BIC
			255,069	255,143			168,998	169,068
Part-time work								
+ 10% RN understaffing per shift	1.013	0.103	0.998	1.028	1.022	0.006	1.006	1.038
+ 10% NA understaffing per shift	0.998	0.587	0.989	1.006	1.007	0.317	0.994	1.019
Part-time (yes vs no)	1.033	0.147	0.989	1.079	1.041	0.119	0.99	1.094
Model fit			AIC	BIC			AIC	BIC
			220,024	220,098			144,073	144,144
Long shifts								
+ 10% RN understaffing per shift	1.016	0.026	1.002	1.029	1.025	0.0003	1.011	1.039
+ 10% NA understaffing per shift	0.999	0.712	0.991	1.006	1.002	0.711	0.991	1.014

TABLE 22 Association between understaffing and sickness absence, adjusted for skill-mix, temporary staffing, shift work (*continued*)

	RNs				NAs			
Term	**OR**	***p***	**LCL**	**UCL**	**OR**	***p***	**LCL**	**UCL**
+ 10% long shifts worked	1.026	0	1.02	1.031	1.003	0.303	0.997	1.009
Model fit			AIC	BIC			AIC	BIC
			254,972	255,047			170,349	170,420

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

Note

All models include staff ID and ward as random effects. All exposures are averaged per shift over the prior 7 days.

TABLE 23 Association between understaffing and sickness absence, adjusted for statistically significant variables added to the core models

Term	OR	*p*	LCL	UCL
RNs' sickness absence				
+ 10% RN understaffing per shift	1.012	0.114	0.997	1.026
+ 10% NA understaffing per shift	0.999	0.820	0.990	1.008
+ 10% bank hours	0.974	0.003	0.957	0.991
+ 10% RN skill-mix in previous 7 days	0.987	0.235	0.967	1.008
+ 10% long shifts worked in previous 7 days	1.026	0	1.021	1.031
			AIC	BIC
			254,967.2	255,066.3
NA s' sickness absence				
+ 10% RN understaffing per shift	1.030	< 0.001	1.014	1.046
+ 10% NA understaffing per shift	0.991	0.181	0.978	1.004
+ 10% bank hours	0.957	0	0.938	0.976
+ 10% RN skill-mix in previous 7 days	1.024	0.064	0.999	1.051
			AIC	BIC
			170,270	170,352

LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

Note

All models include staff ID and ward as random effects. All exposures are averaged per shift over the prior 7 days.

TABLE 24 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing (base-case and sensitivity analysis)

Assumptions	Lives saved	QALY gains	Additional staff cost per admission (£)	Value of reduced sickness absence (% of staff costs)	Value of averted re-admissions (% of staff costs)	Value of reduced stays (% of staff costs)	Additional staff cost per life saved (£)	Additional staff cost per QALY gained (£)	Net additional cost per QALY gained (sickness costed) (£)[a]	Net additional cost per QALY gained (sickness and re-admissions costed) (£)[b]	Net additional cost per QALY gained (sickness, re-admissions and LOS costed) (£)[c]
Base case	**6527**	**44,483**	**197**	**1.4%**	**2.0%**	**267%**	**18,930**	**2778**	**2739**	**2685**	**−4728**
Sensitivities											
Lower discount rate (3%)	6527	46,183	197	1.4%	2.0%	267%	18,930	2675	2638	2586	−4554
Higher discount rate (5%)	6527	40,097	197	1.4%	2.0%	267%	18,930	3081	3038	2978	−5245
High-risk populations saved (over 80)	6527	22,605	197	1.4%	2.0%	267%	18,930	5466	5389	5283	−9304
Lower-risk populations saved	6527	67,843	197	1.4%	2.0%	267%	18,930	1821	1796	1760	−3100
LCL for all consequences	5746	39,158	197	0.2%	−1.7%	261%	21,504	3155	3150	3203	−5041
UCL for all consequences	7303	49,773	197	3.4%	5.6%	272%	16,918	2482	2399	2261	−4503
Lower value for unit costs	6527	44,483	181	1.4%	0.8%	267%	17,363	2548	2513	2492	−4317
Higher value for unit costs	6527	44,483	227	1.4%	2.0%	265%	21,788	3197	3153	3088	−5395
Optimistic cost assumptions[d]	6527	44,483	181	1.4%	2.5%	333%	17,363	2548	2513	2448	−6035
Pessimistic cost assumptions[e]	6527	44,483	227	1.4%	0.7%	213%	21,788	3197	3153	3132	−3677

LCL, lower confidence limit; UCL, upper confidence limit.

a Costs of avoided sickness absence removed.

b Costs of avoided sickness absence and re-admissions removed.

c Costs of avoided sickness absence, re-admissions and days of stay removed.

d Lower limit for staffing unit costs, upper limit for re-admission/LOS unit costs (see *Table 33* for costs).

e Upper limit for staffing unit costs, lower limit for re-admission/LOS unit costs (see *Table 33* for costs).

made (high staff cost and low costs for days of stay and re-admissions averted), cost per QALY was 15% higher (£3153). The major sensitivity was the assumption about the quality-adjusted life expectancy of lives saved. If it was assumed that the expectancy was lower than in the base case (mean expectancy), matching that of the (approximately) 50% of the cohort of deaths who were 80 years old or older, the cost per QALY was 97% higher (£5389).

To estimate the effect of more targeted reductions in understaffing, we considered alternative scenarios: avoiding some, but not all, understaffing and focusing on subgroups (*Table 25*). Avoiding only 50% of understaffing cost less per patient, but was similarly cost-effective to avoiding all understaffing, with reduced benefits in terms of lives saved (and other outcomes). Focusing efforts to avoid low staffing on only the most acute patients (those with NEWS of at least five) cost more per QALY gained (£9848) than the base case (£2778), although different populations make these estimates hard to compare directly. Focusing on only less acute patients (those with NEWS < 5) cost £3820 per QALY gained. For older people's wards, the estimated effect of low staffing on mortality was lower than in the population as a whole and the estimated cost per QALY gained was higher at £11,416.

If only RN understaffing was avoided, the staff cost was £3532 per QALY gained. While staff costs from avoiding only NA staffing were lower (£1626 per QALY gained), the relative advantage reduced when net staff costs (considering sickness absence) and re-admissions were considered, because of the adverse effects of higher assistant staffing (i.e. avoided low staffing) on these outcomes. Similarly, while avoiding low RN staffing for highly acute patients was cost-effective at similar levels to the base case (£2701 per QALY gained), avoidance of low assistant understaffing was the only scenario where there was an adverse effect from higher staffing levels, and so the decision not to intervene dominates.

We considered the costs and consequences of using temporary staff to avoid staffing shortfalls, focusing on agency staff because these are potentially more expensive, and the adverse effects of temporary staffing were either similar or worse than for bank staff. We considered the adverse effect of a higher proportion of agency staff on mortality based on the mean levels of understaffing observed on days of low staffing and combined this with the estimated benefit of removing low staffing to give an attenuated estimate. For both RN and NA staff, the mean staffing shortfall was approximately 17% and so we used these figures to estimate the proportion of agency staff that would be observed using agency staff to avoid low staffing still reduced the risk of death, but by a lower amount, with 62% fewer lives saved if low staffing was rectified using agency staff as opposed to permanently employed staff. If agency staff are assumed to cost the same as permanent staff, the cost per QALY was £7320. As prices paid for agency staff can be highly variable, we estimated various proportional increases. If agency staff cost 150% of the cost of substantive staff, the cost per life saved was £10,980. If agency staff cost 200% of the cost of substantive staff, the cost per life saved was £14,435. If no assumption of reduced effectiveness from agency staff is made, the cost-effectiveness of using agency staff to rectify staffing shortfalls is more favourable and, in the extreme (no cost difference, no adverse effects), it is the same as the base case above (*Table 26*).

Summary of main results

- For adult inpatients, exposure to days with lower-than-expected RNs or NA staff was associated with increased hazard of death (aHR 1.08/1.07, 95% CI 1.07 to 1.09/1.06 to 1.08) and LOSs.
- Low RN staffing was also associated with increased hazard of re-admission (aHR 1.01, 95% CI 1.01 to 1.02).
- Results for paediatric admissions were similar, although low NA staffing was also significantly associated with increased risk ofre-admission.
- The effects of low staffing on general wards and for less acute patients (NEWS < 5) only were similar to original results.
- Effects in highly acute patients (NEWS 5+), and older people's wards were smaller. For highly acute patients, there was evidence of an adverse effect from NA staffing as low staffing was associated with significant decreases in the hazard of death (aHR 0.98, 95% CI 0.96 to 1.00) and increased risk of re-admission (aHR 0.97, 95% CI 0.95 to 0.99).
- NA low staffing was also associated with lower hazard of re-admission for people over 75.
- Days of low staffing were also associated with increased risk of the potentially nurse-sensitive adverse events of DVT, pneumonia and PUs in surgical admissions. Low RN staffing had a larger effect than low assistant staffing.

TABLE 25 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing for patient/ward subgroups

Scenario	Lives saved	QALY gains	Additional staff cost per admission (£)	Additional staff cost per life saved (£)	Additional staff cost per QALY (£)	Net additional cost per QALY gained (sickness costed) (£)[a]	Net additional cost per QALY gained (sickness and re-admissions costed) (£)[b]	Net additional cost per QALY gained (sickness, re-admissions and LOS costed) (£)[c]
Remove 50% understaffing	3356	22,875	99	18,405	2701	2663	2611	−5110
Remove RN understaffing only	3627	24,717	139	24,071	3532	3459	3257	−4205
Remove NA understaffing only	3272	22,303	58	11,080	1626	1629	1747	−5819
High NEWS subgroup[d]	141	926	311	64,786	9848	7986	8361	−16,410
High NEWS subgroup – remove RN understaffing only[d]	333	2190	202	17,769	2701	1880	1739	−4115
High NEWS subgroup – remove NA understaffing only[d]	−205	−1346	109	−15,649	−2379	−2435	−2932	5974
Lower NEWS subgroup[e]	4902	32,577	285	25,388	3820	3767	3710	−2922
Care of older people ward subgroup	384	2617	427	77,806	11,416			

a Costs of avoided sickness absence removed.
b Costs of avoided sickness absence and re-admissions removed.
c Costs of avoided sickness absence, re-admissions and days of stay removed.
d NEWS at least 5.
e NEWS below 5.

TABLE 26 Costs, consequences and cost-effectiveness (incremental cost-effectiveness ratios) of eliminating low staffing with agency staff (base-case and sensitivity analysis)

Assumptions	Lives saved	QALY gains	Additional staff cost per admission (£)	Additional staff cost per life saved (£)	Additional staff cost per QALY (£)	Net additional cost per QALY gained (sickness costed) (£)[a]	Net additional cost per QALY gained (sickness and re-admissions costed) (£)[b]	Net additional cost per QALY gained (sickness, re-admissions and LOS costed) (£)[c]
Agency staff less effective than permanent								
Agency staff cost same as substantive staff[d]	2477	16,879	197	49,887	7320	7218	7075	−12,460
Higher agency cost (125% of substantive staff cost)[d]	2477	16,879	247	62,358	9150	9022	8879	−10,656
Higher agency cost (150% of substantive staff cost)[d]	2477	16,879	296	74,830	10,980	10,826	10,684	−8851
Higher agency cost (200% of substantive staff cost)[d]	2477	16,879	395	99,773	14,639	14,435	14,292	−5243
Agency staff same effect as permanent								
Higher agency cost (125% of substantive staff cost)	6527	44,483	247	23,663	3472	3424	3369	−4043
Higher agency cost (150% of substantive staff cost)	6527	44,483	296	28,395	4166	4108	4054	−3359
Higher agency cost (200% of substantive staff cost)	6527	44,483	395	37,860	5555	5478	5423	−1989

a Costs of avoided sickness absence removed.
b Costs of avoided sickness absence and re-admissions removed.
c Costs of avoided sickness absence, re-admissions and days of stay removed.
d Mortality effect adjusted for use of agency.

- Exposure to days of low RN staffing increased the odds of sickness absence for both RNs and NAs (adjusted OR 1.02/1.03, 95% CI 1.00 to 1.03/1.01 to 1.04 for each 10% of the past 7 days worked that were understaffed).
- A higher proportion of band 6+ RNs and band 4 NAs were associated with reduced hazard of death (aHR 0.98), although the result was not statistically significant for NAs nor when all staff-mix variables were considered simultaneously.
- Higher proportions of bank and agency staff (both RNs and NAs) were associated with increased hazard of death, with the strongest adverse effect associated with agency assistant staff.

Economic analysis

- Eliminating low staffing cost £2778 per QALY gained. Savings from avoided sickness absence and re-admissions made cost-effectiveness estimates more favourable but did not have a major impact.
- If costs of avoided hospital stays are included, avoiding low staffing generates a net cost saving under all scenarios modelled except reducing understaffing by NAs for highly acute patients, which led to a net cost increase and worse outcomes.
- Avoidance of RN rather than assistant understaffing for highly acute patients gave more benefits and was more cost-effective.
- Avoiding low staffing using temporary agency staff reduced mortality but was less cost-effective than using substantive staff because benefits were reduced, and costs increased.

Chapter 6 Discussion and conclusions

Summary of main findings

In this study, we set out to add to the limited body of economic evidence to guide decisions about staffing to provide nursing care on hospital wards. Noting that evidence showing associations between high workloads for nurses and adverse outcomes has rarely considered staffing by other staff groups, we undertook cross-sectional studies exploring associations between workloads, expressed as staff per occupied bed, and mortality, patient experience and staff experience. In these studies, staffing and outcomes were assessed annually at the hospital Trust level. We found that Trusts with more occupied beds per doctor and AHP had higher mortality rates than expected, given the case mix. An association between high RN workloads and high mortality was reduced and became nonsignificant in models including other staff groups. We found that Trusts with higher levels of support staff for nurses and AHPs had higher mortality. Higher workloads for RNs were associated with reductions in the quality of patient experience, staff-reported health and well-being, and quality of care. Higher nurse staffing levels (measured as HPPD) on wards were associated with reduced risk of avoidable patient harms.

In our longitudinal patient-level study in four hospital Trusts, we considered the nurse staffing patients directly experienced during their hospital stays. This study aimed to avoid the potential biases of hospital-level cross-sectional studies where staffing levels by different staff groups are strongly correlated and provided a stronger basis for modelling the economic costs and consequences of staffing deficits. We found that when patients are exposed to days of low staffing on inpatient wards, from either RNs or NA staff, the hazard of death is increased, and the hospital stay is longer. Exposure to low RN staffing also increased the risk of re-admission. We found similar results when looking at patient subgroups, considering general ward staffing only and for less highly acute patients. Results in paediatric patients were also broadly similar, although mortality effects were not significant. For the most highly acute patients, there was evidence of harm from higher staffing from assistants, because days where patients experienced lower than average staffing from NA staff were associated with reduced hazard of death and re-admission. We found that greater use of temporary staff was associated with increased hazard of death, while higher proportions of senior grade staff reduced the hazard, although this was not statistically significant in a model considering multiple staff-mix factors. There was also evidence of additional benefits of RN staffing at higher levels than the threshold we used, which are based on current ward mean averages. Low staffing was also associated with increased risk of DVT, pneumonia and PUs in surgical patients, with low RN staffing having a stronger adverse effect than assistant staffing. We found that incident reports were more frequent when staffing levels were higher. RN understaffing was associated with increased risk of sickness absence for both RNs and NAs.

We estimated that the cost of staff required to eliminate days of low staffing was £197 per patient per day and in our base-case scenario the cost per QALY gained was £2788. Reductions in staff sickness absence and avoided re-admissions reduced net costs by a small amount (2% or less), but the value of bed days saved far exceeded the costs of additional staff. While focusing on RN understaffing only was somewhat less cost-effective (£3459 per QALY gained), there were more benefits that arose in terms of reduced sickness absence, fewer re-admissions, reduced mortality for the sickest patients and fewer adverse events (associated with avoiding low staffing from RNs compared to NAs). Using agency staff to avoid low staffing was less cost-effective than using substantive staff, but the effect on mortality was still protective. If the cost of agency staff was 50% more than that of substantive staff, the cost per QALY was £10,980.

Relationship to existing research

Our cross-sectional study adds to a small body of research that explores associations between staffing levels and outcomes from staff groups other than nurses.[103] While the association between high medical workloads and reduced patient safety has been shown previously, ours is the first study to show a relationship between AHP staffing and safety. Most cross-sectional studies addressing nursing workloads have given scant consideration to other staff groups, but our results show that estimates of nurse staffing effects from such studies could be biased because nurse staffing levels are correlated with other staff groups. This in turn has implications for economic analyses examining the benefits of increased nurse staffing because the benefits of increased nurse staffing could be overestimated because of this bias. On the other hand, recent work has shown that associations based on annual aggregate data could bias estimates of nurse staffing effects in the opposite direction, in part because much of the variation experienced by patients is averaged out across a longer time period.[156] Most previous economic studies rely on estimates from such cross-sectional studies.[59]

The results of our longitudinal patient-level study provide granular estimates of effects based on daily exposures to varying staffing levels. The findings are consistent with the limited number of previous studies using similar methods, although direct comparisons are difficult because of the diversity of staffing measures used.[34] In comparison to our previous single-site study using similar methods, where a day of low RN staffing was associated with a 3% increase in the hazard of death,[22] the effects observed in this study were larger. Exposure to a day of low RN staffing was associated with a 6% increase in hazard of death in comparable models where workload associated with transfers was included. Needleman and colleagues found a 3% increase in hazard of death per low staffed shift, which would give a similar daily estimate to our finding.[157]

Our study adds additional evidence on the consequences of low nurse staffing in terms of increased re-admissions, LOS and other potentially nurse-sensitive adverse events. The estimated effect on LOS in this study is relatively large, although similarly large effects have been observed previously.[81] Because most stays are short (median is < 4 days), any substantial extension to one person's stay resulting from adverse events can occupy the same number of bed-days as many 'typical' admissions, which is significant both because of the resources used, but also because of the inability to use the bed for new admissions. While we found harmful effects from low staffing for adverse events recorded in patient records, we found that higher staffing levels were associated with higher levels of incident reporting. The pattern of results was consistent with ascertainment and reporting bias, leading to higher recording when staffing is higher, which has been observed previously in studies using UK incident reporting systems.[158] Our study is the first longitudinal study to demonstrate an adverse effect of low staffing on an important and objectively measured staff outcome, that is sickness absence. While the effect on sickness absence was relatively small (about 2%), other research has identified adverse effects on subjective well-being for nurses who experience understaffing[40] and our cross-sectional study also showed that higher workloads were associated with worse scores for staff health and well-being on the staff experience survey.

We found that the mix of staff was related to patient outcomes. Richer grade mix for RNs (i.e. a higher proportion of band 6+ staff) was associated with improved outcomes, a finding reflected by recent research from the NHS in England, which found that senior RNs were associated with a larger reduction in mortality than more junior nurses.[159] However, in our study, the grade mix of staff was no longer significant when the mix of temporary staff was considered simultaneously. Zaranko and colleagues did not find any benefit from adding agency staff but they found that bank staff had a similar effect in reducing mortality to that of permanent staff.[159] In contrast, our study, which approached this problem differently, found that after controlling for low staffing, both a higher proportion of bank staff and a higher proportion of agency staff had a negative effect on mortality, especially for agency assistant staff. Our previous study also found that higher proportions of temporary staff were associated with increased risk of mortality after controlling for low staffing.[16] These findings are not entirely incompatible, as our economic model indicated that the relative harm from a high proportion of temporary staff was smaller than the beneficial effect of eliminating low staffing. The effect of temporary staff can only be understood in the context of the achieved staffing level because the two effects are independent. The benefits of low staffing avoidance are reduced but not removed by using agency staff. Avoiding low staffing using agency staff is associated with reduced harm even if it is a less effective (and cost-effective) option than using permanent staff.

We found statistical interaction effects between levels of RNs and NAs. Visual inspection of the plots did not indicate that these effects were of substantial material importance, but the cumulative nature of the variable may mask important effects on the day and there are strong theoretical reasons to believe that such effects may exist related to, among other things, the capacity of RNs to properly supervise assistant staff.[160] The adverse effect of a higher proportion of temporary staff was independent of the adverse effect of low staffing and we found no interaction between these two variables.

This is the first study to estimate the QALY gains associated with changes in nurse staffing in a longitudinal study, and the first UK study to consider life expectancies. Thus, our economic results cannot be directly compared with previous studies. In our base case, we estimated that eliminating all understaffing cost £2778 per QALY. In our recent study, we estimated the costs and consequences of a blanket 1-hour-per-patient day increase in RN staffing levels, as opposed to a strategy of eliminating relative low staffing. The cost of this per patient was £219 and the cost per life saved was £47,376, although the finding was highly sensitive to assumptions about the mortality effect (using upper and lower 95% confidence limits as a sensitivity check) in that much smaller study. If the estimate for QALYs gained from this study is applied to that estimate, the cost per QALY (£6946) would be close to some values in scenarios/sensitivity analyses observed here (albeit with costs based on an earlier year) and of a similar scale to estimates based on life expectancies seen in previous economic studies (albeit in different contexts and with varying interventions).[79,80]

There is no clear criterion or threshold for establishing whether an intervention is 'cost-effective'. Per capita GDP is sometimes used, but in a resource-constrained system, consideration must be given to the opportunity costs when considering whether or not the health benefits gained are greater than the health that is likely to be lost because resources are not deployed elsewhere.[32] In the UK context, the NICE identified £10,000 per QALY ($15,572 is the 2021 US$ equivalent) as representing 'exceptional value for money', meaning that a drug could be fast-tracked for availability in the NHS.[161] In almost all cases, our cost-effectiveness estimates for eliminating low staffing were below £7843, the median for public health interventions recommended by NICE between 2011 and 2016.[162] If the value of hospital stays averted is included in the cost estimate, then increased staffing is economically dominant, as net costs are reduced and outcomes improved. However, while there is clearly value in releasing a bed for other uses and an opportunity cost of using a bed to care for someone experiencing an avoidable stay due to an adverse event or organisational inefficiency arising from low staffing, hospital revenues are not necessarily increased within the NHS.

Much previous research and our review of economic evidence have pointed to a conclusion that adding assistant staff as substitutes for RNs is potentially harmful and is not cost-effective. Our findings, in common with some other research, show that low levels of assistant staffing are associated with harm.[22,87,157] However, we do find that there are greater benefits from rectifying RN staffing shortages. Using continuous measures of staffing, we found that staffing higher than the norm improves outcomes further. For RNs (but not assistants), we also found evidence that there may be increasing marginal gains as staffing increases, and we found larger effect sizes when low staffing was measured relative to thresholds higher than the mean. In most cases, beneficial effects from increased RN staffing were higher than those from NA staffing, in particular adverse events. For some patient subgroups, in particular the most acutely ill, there was indication that higher assistant staffing levels might cause harm, consistent with our previous study,[22] although the specific pattern of results observed in our previous study, with a clear increase in harm from higher assistant staffing levels using a continuous staffing measure, was not replicated here. Finally, we identified that assistant workload associated with patient transfers was also associated with mortality risk, but the relationship was much stronger and only statistically significant for RNs, presumably because this work is primarily undertaken by RNs. These findings in tandem with previous research do not support a conclusion that substitution is either safe or effective. However, they do show that NAs represent an important part of care delivery and appropriate staffing levels need to be considered alongside staffing by RNs.

Strengths and limitations relative to other research

In many respects, this study is stronger than previous economic modelling studies, especially in relation to UK policy decision-making. The underlying longitudinal study with patient-level staffing exposures is a far stronger design than cross-sectional studies and is less prone to bias, although it remains an observational study and causal inferences cannot be made from these results in isolation. This is the first multicentre economic study of its type and the first to estimate costs per QALY. QALY gains were estimated, although conclusions about costs per QALY are not sensitive to a range of assumptions about gains, costs or discount rates. We have considered a wider range of outcomes and costs than many previous studies, although data on incidents and adverse events proved limited.

Some previous patient-level studies have been able to define staffing requirements relative to an assessed daily or shift-level demand or expectation based on detailed patient and shift characteristics.[85,97] We did not have these data available for most of our sample, and so we used the ward mean as a reference point. While our previous study demonstrated a close correspondence between the ward mean and the planned staffing level,[19] it is likely that the observed staffing is, on average, lower than the planned establishment, due to vacancies. This need not be a limitation as most other studies have used a threshold somewhere below the planned staffing level to define low staffing, but the uncertainty and likely variation in this threshold between wards is a problem. We had access to planned staffing levels from the data supplied by the NHS England for the national ward-level cross-sectional study. However, access to these data was given late and the analysis suggests that these planned staffing levels may not give a valid reference point for staffing requirements. In the absence of detailed information about how planned CHPPD were determined, it is unclear if the data have value. Given that there is a relationship between RN staffing levels (HPPD) and harms (in models with ward random effect), the absence of any association between RN staffing shortfalls and harms does not give confidence that the planned staffing levels in the NHS England data are accurate measures of staffing requirements, as the shortfall would be expected to provide equal or stronger associations. Additionally, although using the mean as a reference point may have validity, it cannot account for daily variation in demand caused by variation in case mix that is not directly related to mortality risk, which is controlled for in the model.

Challenges faced and other considerations

This study launched in early 2020, just as the first wave of the COVID-19 pandemic occurred in the UK. Trusts made exceptional efforts to participate despite difficult working conditions and competing demands. Although we met our initial target sample of four Trusts, working through complex data governance processes was time-consuming. We had recruited additional Trusts which were unable, ultimately, to find the resources to allow the project to progress, meaning late withdrawals. For these reasons, data transfers were substantially delayed and although the project was extended, we still had less time with a large and immensely complex data set than we had originally planned. Moreover, the complexity of data governance and challenges for resources within Trusts meant that we had to abandon any aspiration to link roster data to other staff characteristics from the electronic staff record. Information from the electronic staff record could enhance analysis of sickness absence by providing information about age and turnover as an outcome. The circumstances under which the research was conducted also limited our planned engagement activities and our opportunities to work as closely with participating Trusts as we would have liked, especially during the final analysis stages, which were compressed. However, we did successfully adopt several creative strategies for engagement as a result of the pandemic. See *Staff and service engagement*.

While we completed analyses as per the protocol, we were unable to explore some issues that arose from the data analysis or exploit several opportunities due to lack of time. Potential future use for the data includes further exploration of temporary staffing effects, including potential nonlinear associations and thresholds for adverse effects (a possibility identified in the results of our previous study[16]) and using patient and shift information to better characterise expected staffing levels. Although we model longitudinal associations with outcomes, which we consider preferable, Zaranko's recent study used within-day associations to identify some potentially important relationships. Our data could be used to replicate this analysis, which would provide valuable confirmation (or otherwise) of their findings about seniority of staff.[159]

Patient, public and service engagement and involvement

As noted in *Chapter 3*, we took an integrated approach to involving patients, the public and staff members because all are potential beneficiaries of the research. The study began concurrently with the beginning of the COVID-19 pandemic, and this impacted on many of our plans for ongoing involvement and engagement. Planned events involving NHS staff had to be cancelled as COVID-19-related work had to be prioritised. Early events with service users and the public had to be conducted online. Despite this, we received significant ongoing engagement from the Trusts that volunteered to provide data for the study, and many members of staff worked hard to progress approvals for a study that would not otherwise have been prioritised in the context of COVID-19 because it was recognised as addressing an issue that was vital for them. Additional information is provided in the Supplementary Engagement and Involvement Report and a summary is given here.

Public engagement

In late spring 2020, we met with a group of experienced lay patient and public participation and engagement contributors who are part of a group involved with the NIHR Applied Research Collaboration (Wessex) Wessex Inclusion in Service Research and Design (WISeRD) Group. FL (the project Patient and Public Involvement and Engagement Lead who has expertise by experience in health services) and lead investigator PG presented the background to this project and shared some previous research results to stimulate discussion. The discussion highlighted a strong sense from participants in the group that staffing in the NHS was a vital topic, but also a belief from some that NAs might be suitable substitutes for RNs. Participants were surprised by results of previous research pointing to adverse effects from higher assistant staffing and were surprised at lack of regulation around staffing numbers. The discussion highlighted the importance of considering the mix of staff within job groups and highlighted a limitation of this study – which is that we are unable to identify experience and background training, although our approach to measuring grade mix may address the issues highlighted.

We discussed the approach we should take for additional engagement with lay representatives on our advisory group who noted that this (WISeRD) group had very particular experience – generally they were experienced in patient and public involvement and engagement (PPIE), older and affluent. We actively sought out a different group of non-experienced lay people and those from a younger generation who may have fresh viewpoints and experiences of the NHS. FL contacted the Head of Health and Social Care modules/courses at a local further education/sixth form college near the University. Eleven students (age range was 16–19 years old) and their tutor attended an event we organised at the University with several members of the research team. The students came from a wide variety of backgrounds and ethnicities, which reflected the demographic of the city well. Some of the group had experience as young carers and most were looking to either go on to university and work in health and social care, or child care.

Francesca Lambert facilitated the session, which began with an interactive presentation from PG, including an introduction to some background about the research and methods used. The session included a quiz with questions to stimulate thinking about safe staffing, which used examples from settings and industries such as the airline industry where safe staffing levels are sometimes mandated. We used Vevox, a live polling app, to gather responses. This meant that those who were more reluctant to speak up were able to engage and it appeared to work as a stimulus for more open conversations. The discussion covered several issues and confirmed that like the experienced PPIE group, participants had concerns about low staffing in the NHS, had expectations that NAs had mandatory training and that there were mandatory minimum staffing levels.

We also discussed the issue of using patient data in the study. It was clear that while there was general support for use of anonymised data for research purposes by university researchers, where there was no profit motive, this agreement was contingent and there was less support for the use of similar data by commercial entities. This accords with the results obtained from larger and more formal consultations with the public about the use of health data.[163]

Staff and service engagement

Planned national launch events were cancelled due to the COVID-19 pandemic. We focused our efforts on wider engagement with the Trusts that we were working with. There were three primary activities. Firstly, we invited all local contacts, including data and systems staff, to participate as observers in our regular advisory/steering group meetings, although attendance from Trusts was limited. The steering group itself included a range of interested parties including representatives from nursing unions and local government.

Secondly, at the suggestion of staff at one Trust who noted that previous experience led them to expect hard work to extract data but no ongoing engagement, we organised a series of seminars to set the research in context and to discuss emerging ideas and priorities. Three seminars were delivered online in 2021 and a fourth was planned but cancelled subsequently due to winter pressures which made it difficult for staff to attend. Thirdly, we produced a series of newsletters and have undertaken one-to-one briefings of emerging results. We drew on this group of interested parties to guide the selection of areas to focus on in the staff experience analysis (see *Outcomes*) where we conducted an online consultation.

Additionally, we undertook one-to-one briefings with senior staff at NHS England, focusing on the AHP workforce (sharing early results from the national cross-sectional study) and the nursing workforce directorates. We worked closely with the workforce directorate and incorporated the expected HPPD into our ward-level analyses of CHPPD at their request because this addressed an issue of concern to them – whether deviation from planned hours provided a better indicator of workforce adequacy.

Equality, diversity and inclusion

Patient safety in acute hospitals is a matter of wide public concern, potentially affecting all members of all communities. Similarly, NHS staff are diverse and so issues affecting them are relevant to a diverse community. The Trusts that participated in our study service diverse communities including significant ethnic minority communities and some areas of high deprivation. The nature of our study is such that we were able to include all patients. While we had no a priori hypotheses or explicitly stated objectives related to discrimination, one effect of low nurse staffing has been described as 'implicit rationing' of care.[164] It is certainly possible that different communities are disproportionately affected by the issues we consider in this report. One potentially vulnerable community that could be affected by such rationing is older people and we have considered people over 75 years old and wards specialising in the care of older people in our analyses. However, we did not have any data that would allow us to explore issues that might disproportionately affect specific communities of patients or staff.

Impact and learning

In this report, we have given an overview of the vast existing body of research exploring associations between nurse staffing levels and outcomes in acute hospitals. We have also demonstrated that estimates of nurse staffing effects arising from cross-sectional studies may suffer from multiple sources of bias, in particular the omission of other staff groups, including AHP staffing, because staffing levels are correlated and staffing by other groups are also associated with quality and patient safety. The added value of further research exploring hospital-level associations is low, particularly if studying nursing in isolation. Increasing access to electronic roster data opens up more opportunities for patient-level analyses, although equivalent data are not widely available for other staff groups and even, where available, the appropriate way to link staff availability and capacity to patients is still unclear. The design of this study, with patient-level exposure to variation in nurse staffing levels, means that there is no intrinsic bias from the omission of other staff groups, although the extent to which within-hospital variation in low staffing across different staff groups is correlated remains unclear. In our study, we modelled the potential sensitivity of results to this omission by considering weekend and seasonal effects, both factors that could foreseeably lead to nurse staffing shortfalls being correlated with lower availability of other staff. Although we found no evidence of bias introduced by this omission, staffing levels from other groups are of interest in their own right, especially as AHP staffing was the most variable across all the staff groups we considered in our cross-sectional study.

Our results on the cost-effectiveness of removing nurse staffing shortages are challenging in the face of ongoing shortages of supply. We have shown for the first time that temporary staff may reduce the adverse effects of short staffing in a cost-effective manner, but addressing underlying staffing shortages with permanently employed staff remains the preferred solution as it is likely to be both less costly and deliver better outcomes. Focusing on RN staffing shortages delivers the best outcomes for patients, although shortages of assistants also present a risk for patients. Our findings did not clearly or directly demonstrate the added value of band 4 assistant staff (which would include nursing associates). Although we did find that a higher proportion of band 4 assistants were associated with lower hazard of death in staff-mix models, this result was not statistically significant. If the value of well-trained staffing in all staff groups were fully accepted, the priority for future research could turn to how best to attract, train and retain recruits into the profession.

This research has revealed limitations in both the available data and methods used to determine staffing requirements for nursing teams on hospital wards. Our previous NIHR-funded project highlighted the limited evidence for the validity and effectiveness of existing tools to determine staffing requirements.[78] More work is required to find methods to estimate staffing requirements relative to patient needs, including approaches that draw on the increasingly rich sources of routine data, both for prospective planning and to monitor real-time staffing adequacy to guide staff deployment. While not formally reviewed recently, there appears to be a dearth of tools for planning allied health staffing capacity and an equally limited evidence base.[165]

Findings from this study about temporary staff provide reassurance that money spent on temporary staff could still deliver a cost-effective reduction in harm from short-staffing and real-time systems can potentially be used to guide flexible staffing deployments to target areas of greatest need. However, results of this study and previous research do show that staffing solutions based on permanent staff are likely to be more cost-effective.[15,16] Further research is needed to define acceptable levels of temporary staff usage and to determine optimal approaches to deployment. The trade-offs between benefits and harms associated with different levels of temporary staff require further exploration.

Implications for decision-makers

The NHS faces multiple competing demands for scarce resources. The evidence presented here suggests that investment in nurse staffing in acute hospitals could be a cost-effective one, on a par with many public health interventions. Investment to avoid low staffing would be a dominant strategy from a standpoint of health-economic decision-making if the benefits of reduced LOS are considered, because it delivers improved outcomes at reduced cost, but this is only the case if the economic benefits of reduced stays are realised. The extent to which the NHS can realise the benefit is constrained, but shorter stays in a system where demand is high increases bed availability which can, in turn, improve flow through emergency departments or improve bed availability for elective surgery. The possibility of a net gain to the system should not be discounted.

The relative increase in staff costs to avoid low staffing is modest, although the supply of staff to meet demand remains challenging. It is important that this scarcity does not obscure the need for and demonstrated value for money in improving RN staffing levels. While decision-makers may, of necessity, need to experiment with novel approaches to addressing staffing shortages, this needs to be done in the context of a full understanding of what is already known about the consequences of changing staffing levels and configurations. The safety and cost-effectiveness of alternatives should not be assumed.

Our findings indicate that understaffing from both RNs and NAs presents a significant risk to patients. Scarcity of RNs and lower costs for employing assistants should not be used to justify reliance on temporary staff, prioritising assistant staffing or to support substitution. The evidence presented in this report shows that harms associated with RN understaffing are generally greater than those associated with assistant understaffing, and we found evidence that harms extend across outcomes including re-admissions, adverse events and staff sickness in addition to increased mortality risk. Higher use of temporary staff was associated with increased risk and although a fully staffed shift using temporary staff had a lower risk than an understaffed shift, risk remained elevated. Use of agency assistant staff to manage staffing shortfalls was not an effective strategy, leaving the harms associated with short staffing unaltered. While we did not find consistent evidence of harms associated with increased assistant staffing, this contrasts with other research. We also saw evidence that if low staffing was defined relative to a higher threshold than current norms, the benefits from avoiding low RN staffing increase. Therefore, the priority for decision-makers should be to address shortages of RNs and to do so with permanent staff. We found that the clearest evidence of adverse effects of low staffing was in the general population and the subgroup of lower acuity populations, which indicates limited scope for targeted interventions or remedies for low staffing.

Research recommendations

Several priorities for future research emerge from this work:

- More research is needed into methods to determine nurse staffing requirements on hospital wards, for planning, real-time monitoring and for use in research. The requirements of service should inform decisions about the required timeliness of data, acceptable data gathering load and the necessary precision.
- Our findings, combined with the results of previous research, leave uncertainty about the trade-offs between staff shortages and temporary staffing levels, including the relative (adverse) effects of temporary staff at different levels and from different sources. Both qualitative and quantitative research would be of value here.
- There remains uncertainty about the interaction between RN and NA staffing levels which should be addressed through both qualitative and quantitative research.
- Research is required to better understand whether the observed variation in AHP staffing is based on variation in service and patient need. The observed association with mortality rates in this study suggests it may not be, and, if that is the case, evidence-based methods for determining appropriate staffing levels need to be developed.

Conclusions

Our results not only show the adverse effects of low nurse staffing but also show that medical and AHP staffing are important considerations for patient safety. Eliminating low RN staffing gives more benefits than eliminating assistant staffing, but both interventions are cost-effective in terms of QALYs gained relative to many public health interventions. These findings lend support to policy initiatives aimed at increasing the supply of RNs.

Additional information

CRediT contribution statement

Peter Griffiths (https://orcid.org/0000-0003-2439-2857): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Visualisation, Writing – original draft (lead), Writing – reviewing and editing.

Christina Saville (https://orcid.org/0000-0001-7718-5689): Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – reviewing and editing.

Jane Ball (https://orcid.org/0000-0002-8655-2994): Conceptualisation, Funding acquisition, Investigation, Writing – reviewing and editing.

David Culliford (https://orcid.org/0000-0003-1663-0253): Formal analysis, Funding acquisition, Investigation, Methodology, Writing – reviewing and editing.

Jeremy Jones (https://orcid.org/0000-0002-2725-0937): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – reviewing and editing.

Francesca Lambert (https://orcid.org/0000-0003-0327-4325): Conceptualisation, Funding acquisition, Investigation, Methodology, Project administration, Writing – reviewing and editing.

Paul Meredith (https://orcid.org/0000-0002-5464-371X): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Visualisation, Writing – reviewing and editing.

Bruna Rubbo (https://orcid.org/0000-0002-1629-8601): Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – reviewing and editing.

Lesley Turner (https://orcid.org/0000-0003-1489-3471): Formal analysis, Investigation, Writing – reviewing and editing.

Chiara Dall'Ora (https://orcid.org/0000-0002-6858-3535): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – reviewing and editing.

Paul Schmidt: Conceptualisation, Funding acquisition.

Andrew Barraclough: Resources.

Louise Bramley: Resources.

Greg Bull: Resources.

Keith Elkin: Resources.

Liam Flynn: Resources.

Fiona Hyett: Resources.

Liz Rix: Resources.

Carol Stiles: Resources.

Rachel Taylor: Resources.

Acknowledgements

We would like to express our gratitude to the following:

- Paul Schmidt, Portsmouth Hospitals University NHS Trust

Paul Schmidt was a co-applicant and involved in the development of the project. His ongoing involvement in the project was limited due clinical duties linked to the COVID-19 pandemic.

Project Steering Group

- Andy Charlwood – Chair, University of Leeds
- Chris Bojke, University of Leeds
- Lara Carmona (formerly of RCN)
- Dave Shields – Patient and public representative
- Jonathan Drennan, University College Dublin
- Jo Fillingham (formerly NHS England)
- Oliver Redfern, University of Oxford
- Karen Swinson (formerly NHS England)
- Neil Tape (formerly University Hospital Southampton NHS Foundation Trust)

Advisory Group and our Trust partners

- Andrew Barraclough, Nottingham University Hospitals NHS Trust
- Louise Bramley, Nottingham University Hospitals NHS Trust
- Liam Flynn (formerly Nottingham University Hospitals NHS Trust)
- Fiona Hyett, Salisbury NHS Foundation Trust
- Liz Rix, Portsmouth Hospitals University NHS Trust
- Rachel Taylor- University Colleges London NHS Foundation Trust
- Carol Stiles (formerly University Colleges London NHS Foundation Trust) – NHS England

Others

- Greg Bull, Oceansblue
- Keith Elkin, Oceansblue
- Liz Price, Curriculum Manager, Richard Taunton Sixth Form College
- The Health and Social Care students and the ARC Wessex PPIE group participants

Patient data statement

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review but is subject to agreement from the organisations who supplied the original data and our data-sharing agreements with them.

Ethics statement

We used the screening questions for the Health Research Authority integrated research application service to determine the required approvals. We determined that the study was exempt from the requirement for approval by the National Research Ethics Service. The analyses of national routine data involved aggregate data that was publicly available and so no ethical approval was required. Our longitudinal study using patient level records was assessed as exempt from NHS research ethics approval as all data was functionally anonymised with no opportunity to reidentify any living person from the data shared.

Health Research Authority approval was granted to undertake the research in NHS Trusts (IRAS ref 273185). Additionally, as a requirement of the University of Southampton we sought ethical review from the University's system and were granted approval (ERGO 52957, approved 18 January 2020).

Information governance statement

The University of Southampton is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679. Under the Data Protection legislation, the University of Southampton is the Data Controller, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here: https://www.southampton.ac.uk/about/governance/information-publications/data-protection-information-breach This study did not use personal data. Within the Trusts, pseudonymisation was used in the preparation of data sets where different sources had to be linked but the data transferred to the research team was functionally anonymous because we had no access to pseudonymisation keys. A similar approach was taken with workforce data, where individual records of shifts worked were linked via a staff identifier which was then transformed with no keys transferred to the research team or retained at source. We minimised the data requested to ensure that any residual risk of reidentification was eliminated. Hence, we asked for age bands, not exact age or date of birth and we did not obtain any additional information such as area of residence or measures of deprivation. All procedures and data flows were subject to data protection impact assessment by the participating Trusts and the University of Southampton and authorised by the Caldicot guardians in the Trusts. All data were stored on secure research file storage at the University of Southampton with access limited to research team members. Data processing occurred on university-owned and -managed computers with password protection and encryption. Analysis files were temporarily stored on secure cloud services with high levels of encryption in transit and at rest.

Disclosure of interest

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/ZBAR9152.

Primary conflicts of interest: Peter Griffiths declares funding from the European Union Horizon 2020 programme National Institute for Health Research and the Health Research Board for Ireland for projects related to this. He is a member of the NHS England Safe Staffing Faculty Steering Group, for which he received no remuneration.

Christina Saville declares funding from the National Institute for Health Research and for projects related to this.

Jane Ball declares funding from the European Union Horizon 2020 programme National Institute for Health Research and the Health Research Board for Ireland for projects related to this.

David Culliford declares no conflicting interests.

Jeremy Jones declares no conflicting interests.

Francesca Lambert declares no conflicting interests.

Paul Meredith declares funding from the National Institute for Health Research and for projects related to this.

Bruna Rubbo declares no conflicting interests.

Lesley Turner declares no conflicting interests.

Chiara Dall'Ora declares funding from the National Institute for Health Research and for projects related to this.

Publications

Turner L, Griffiths P, Kitson-Reynolds E. Midwifery and nurse staffing of inpatient maternity services – a systematic scoping review of associations with outcomes and quality of care. *Midwifery* 2021;**103**:103118. https://doi.org/10.1016/j.midw.2021.103118

Dall'Ora C, Saville C, Rubbo B, Turner L, Jones J, Griffiths P. Nurse staffing levels and patient outcomes: a systematic review of longitudinal studies. *Int J Nurs Stud* 2022;**134**:104311.

Turner L, Ball J, Culliford D, Kitson-Reynolds E, Griffiths P. Exploring the relationship between women's experience of postnatal care and reported staffing measures: an observational study. *PLOS ONE* 2022;**17**:e0266638.

Turner L, Culliford D, Ball J, Kitson-Reynolds E, Griffiths P. The association between midwifery staffing levels and the experiences of mothers on postnatal wards: cross sectional analysis of routine data. *Women Birth* 2022;**35**:e583–9.

Dall'Ora C, Rubbo B, Saville C, Turner L, Ball J, Ball C, Griffiths P. The association between multi-disciplinary staffing levels and mortality in acute hospitals: a systematic review. *Hum Resour Health* 2023;**21**:30.

Griffiths P, Saville C, Ball J, Dall'Ora C, Meredith P, Turner L, Jones J. Costs and cost-effectiveness of improved nurse staffing levels and skill mix in acute hospitals: a systematic review. *Int J Nurs Stud* 2023;**147**:104601.

Rubbo B, Saville C, Dall'Ora C, Turner L, Jones J, Ball J, *et al.* Staffing levels and hospital mortality in England: a national panel study using routinely collected data. *BMJ Open* 2023;**13**:e066702.

Griffiths P, Saville C, Ball J, Culliford D, Jones J, Lambert F, *et al.*; Workforce Health Outcomes Study Group. Nursing Team Composition and Mortality Following Acute Hospital Admission. *JAMA Netw Open* 2024;**7**:e2428769.

Meredith P, Turner L, Saville C, Griffiths P. Nurse understaffing associated with adverse outcomes for surgical admissions. *Br J Surg* 2024;**111**:znae215.

Turner L, Ball J, Meredith P, Kitson-Reynolds E, Griffiths P. The association between midwifery staffing and reported harmful incidents: a cross-sectional analysis of routinely collected data. *BMC Health Serv Res* 2024;**24**:391.

Turner LY, Saville C, Ball J, Culliford D, Dall'Ora C, Jones J, *et al.* Inpatient midwifery staffing levels and postpartum readmissions: a retrospective multicentre longitudinal study. *BMJ Open* 2024;**14**:e077710.

References

1. NHS Digital. *NHS Vacancy Statistics England April 2015–March 2022 Experimental Statistics*. NHS Digital; 2022. URL: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-vacancies-survey/april-2015---march-2022-experimental-statistics (accessed 6 July 2022).
2. Ball J, Barker H, Griffiths P, Jones J, Lawless J. *Implementation, Impact and Costs of Policies for Safe Staffing in Acute Trusts*. Southampton: University of Southampton; 2019.
3. Buchan J, Ball J, Shembavnekar N, Charlesworth A. *Building the NHS Nursing Workforce in England. Workforce Pressure Points*. London: Health Foundation; 2020.
4. Griffiths P, Ball J, Drennan J, Dall'Ora C, Jones J, Maruotti A, *et al*. 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* 2016;**63**:213–25. https://doi.org/10.1016/j.ijnurstu.2016.03.012
5. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office; 2013.
6. Keogh B. *Review into the Quality of Care and Treatment Provided by 14 Hospital Trusts in England: Overview Report*. NHS; 2013.
7. National Quality Board. *Supporting NHS Providers to Deliver the Right Staff, with the Right Skills, in the Right place at the Right Time – Safe Sustainable and Productive Staffing*. NHS England; 2016. URL: www.england.nhs.uk/publication/national-quality-board-guidance-on-safe-staffing/ (accessed 28 November 2024).
8. National Institute for Health and Clinical Excellence. *Safe Staffing for Nursing in Adult Inpatient Wards in Acute Hospitals*. London: NICE; 2014.
9. Aiken LH, Sloane DM, Bruyneel L, Van den Heede K, Griffiths P, Busse R, *et al*.; RN4CAST consortium. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. *Lancet* 2014;**383**:1824–30. https://doi.org/10.1016/S0140-6736(13)62631-8
10. Mitchell BG, Gardner A, Stone PW, Hall L, Pogorzelska-Maziarz M. Hospital staffing and health care-associated infections: a systematic review of the literature. *Jt Comm J Qual Patient Saf* 2018;**44**:613–22. https://doi.org/10.1016/j.jcjq.2018.02.002
11. Patrician PA, Loan L, McCarthy M, Fridman M, Donaldson N, Bingham M, Brosch LR. The association of shift-level nurse staffing with adverse patient events. *J Nurs Adm* 2011;**41**:64–70. https://doi.org/10.1097/NNA.0b013e31820594bf
12. Aiken LH, Sloane DM, Ball J, Bruyneel L, Rafferty AM, Griffiths P. Patient satisfaction with hospital care and nurses in England: an observational study. *BMJ Open* 2018;**8**:e019189. https://doi.org/10.1136/bmjopen-2017-019189
13. Griffiths P, Recio-Saucedo A, Dall'Ora C, Briggs J, Maruotti A, Meredith P, *et al*.; Missed Care Study Group. The association between nurse staffing and omissions in nursing care: a systematic review. *J Adv Nurs* 2018;**74**:1474–87. https://doi.org/10.1111/jan.13564
14. Aiken LH, Sermeus W, Van den Heede K, Sloane DM, Busse R, McKee M, *et al*. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. *BMJ* 2012;**344**:e1717. https://doi.org/10.1136/bmj.e1717
15. Griffiths P, Saville C, Ball JE, Jones J, Monks T; Safer Nursing Care Tool study team. Beyond ratios – flexible and resilient nurse staffing options to deliver cost-effective hospital care and address staff shortages: a simulation and economic modelling study. *Int J Nurs Stud* 2021;**117**:103901. https://doi.org/10.1016/j.ijnurstu.2021.103901

16. Dall'Ora C, Maruotti A, Griffiths P. Temporary staffing and patient death in acute care hospitals: a retrospective longitudinal study. *J Nurs Scholarsh* 2020;**52**:210–6. https://doi.org/10.1111/jnu.12537

17. Griffiths P, Ball J, Murrells T, Jones S, Rafferty AM. Registered nurse, healthcare support worker, medical staffing levels and mortality in English hospital trusts: a cross-sectional study. *BMJ Open* 2016;**6**:e008751. https://doi.org/10.1136/bmjopen-2015-008751

18. Carter P. *Operational Productivity and Performance in English NHS Acute Hospitals: Unwarranted Variations*. London: Department of Health; 2016.

19. Griffiths P, Ball J, Bloor K, Böhning D, Briggs J, Dall'Ora C, *et al*. Nurse staffing levels, missed vital signs and mortality in hospitals: retrospective longitudinal observational study. *Health Soc Care Deliv Res* 2018;**6**:1–120. https://doi.org/10.3310/hsdr06380

20. Aiken LH, Sloane D, Griffiths P, Rafferty AM, Bruyneel L, McHugh M, *et al*.; RN4CAST Consortium. Nursing skill mix in European hospitals: cross-sectional study of the association with mortality, patient ratings, and quality of care. *BMJ Qual Saf* 2017;**26**:559–68. https://doi.org/10.1136/bmjqs-2016-005567

21. Bruyneel L, Li B, Ausserhofer D, Lesaffre E, Dumitrescu I, Smith HL, *et al*. Organization of hospital nursing, provision of nursing care, and patient experiences with care in Europe. *Med Care Res Rev* 2015;**72**:643–64. https://doi.org/10.1177/1077558715589188

22. Griffiths P, Maruotti A, Recio Saucedo A, Redfern OC, Ball JE, Briggs J, *et al*.; Missed Care Study Group. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. *BMJ Qual Saf* 2019;**28**:609–17. https://doi.org/10.1136/bmjqs-2018-008043

23. Bridges J, Griffiths P, Oliver E, Pickering RM. Hospital nurse staffing and staff-patient interactions: an observational study. *BMJ Qual Saf* 2019;**28**:706–13. https://doi.org/10.1136/bmjqs-2018-008948

24. Nursing and Midwifery Council. *The NMC register 1 April 2021–31 March 2022*. London: Nursing and Midwifery Council; 2022.

25. NHS Digital. *Summary Hospital-level Mortality Indicator (SHMI) – Deaths Associated with Hospitalisation, England*. 2023. URL: https://digital.nhs.uk/data-and-information/publications/clinical-indicators/shmi (accessed 15 August 2023).

26. Avalere Health LLC. *Optimal Nurse Staffing to Improve Quality of Care and Patient Outcomes: Executive Summary*. Silver Spring, MD: American Nurses Association; 2015.

27. Buchan J, Seccombe I, O'May F. *Safe Staffing Levels – A National Imperative. The UK Nursing Labour Market Review 2013*. London: The Royal College of Nursing; 2013.

28. Schiff G, Shojania KG. Looking back on the history of patient safety: an opportunity to reflect and ponder future challenges. *BMJ Qual Saf* 2022;**31**:148–52. https://doi.org/10.1136/bmjqs-2021-014163

29. Health Education England. *Facing the Facts, Shaping the Future, a Draft Health and Care Workforce Strategy for England to 2027*. Public Health England; 2017. URL: www.hee.nhs.uk/our-work/workforce-strategy (accessed 28 November 2024).

30. NHS Digital. *NHS Workforce Statistics, September 2017, Provisional Statistics*. 2017. URL: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/nhs-workforce-statistics-september-2017-provisional-statistics (accessed 28 August 2018).

31. Curtis L, Burns A. *Unit Costs of Health and Social Care 2017*. Canterbury: PSSRU, University of Kent; 2017.

32. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al*. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess* 2015;**19**:1–503, v–vi. https://doi.org/10.3310/hta19140

33. Cayton H. Mandating staffing levels is not the answer to reducing poor care. *Health Serv J (Online)* 2012. URL: www.hsj.co.uk/comment/mandating-staffing-levels-is-not-the-answer-to-reducing-poor-care/5042418.article (accessed 29 April 2025).

34. Dall’Ora C, Saville C, Rubbo B, Turner L, Jones J, Griffiths P. Nurse staffing levels and patient outcomes: a systematic review of longitudinal studies. *Int J Nurs Stud* 2022;**134**:104311. https://doi.org/10.1016/j.ijnurstu.2022.104311

35. Rae PJL, Pearce S, Greaves PJ, Dall'Ora C, Griffiths P, Endacott R. Outcomes sensitive to critical care nurse staffing levels: a systematic review. *Intensive Crit Care Nurs* 2021;**67**:103110. https://doi.org/10.1016/j.iccn.2021.103110

36. Blume KS, Dietermann K, Kirchner-Heklau U, Winter V, Fleischer S, Kreidl LM, *et al*. Staffing levels and nursing-sensitive patient outcomes: umbrella review and qualitative study. *Health Serv Res* 2021;**56**:885–907. https://doi.org/10.1111/1475-6773.13647

37. Shin S, Park JH, Bae SH. Nurse staffing and nurse outcomes: a systematic review and meta-analysis. *Nurs Outlook* 2018;**66**:273–82. https://doi.org/10.1016/j.outlook.2017.12.002

38. Penoyer DA. Nurse staffing and patient outcomes in critical care: a concise review. *Crit Care Med* 2010;**38**:1521–8; quiz 1529. https://doi.org/10.1097/CCM.0b013e3181e47888

39. Kane RL, Shamliyan TA, Mueller C, Duval S, Wilt TJ. The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. *Med Care* 2007;**45**:1195–204. https://doi.org/10.1097/MLR.0b013e3181468ca3

40. Dall’Ora C, Ball J, Reinius M, Griffiths P. Burnout in nursing: a theoretical review. *Hum Resour Health* 2020;**18**:41. https://doi.org/10.1186/s12960-020-00469-9

41. Wilson S, Bremner A, Hauck Y, Finn J. The effect of nurse staffing on clinical outcomes of children in hospital: a systematic review. *Int J Evid Based Healthc* 2011;**9**:97–121. https://doi.org/10.1111/j.1744-1609.2011.00209.x

42. Shekelle PG. Nurse-patient ratios as a patient safety strategy: a systematic review. *Ann Intern Med* 2013;**158**:404–9. https://doi.org/10.7326/0003-4819-158-5-201303051-00007

43. Driscoll A, Grant MJ, Carroll D, Dalton S, Deaton C, Jones I, *et al*. 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* 2018;**17**:6–22. https://doi.org/10.1177/1474515117721561

44. Shin S, Park JH, Bae SH. Nurse staffing and hospital-acquired conditions: a systematic review. *J Clin Nurs* 2019;**28**:4264–75. https://doi.org/10.1111/jocn.15046

45. West E, Mays N, Rafferty AM, Rowan K, Sanderson C. Nursing resources and patient outcomes in intensive care: a systematic review of the literature. *Int J Nurs Stud* 2009;**46**:993–1011. https://doi.org/10.1016/j.ijnurstu.2007.07.011

46. Fogg C, Bridges J, Meredith P, Spice C, Field L, Culliford D, Griffiths P. 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* 2021;**50**:431–9. https://doi.org/10.1093/ageing/afaa133

47. Friese CR, Kalisch BJ, Lee KH. Patterns and correlates of missed nursing care in inpatient oncology units. *Cancer Nurs* 2013;**36**:E51–7. https://doi.org/10.1097/NCC.0b013e318275f552

48. Ozdemir BA, Karthikesalingam A, Sinha S, Poloniecki JD, Vidal-Diez A, Hinchliffe RJ, *et al*. Association of hospital structures with mortality from ruptured abdominal aortic aneurysm. *Br J Surg* 2015;**102**:516–24. https://doi.org/10.1002/bjs.9759

49. Bray BD, Ayis S, Campbell J, Cloud GC, James M, Hoffman A, *et al*. Associations between stroke mortality and weekend working by stroke specialist physicians and registered nurses: prospective multicentre cohort study. *PLOS Med* 2014;**11**:e1001705.

50. Beech J, Bottery S, Charlesworth A, Evans H, Gershlick B, Hemmings N, *et al*. *Closing the Gap: Key Areas for Action on the Health and Care Workforce*. London: The Health Foundation/Nuffield Trust/The King’s Fund; 2019.

51. Twigg DE, Myers H, Duffield C, Giles M, Evans G. Is there an economic case for investing in nursing care – what does the literature tell us? *J Adv Nurs* 2015;**71**:975–90. https://doi.org/10.1111/jan.12577

52. Saville C, Griffiths P, Ball J, *et al*. Economic evaluations of alternative nurse staffing configurations: a systematic review of the literature. *PROSPERO* 2020. URL: www.crd.york.ac.uk/PROSPERO/view/CRD42021281202 (accessed 29 April 2025).
53. Dall'Ora C, Griffiths P, Rubbo B, *et al*. What is the association between nurse and midwife staffing levels and patient outcomes, quality of care and resource use? A systematic review and meta-analysis of longitudinal studies. *PROSPERO* 2020:CRD42020191798.
54. Rubbo B, Griffiths P, Saville C, *et al*. Associations between healthcare professional staffing levels and mortality in acute hospitals, a systematic review (protocol). *PROSPERO* 2020:CRD42020219869.
55. Griffiths P, Ball J, Drennan J, James L, Jones J, Recio A, Simon M. *The Association Between Patient Safety Outcomes and Nurse/Healthcare Assistant Skill Mix and Staffing Levels and Factors that May Influence Staffing Requirements (NICE Evidence Review)*. Southampton, GB: University of Southampton; 2014.
56. Sterne JA, Hernan MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919. https://doi.org/10.1136/bmj.i4919
57. National Institute for Clinical Excellence. *Methods for the Development of NICE Public Health Guidance*. London: NICE; 2012.
58. Henrikson NB, Skelly AC. Economic studies part 2: evaluating the quality. *Evid Based Spine Care J* 2013;**4**:2–5. https://doi.org/10.1055/s-0033-1341594
59. Griffiths P, Saville C, Ball J, Dall'Ora C, Meredith P, Turner L, Jones J. Costs and cost-effectiveness of improved nurse staffing levels and skill mix in acute hospitals: a systematic review. *Int J Nurs Stud* 2023;**147**:104601. https://doi.org/10.1016/j.ijnurstu.2023.104601
60. Kim Y, Kim SH, Ko Y. Effect of nurse staffing variation and hospital resource utilization. *Nurs Health Sci* 2016;**18**:473–80. https://doi.org/10.1111/nhs.12294
61. Li YF, Wong ES, Sales AE, Sharp ND, Needleman J, Maciejewski ML, *et al*. Nurse staffing and patient care costs in acute inpatient nursing units. *Med Care* 2011;**49**:708–15. https://doi.org/10.1097/MLR.0b013e318223a9f1
62. Pang D, Liu Z, Wang L. Comparison of nursing aids and registered nurses mixed nursing staffing model with different ratios on the nursing outcomes and cost in Neurology and Neurosurgery Center. *Ir J Med Sci* 2019;**188**:1435–41. https://doi.org/10.1007/s11845-019-01988-8
63. Ross HI, Jones MC, Hendriksen BS, Hollenbeak CS. Nurse staffing and outcomes for pulmonary lobectomy: cost and mortality trade-offs. *Heart Lung* 2021;**50**:206–12. https://doi.org/10.1016/j.hrtlng.2020.12.001
64. Behner KG, Fogg LF, Fournier LC, Frankenbach JT, Robertson SB. Nursing resource management: analyzing the relationship between costs and quality in staffing decisions. *Health Care Manage Rev* 1990;**15**:63–71. https://doi.org/10.1097/00004010-199001540-00008
65. Clark SL, Saade GA, Meyers JA, Frye DR, Perlin JB. The clinical and economic impact of nurse to patient staffing ratios in women receiving intrapartum oxytocin. *Am J Perinatol* 2014;**31**:119–24. https://doi.org/10.1055/s-0033-1338175
66. Cookson G, Jones S, van Vlymen J, Laliotis I. *The Cost-Effectiveness of Midwifery Staffing and Skill Mix on Maternity Outcomes: A Report for the National Institute of Health and Care Excellence*. Guildford: University of Surrey; 2014.
67. Dall TM, Chen YJ, Seifert RF, Maddox PJ, Hogan PF. The economic value of professional nursing. *Med Care* 2009;**47**:97–104. https://doi.org/10.1097/MLR.0b013e3181844da8
68. Lasater KB, Aiken LH, Sloane DM, French R, Anusiewicz CV, Martin B, *et al*. Is hospital nurse staffing legislation in the public's interest?: an observational study in New York State. *Med Care* 2021;**59**:444–50. https://doi.org/10.1097/MLR.0000000000001519
69. Lasater KB, McHugh M, Rosenbaum PR, Aiken LH, Smith H, Reiter JG, *et al*. Valuing hospital investments in nursing: multistate matched-cohort study of surgical patients. *BMJ Qual Saf* 2021;**30**:46–55. https://doi.org/10.1136/bmjqs-2019-010534

70. Lasater KB, McHugh MD, Rosenbaum PR, Aiken LH, Smith HL, Reiter JG, *et al*. Evaluating the costs and outcomes of hospital nursing resources: a matched cohort study of patients with common medical conditions. *J Gen Intern Med* 2021;**36**:84–91. https://doi.org/10.1007/s11606-020-06151-z
71. Li X, Bowman SM, Smith TC. Effects of registered nurse staffing level on hospital-acquired conditions in cardiac surgery patients: a propensity score matching analysis. *Nurs Outlook* 2016;**64**:533–41. https://doi.org/10.1016/j.outlook.2016.05.002
72. Martsolf GR, Auerbach D, Benevent R, Stocks C, Jiang HJ, Pearson ML, *et al*. Examining the value of inpatient nurse staffing: an assessment of quality and patient care costs. *Med Care* 2014;**52**:982–8. https://doi.org/10.1097/MLR.0000000000000248
73. Needleman J, Buerhaus PI, Stewart M, Zelevinsky K, Mattke S. Nurse staffing in hospitals: is there a business case for quality? *Health Aff (Millwood)* 2006;**25**:204–11. https://doi.org/10.1377/hlthaff.25.1.204
74. Rothberg MB, Abraham I, Lindenauer PK, Rose DN. Improving nurse-to-patient staffing ratios as a cost-effective safety intervention. *Med Care* 2005;**43**:785–91. https://doi.org/10.1097/01.mlr.0000170408.35854.fa
75. Shamliyan TA, Kane RL, Mueller C, Duval S, Wilt TJ. Cost savings associated with increased RN staffing in acute care hospitals: simulation exercise. *Nurs Econ* 2009;**27**:302–14, 331.
76. Weiss ME, Yakusheva O, Bobay KL. Quality and cost analysis of nurse staffing, discharge preparation, and post-discharge utilization. *Health Serv Res* 2011;**46**:1473–94. https://doi.org/10.1111/j.1475-6773.2011.01267.x
77. Yakusheva O, Lindrooth R, Weiss M. Economic evaluation of the 80% baccalaureate nurse workforce recommendation: a patient-level analysis. *Med Care* 2014;**52**:864–9. https://doi.org/10.1097/MLR.0000000000000189
78. Griffiths P, Saville C, Ball JE, Chable R, Dimech A, Jones J, *et al*. The Safer Nursing Care Tool as a guide to nurse staffing requirements on hospital wards: observational and modelling study. *Health Soc Care Deliv Res* 2020;**8**:1–162. https://doi.org/10.3310/hsdr08160
79. Twigg DE, Geelhoed EA, Bremner AP, Duffield CM. The economic benefits of increased levels of nursing care in the hospital setting. *J Adv Nurs* 2013;**69**:2253–61. https://doi.org/10.1111/jan.12109
80. Van den Heede K, Simoens S, Diya L, Lesaffre E, Vleugels A, Sermeus W. Increasing nurse staffing levels in Belgian cardiac surgery centres: a cost-effective patient safety intervention? *J Adv Nurs* 2010;**66**:1291–6. https://doi.org/10.1111/j.1365-2648.2010.05307.x
81. McHugh MD, Aiken LH, Sloane DM, Windsor C, Douglas C, Yates P. 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* 2021;**397**:1905–13. https://doi.org/10.1016/S0140-6736(21)00768-6
82. Nixon J, Khan KS, Kleijnen J. Summarising economic evaluations in systematic reviews: a new approach. *BMJ* 2001;**322**:1596–8. https://doi.org/10.1136/bmj.322.7302.1596
83. Alonso-Echanove J, Edwards JR, Richards MJ, Brennan P, Venezia RA, Keen J, *et al*. Effect of nurse staffing and antimicrobial-impregnated central venous catheters on the risk for bloodstream infections in intensive care units. *Infect Control Hosp Epidemiol* 2003;**24**:916–25. https://doi.org/10.1086/502160
84. Cimiotti JP, Haas J, Saiman L, Larson EL. Impact of staffing on bloodstream infections in the neonatal intensive care unit. *Arch Pediatr Adolesc Med* 2006;**160**:832–6. https://doi.org/10.1001/archpedi.160.8.832
85. Needleman J, Buerhaus P, Pankratz VS, Leibson CL, Stevens SR, Harris M. Nurse staffing and inpatient hospital mortality. *N Engl J Med* 2011;**364**:1037–45. https://doi.org/10.1056/NEJMsa1001025
86. Patrician PA, McCarthy MS, Swiger P, Raju D, Breckenridge-Sproat S, Su X, *et al*. Association of temporal variations in staffing with hospital-acquired pressure injury in military hospitals. *Res Nurs Health* 2017;**40**:111–9. https://doi.org/10.1002/nur.21781
87. Shang J, Needleman J, Liu J, Larson E, Stone PW. Nurse staffing and healthcare-associated infection, unit-level analysis. *J Nurs Adm* 2019;**49**:260–5. https://doi.org/10.1097/NNA.0000000000000748

88. Tschannen DJ. *Organizational Structure, Process, and Outcome: The Effects of Nurse Staffing and Nurse-physician Collaboration on Patient Length of Stay*. Ann Arbor: University of Michigan; 2005.

89. Mark BA, Belyea M. Nurse staffing and medication errors: cross-sectional or longitudinal relationships? *Res Nurs Health* 2009;**32**:18–30. https://doi.org/10.1002/nur.20305

90. Al-Abdwani R, Williams CB, Dunn C, Macartney J, Wollny K, Frndova H, *et al*. Incidence, outcomes and outcome prediction of unplanned extubation in critically ill children: an 11year experience. *J Crit Care* 2018;**44**:368–75. https://doi.org/10.1016/j.jcrc.2017.12.017

91. Rochefort CM, Beauchamp ME, Audet LA, Abrahamowicz M, Bourgault P. Associations of 4 nurse staffing practices with hospital mortality. *Med Care* 2020;**58**:912–8. https://doi.org/10.1097/MLR.0000000000001397

92. Beltempo M, Blais R, Lacroix G, Cabot M, Piedboeuf B. Association of nursing overtime, nurse staffing, and unit occupancy with health care-associated infections in the NICU. *Am J Perinatol* 2017;**34**:996–1002. https://doi.org/10.1055/s-0037-1601459

93. Beltempo M, Lacroix G, Cabot M, Blais R, Piedboeuf B. Association of nursing overtime, nurse staffing and unit occupancy with medical incidents and outcomes of very preterm infants. *J Perinatol* 2018;**38**:175–80. https://doi.org/10.1038/jp.2017.146

94. Hamilton KE, Redshaw ME, Tarnow-Mordi W. Nurse staffing in relation to risk-adjusted mortality in neonatal care. *Arch Dis Child Fetal Neonatal Ed* 2007;**92**:F99–103. https://doi.org/10.1136/adc.2006.102988

95. Hugonnet S, Chevrolet JC, Pittet D. The effect of workload on infection risk in critically ill patients. *Crit Care Med* 2007;**35**:76–81. https://doi.org/10.1097/01.CCM.0000251125.08629.3F

96. Hugonnet S, Uckay I, Pittet D. Staffing level: a determinant of late-onset ventilator-associated pneumonia. *Crit Care* 2007;**11**:R80. https://doi.org/10.1186/cc5974

97. Musy SN, Endrich O, Leichtle AB, Griffiths P, Nakas CT, Simon M. The association between nurse staffing and inpatient mortality: a shift-level retrospective longitudinal study. *Int J Nurs Stud* 2021;**120**:103950. https://doi.org/10.1016/j.ijnurstu.2021.103950

98. Ambrosi E, De Togni S, Guarnier A, Barelli P, Zambiasi P, Allegrini E, *et al*. In-hospital elderly mortality and associated factors in 12 Italian acute medical units: findings from an exploratory longitudinal study. *Aging Clin Exp Res* 2017;**29**:517–27. https://doi.org/10.1007/s40520-016-0576-8

99. Jansson MM, Syrjala HP, Ala-Kokko TI. Association of nurse staffing and nursing workload with ventilator-associated pneumonia and mortality: a prospective, single-center cohort study. *J Hosp Infect* 2019;**101**:257–63. https://doi.org/10.1016/j.jhin.2018.12.001

100. Palese A, Gonella S, Moreale R, Guarnier A, Barelli P, Zambiasi P, *et al*.; ESAMED Group. Hospital-acquired functional decline in older patients cared for in acute medical wards and predictors: findings from a multicentre longitudinal study. *Geriatr Nurs* 2016;**37**:192–9. https://doi.org/10.1016/j.gerinurse.2016.01.001

101. Callaghan LA, Cartwright DW, O'Rourke P, Davies MW. Infant to staff ratios and risk of mortality in very low birthweight infants. *Arch Dis Child Fetal Neonatal Ed* 2003;**88**:F94–97. https://doi.org/10.1136/fn.88.2.f94

102. Twigg DE, Myers H, Duffield C, Pugh JD, Gelder L, Roche M. 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* 2016;**63**:189–200. https://doi.org/10.1016/j.ijnurstu.2016.09.008

103. Dall'Ora C, Rubbo B, Saville C, Turner L, Ball J, Ball C, Griffiths P. The association between multi-disciplinary staffing levels and mortality in acute hospitals: a systematic review. *Hum Resour Health* 2023;**21**:30. https://doi.org/10.1186/s12960-023-00817-5

104. Neuraz A, Guerin C, Payet C, Polazzi S, Aubrun F, Dailler F, *et al*. Patient mortality is associated with staff resources and workload in the ICU: a multicenter observational study. *Crit Care Med* 2015;**43**:1587–94. https://doi.org/10.1097/CCM.0000000000001015

105. Bond CA, Raehl CL, Pitterle ME, Franke T. Health care professional staffing, hospital characteristics, and hospital mortality rates. *Pharmacotherapy* 1999;**19**:130–8. https://doi.org/10.1592/phco.19.3.130.30915

106. Smith DL, Elting LS, Learn PA, Raut CP, Mansfield PF. Factors influencing the volume-outcome relationship in gastrectomies: a population-based study. *Ann Surg Oncol* 2007;**14**:1846–52. https://doi.org/10.1245/s10434-007-9381-0

107. Peelen L, de Keizer NF, Peek N, Scheffer GJ, van der Voort PHJ, de Jonge E. The influence of volume and intensive care unit organization on hospital mortality in patients admitted with severe sepsis: a retrospective multicentre cohort study. *Crit Care* 2007;**11**:R40. https://doi.org/10.1186/cc5727

108. Robertson RH, Hassan M. Staffing intensity, skill mix and mortality outcomes: the case of chronic obstructive lung disease. *Health Serv Manage Res* 1999;**12**:258–68. https://doi.org/10.1177/095148489901200407

109. Bjerregaard U, Hølge-Hazelton B, Rud Kristensen S, Rose Olsen K. Nurse staffing and patient outcomes: analyzing within-and between-variation. *DaCHE Discussion Papers* 2020;**3**:1–23. https://doi.org/10.21996/d5ef-1y69.

110. Checkley W, Martin GS, Brown SM, Chang SY, Dabbagh O, Fremont RD, *et al*.; United States Critical Illness and Injury Trials Group Critical Illness Outcomes Study Investigators. Structure, process, and annual ICU mortality across 69 centers: United States Critical Illness and Injury Trials Group Critical Illness Outcomes Study. *Crit Care Med* 2014;**42**:344–56. https://doi.org/10.1097/CCM.0b013e3182a275d7

111. Chung W, Sohn M. The impact of nurse staffing on in-hospital mortality of stroke patients in Korea. *J Cardiovasc Nurs* 2018;**33**:47–54. https://doi.org/10.1097/JCN.0000000000000415

112. Ricciardi R, Nelson J, Roberts PL, Marcello PW, Read TE, Schoetz DJ. Is the presence of medical trainees associated with increased mortality with weekend admission? *BMC Med Educ* 2014;**14**:4. https://doi.org/10.1186/1472-6920-14-4

113. West E, Barron DN, Harrison D, Rafferty AM, Rowan K, Sanderson C. Nurse staffing, medical staffing and mortality in intensive care: an observational study. *Int J Nurs Stud* 2014;**51**:781–94. https://doi.org/10.1016/j.ijnurstu.2014.02.007

114. Kim Y, Kim J. In-hospital mortality in patients receiving percutaneous coronary intervention according to nurse staffing level: an analysis of national administrative health data. *Int J Environ Res Public Health* 2020;**17**:3799. https://doi.org/10.3390/ijerph17113799

115. NHS Digital. *National Workforce Data Set (NWD) and NHS Occupational Codes*. No date. URL: https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/national-workforce-data-set-nwd-and-nhs-occupation-codes#top (accessed 15 January 2021).

116. NHS England. *NHS Data Model and Dictionary*. URL: https://datadictionary.nhs.uk/index.html (accessed 17 November 2020).

117. NHS Digital. *NHS Workforce Statistics*. No date. URL: https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics (accessed 15 March 2022).

118. NHS Digital. *Hospital Estates and Facilities Statistics 2018/19: NHS Digital*. 2019. URL: https://digital.nhs.uk/data-and-information/publications/statistical/estates-returns-information-collection (accessed 30 August 2023).

119. NHS Digital. *About the Summary Hospital-Level Mortality Indicator (SHMI)*. No date. URL: https://digital.nhs.uk/shmi (accessed 8 March 2022).

120. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83. https://doi.org/10.1016/0021-9681(87)90171-8

121. Campbell MJ, Jacques RM, Fotheringham J, Maheswaran R, Nicholl J. Developing a summary hospital mortality index: retrospective analysis in English hospitals over five years. *BMJ* 2012;**344**:e1001. https://doi.org/10.1136/bmj.e1001

122. NHS Survey Coordination Centre. *NHS Staff Survey*. No date. URL: www.nhsstaffsurveys.com/ (accessed 15 March 2022).

123. Wake M, Green W. Relationship between employee engagement scores and service quality ratings: analysis of the National Health Service staff survey across 97 acute NHS Trusts in England and concurrent Care Quality Commission outcomes (2012-2016). *BMJ Open* 2019;**9**:e026472. https://doi.org/10.1136/bmjopen-2018-026472

124. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw* 2015;**67**:1–48. https://doi.org/10.18637/jss.v067.i01

125. Bell A, Fairbrother M, Jones K. Fixed and random effects models: making an informed choice. *Qual Quant* 2019;**53**:1051–74.

126. Schunck R. Within and between estimates in random-effects models: advantages and drawbacks of correlated random effects and hybrid models. *Stata J: Promoting Communications on Statistics and Stata* 2013;**13**:65–76.

127. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Qual Quant* 2007;**41**:673–90.

128. Power M, Stewart K, Brotherton A. What is the NHS safety thermometer? *Clin Risk* 2012;**18**:163–9. https://doi.org/10.1258/cr.2012.012038

129. Friedrichs M, Webster C, Marsh B, Dice J, Lee S. *fedmatch: Fast, Flexible, and User-Friendly Record Linkage Methods (v2.05)*. 2021. URL: https://CRAN.R-project.org/package=fedmatch (accessed 18 August 2023).

130. NHS Digital. *ISB 0070: Healthcare Resource Groups (HRGs)*. 2018. URL: https://web.archive.org/web/20230818122712/https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb-0070-healthcare-resource-groups-hrgs (accessed 18 August 2023).

131. Royal College of Physicians. *National Early Warning Score (NEWS) 2: Standardising the Assessment of Acute-illness Severity in the NHS – Updated Report of a Working Party*. London: Royal College of Physicians; 2017.

132. NHS Digital. *About the Summary Hospital-Level Mortality Indicator (SHMI)*. 2022. URL: https://digital.nhs.uk/shmi. (accessed 18 August 2023).

133. NHS Employers. *NHS Terms and Conditions of Service (Agenda for Change)*. 2023. URL: www.nhsemployers.org/topics/pay-pensions-and-reward/nhs-terms-and-conditions-service-agenda-change (accessed 18 August 2023).

134. Dall'Ora C, Ball J, Redfern O, Recio-Saucedo A, Maruotti A, Meredith P, Griffiths P. Are long nursing shifts on hospital wards associated with sickness absence? A longitudinal retrospective observational study. *J Nurs Manag* 2019;**27**:19–26. https://doi.org/10.1111/jonm.12643

135. Therneau TM, Grambsch PM. *Multiple Events per Subject. Modeling Survival Data: Extending the Cox Model*. New York, NY: Springer New York; 2000. pp. 169–229.

136. Ngwa JS, Cabral HJ, Cheng DM, Pencina MJ, Gagnon DR, LaValley MP, Cupples LA. A comparison of time dependent Cox regression, pooled logistic regression and cross sectional pooling with simulations and an application to the Framingham Heart Study. *BMC Med Res Methodol* 2016;**16**:148. https://doi.org/10.1186/s12874-016-0248-6

137. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med* 1990;**9**:1501–15. https://doi.org/10.1002/sim.4780091214

138. Kuha J. AIC and BIC: comparisons of assumptions and performance. *Sociol Methods Res* 2004;**33**:188–229. https://doi.org/10.1177/0049124103262065

139. Royston P, Sauerbrei W. Handling interactions in Stata, especially with continuous predictors. *United Kingdom Stata Users' Group Meetings 2012 16*. Stata Users Group; 2012.

140. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, *et al*. Welcome to the Tidyverse. *J Open Source Softw* 2019;**4**:1686. https://doi.org/10.21105/joss.01686

141. Therneau T. A package for survival analysis in S. *R Package Version* 2023.

142. Therneau TM. Mixed Effects Cox Models (R package coxme version 2.2-18.1). 2022. URL: www.rdocumentation.org/packages/coxme/versions/2.2-18.1 (accessed 29 April 2025).

143. Harrison E, Drake T, Ots R. finalfit: Quickly Create Elegant Regression Results Tables and Plots when Modelling. R package version 1.0.6 2023. URL: https://www.rdocumentation.org/packages/finalfit/versions/1.0.6 (accessed 11 November 2022).

144. Allaire J, Xie Y, Dervieux C, McPherson J, Luraschi J, Ushey K, *et al*. rmarkdown: Dynamic Documents for R_. R package version 2.21. 2023. URL: https://github.com/rstudio/rmarkdown (accessed 29 April 2025).

145. NHS England. *National Cost Collection for the NHS. 2021/22 National Cost Collection Data*. 2023. URL: https://web.archive.org/web/20230818110747/https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ (accessed 12 December 2023).

146. Jones KC, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A, *et al*. *Unit Costs of Health and Social Care 2022 Manual*. Canterbury & York: University of Kent and University of York; 2022.

147. Stare J, Maucort-Boulch D. Odds ratio, hazard ratio and relative risk. *Metodoloski Zvezki* 2016;**13**:59–67.

148. Nurminen M. To use or not to use the odds ratio in epidemiologic analyses? *Eur J Epidemiol* 1995;**11**:365–71. https://doi.org/10.1007/BF01721219

149. Meacock R, Kristensen SR, Sutton M. The cost-effectiveness of using financial incentives to improve provider quality: a framework and application. *Health Econ* 2014;**23**:1–13. https://doi.org/10.1002/hec.2978

150. HM Treasury. *The Green Book (2022) [HM Treasury guidance on how to appraise policies, programmes and projects]*. London: HM Government; 2022. URL: www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent/the-green-book-2020 (accessed 29 April 2025).

151. Haacker M, Hallett TB, Atun R. On discount rates for economic evaluations in global health. *Health Policy Plan* 2020;**35**:107–14. https://doi.org/10.1093/heapol/czz127

152. Office for National Statistics. *National Life Tables – Life Expectancy in the UK: 2018 to 2020: Office for National Statistics*. 2021. URL: https://web.archive.org/web/20230818130005/https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (accessed 18 August 2023).

153. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;**14**:539–45. https://doi.org/10.1016/j.jval.2010.10.029

154. Ball J, Ballinger C, De Iongh A, Dall'Ora C, Crowe S, Griffiths P. Determining priorities for research to improve fundamental care on hospital wards. *Res Involv Engagem* 2016;**2**:31. https://doi.org/10.1186/s40900-016-0045-8

155. Puckrein G. *Nothing About Me Without Me: The Patient Research Exchange*. 2016. URL: http://web.archive.org/web/20230822162804/https://www.patientresearchexchange.org/stories/detail/nothing-about-me-without-me (accessed 22 August 2023).

156. Winter SG, Bartel AP, de Cordova PB, Needleman J, Schmitt SK, Stone PW, Phibbs CS. The effect of data aggregation on estimations of nurse staffing and patient outcomes. *Health Serv Res* 2021;**56**:1262–70. https://doi.org/10.1111/1475-6773.13866

157. Needleman J, Liu J, Shang J, Larson EL, Stone PW. Association of registered nurse and nursing support staffing with inpatient hospital mortality. *BMJ Qual Saf* 2020;**29**:10–18. https://doi.org/10.1136/bmjqs-2018-009219

158. Leary A, Cook R, Jones S, Radford M, Smith J, Gough M, Punshon G. Using knowledge discovery through data mining to gain intelligence from routinely collected incident reporting in an acute English hospital. *Int J Health Care Qual Assur* 2020;**33**:221–34.

159. Zaranko B, Sanford NJ, Kelly E, Rafferty AM, Bird J, Mercuri L, *et al*. Nurse staffing and inpatient mortality in the English National Health Service: a retrospective longitudinal study. *BMJ Qual Saf* 2023;**32**:254–63. https://doi.org/10.1136/bmjqs-2022-015291

160. Griffiths P, Ball J, Bloor K, Böhning D, Briggs J, Dall'Ora C, *et al*. *Nurse Staffing Levels, Missed Vital Signs and Mortality in Hospitals: Retrospective Longitudinal Observational Study*. Southampton (UK): NIHR Journals Library; 2018.

161. National Institute for Health and Care Excellence. *Changes to NICE Drug Appraisals: What You Need to Know*. 2017. URL: https://web.archive.org/web/20240617050131/https://www.nice.org.uk/process/pmg19/chapter/selecting-technologies#selecting-technologies-for-the-fast-track-appraisal-process (accessed 29 April 2025).

162. Owen L, Pennington B, Fischer A, Jeong K. The cost-effectiveness of public health interventions examined by NICE from 2011 to 2016. *J Public Health (Oxf)* 2018;**40**:557–66. https://doi.org/10.1093/pubmed/fdx119

163. Discover-NOW Citizens Advisory Group. *Case Study: Discover-NOW Citizens Advisory Group London: Discover-NOW, the Health Data Research Hub for Real World Evidence*. 2021. URL: https://web.archive.org/web/20230823085239/https://discover-now.co.uk/case-study/discover-now-citizen-advisory-group-cag/ (accessed 23 August 2023).

164. Jones TL, Hamilton P, Murry N. Unfinished nursing care, missed care, and implicitly rationed care: state of the science review. *Int J Nurs Stud* 2015;**52**:1121–37. https://doi.org/10.1016/j.ijnurstu.2015.02.012

165. Department of Human Services. Ridoult L, Schoo A, Santos T. Workload capacity measure for use in allied health workforce planning. Victoria: Department of Human Services; 2006.

166. Needleman J, Buerhaus P, Mattke S. *Nurse-staffing Levels and Patient Outcomes in Hospitals. Final Report for Health Resources and Services Administration*. Boston, MA: Harvard School of Public Health; 2001.

167. Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med* 2002;**346**:1715–22. https://doi.org/10.1056/NEJMsa012247

168. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA* 2002;**288**:1987–93. https://doi.org/10.1001/jama.288.16.1987

169. Van den Heede K, Lesaffre E, Diya L, Vleugels A, Clarke SP, Aiken LH, Sermeus W. The relationship between inpatient cardiac surgery mortality and nurse numbers and educational level: analysis of administrative data. *Int J Nurs Stud* 2009;**46**:796–803. https://doi.org/10.1016/j.ijnurstu.2008.12.018

Appendix 1 Additional material for literature review

TABLE 27 Search strategies

Economics (PubMed example)			
(Economic*[Title/Abstract]) (Cost*[Title/Abstract]) (Business case[Title/Abstract])	((inpatients[MeSH Major Topic]) (hospitals[MeSH Major Topic])) (hospital units[MeSH Major Topic]) inpatient?[Title/Abstract] hospital?[Title/Abstract] (medical-surgical[Title/Abstract]) (surgical-medical[Title/Abstract]) “medical unit’[Title/Abstract] “medical ward’[Title/Abstract] “surgical unit’[Title/Abstract] “surgical ward’[Title/Abstract] “patient surgical’[Title/Abstract] “patient medical’[Title/Abstract]	Staffing “workforce”[MeSH Major Topic] “personnel staffing and scheduling” [MeSH Major Topic] workload [MeSH Major Topic] workload “skill-mix” [Title/Abstract] “skillmix” [Title/Abstract] “skill-mix” [Title/Abstract] (“staff*” [All Fields] AND “level*” [All Fields]) OR (“staff*” [All Fields] AND “ratio*” [All Fields]) OR (“staff*” [All Fields] AND “model*” [All Fields]) OR (“staff*” [All Fields] AND “roster*” [All Fields])	nurs* “nursing staff, hospital” [MeSH Major Topic] “nursing services” [MeSH Major Topic] “nursing assistants” [MeSH Major Topic] “health care assistant*” [All Fields] “healthcare assistant*” [All Fields] “health care support worker*” [All Fields] “healthcare support worker*” [All Fields] “midwi*” [All Fields] “midwifery” [MeSH Major Topic]
Longitudinal (MEDLINE example)			
longitudinal.mp. exp Longitudinal Studies/ “cohort stud*”.mp. exp Cohort Studies/ “panel stud*”.mp. “prospective stud*”.mp. exp Prospective Studies/ “randomi* control trial”.mp. Randomized Controlled Trials as Topic/ RCT.mp. (cluster adj3 trial).tw. exp Survival Analysis/ “survival analysis” ”.mp.		staffing.mp. exp Workforce/ exp “Personnel Staffing and Scheduling”/ (skill* adj1 mix*).tw. (staff* adj3 (level* or ratio* or model* or roster*)).mp. exp Workload/ workload.mp.	nurs*.mp. Nursing Staff, Hospital/ Nursing Services/ exp Nursing Assistants/ (“health care assistant*” or “healthcare assistant*”).mp. (“health care support worker*” or “healthcare support worker*”).mp. “midwi*”.mp. [mp = title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

TABLE 27 Search strategies *(continued)*

Multidisciplinary (MEDLINE example)		
Hospital Mortality/ “hospital death? ”.ab,ti. “mortality rate? ”.ab,ti. “death rate? ”.ab,ti.	Exp Hospitals/ exp Hospital Units/ exp Inpatients/ Inpatient?.ab,ti. Hospital?.ab,ti. (medical adj3 (unit* or ward*)).ab,ti. (surgical adj3 (unit* or ward*)).ab,ti. (patient* adj3 surgical).ab,ti. (patient* adj3 medical).ab,ti. (“medical-surgical” or “surgical-medical”).ab,ti.	(workload* or workforce* or shift or shiftwork* or shifts or overtime).ab,ti. (Skill mix$ or skillmix or skill-mix$ or skills mix).ab,ti. (skill* adj1 mix).ab,ti. (Staff* level$”.ab,ti. Exp Work schedule/ exp “Personnel Staffing and Scheduling”/ exp Health manpower/ exp Workload/ (skillmix*”.ab,ti. (staffmix* or “staff mix*”).ab,ti. Staffing.ab,ti. (understaff*”.ab,ti. (under staff* “.ab,ti. (staff* adj3 (level* or ratio* or nadequa* or model* or number* or mix* or rota* or rosta* or roster* or nadequa* or overtime or supervision or supervisory)).ab,ti. (staff* adj3 (sufficient* or sufficiency or adequate* or nadequa* or target* or insufficient* or insufficienc* or inadequate* or nadequacy* or short or shortage or efficient* or efficienc* or inefficien*)).ab,ti. (Doctor? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Medical staff* adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Physician? Adj5 (number or workforce or work force)).ab,ti. (surgeon? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. Hospital? Volunteer$.ab,ti. Hospital? Assistant?.ab,ti. (consultant? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,(987) (Pharmacist? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,(2450) (Physiotherap* adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Occupational therap* adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Health* professional? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Allied health professional? Adj5 (number or workforce or workforce or staffing or workload or skill mix)).ab,ti. (Health* professional?”.ab,ti.

.ab, abstract; .mp, title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms; .ti, title; .tw, title or abstract; MESH, medical subject heading (also indicated by/after the term).

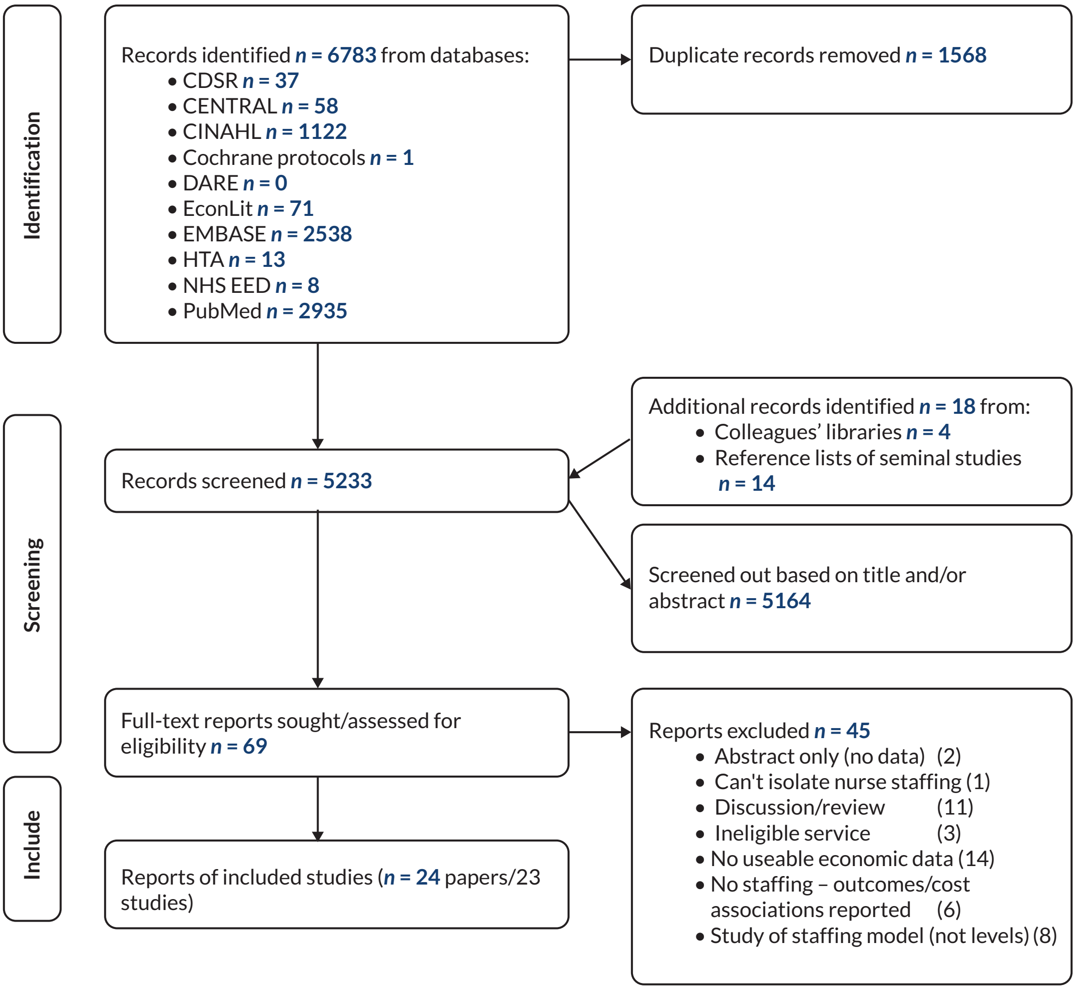


FIGURE 6 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram of sources for the economic review.

TABLE 28 Characteristics of economic studies

Paper	Country	Patient group	Study design	Source of variation	Level of aggregation for staffing	Hospitals	Patients	Economic analysis
Behner *et al.* 1990[64]	USA	Back and neck procedures	Retrospective observational study	Natural variation	Patient stay	1	132	Cost–consequences (disaggregated) and net cost of avoiding low (20% below standard) staffing
Clark *et al.* 2014[65]	USA	Maternity – induction of labour	Retrospective cross-sectional observational study	Natural variation	Hospital	110	101,377	Cost–consequences (disaggregated) of providing universal one-to-one midwifery care
Cookson *et al.* 2014[66]	UK	Maternity	Retrospective cross-sectional observational study	Natural variation	Hospital	157	5,753,551	Cost-effectiveness of 1 additional midwife per 100 deliveries
Dall *et al.* 2009[67]	USA	General med/surg	Secondary modelling (data from cross-sectional studies)[a]	Simulated change	Hospital	610	5,400,000	Cost–(monetary) benefits (disaggregated) and consequences of increased RN staffing
Griffiths *et al.* 2018[19]	UK	General med/surg	Retrospective longitudinal observational	Natural variation	Patient day	1	138,133	Cost-effectiveness of one additional RN HPPD and increasing skill-mix to establishment
Griffiths *et al.* 2020[15,78]	UK	General med/surg	Simulation model (parameter data from retrospective longitudinal study)[b]	Simulated change	Shift	4	N/A	Cost–consequences (of low staffing) and effects of different baseline staffing policies
Kim *et al.* 2016[60]	South Korea	Hip and knee surgery	Retrospective cross-sectional observational	Natural variation	Hospital	222	22,289	Care cost (charges) and consequences (disaggregated) of different patient-to-nurse ratios in hospitals
Lasater *et al.* 2021[68]	USA	General med/surg (select)	Retrospective cross-sectional observational	Natural variation	Hospital	116	417,861	Cost–consequence (disaggregated) of changed patient-to-nurse ratio
Lasater *et al.* 2021[69]	USA	General surg (select)	Matched cohort/retrospective cross-sectional observational	Natural variation	Hospital	306	125,430	Cost-effectiveness of composite nursing resource (staffing, skill-mix, BSN mix and nurse-reported work environment)
Lasater *et al.* 2021[70]	USA	General med (select)	Matched cohort/retrospective cross-sectional observational	Natural variation	Hospital	306	148,090	Cost-effectiveness of composite nursing resource (staffing, skill-mix, BSN mix and nurse-reported work environment)
Li *et al.* 2011 [61]	USA	General med/surg	Retrospective cross-sectional observational	Natural variation	Unit	125	110,646	Costs of additional nursing hour and increased RN skill-mix
Li *et al.* 2016[71]	USA	Cardiac surgery	Propensity-matched cohort/retrospective cross-sectional observational	Natural variation	Hospital	1887	439,365	Cost–consequence (disaggregated) of hospital above median staffing (RN HPPD ≥ 7.07) vs below (RN HPPD < 7.07)

continued

TABLE 28 Characteristics of economic studies *(continued)*

Paper	Country	Patient group	Study design	Source of variation	Level of aggregation for staffing	Hospitals	Patients	Economic analysis
Martsolf *et al.* 2014[72]	USA	General med/surg	Matched cohort/retrospective cross-sectional observational	Natural variation	Hospital	421	18,474,860	Cost–consequence (disaggregated) of additional nurse (RN/LPN) per 1000 admissions and higher RN/LPN skill-mix
McHugh *et al.* 2021[81]	Australia	General med/surg	Quasi-experimental panel	Natural variation and policy implementation	Hospital	55	489,155	Cost–consequences of implementing minimum nurse to patient ratios
Needleman *et al.* 2006[73]	USA	General med/surg	Secondary modelling (data from cross-sectional studies)[a]	Simulated change	Hospital	799	5,075,969	Cost–consequences (disaggregated) of different staffing levels/configurations (in HPPD)
Pang *et al.* 2019[62]	China	Neurology/neuorsurgery	Prospective cross-sectional observational	Natural variation	Hospital	1	5091	Cost–consequence (disaggregated) of care in six wards with different proportions of RNs
Ross *et al.* 2021[63]	USA	Pulmonary lobectomy	Retrospective cross-sectional observational study	Natural variation	Hospital	N/A	16,944	Cost–consequence (disaggregated) of different staffing levels (RN FTEs per 1000 patient days)
Rothberg *et al.* 2005[74]	USA	General med/surg	Secondary modelling (data from cross-sectional studies)[c]	Natural variation	Hospital	799	5,075,969	Cost-effectiveness per unit reduction in patient-to-nurse ratio
Shamliyan *et al.* 2009[75]	USA	General med/surg	Secondary modelling (data from cross-sectional studies)[d]	Natural variation	Hospital	N/A	N/A	Net benefit and cost–benefit arising from avoided deaths (and adverse events – not reported) corresponding to a 1 FTE RN per 1000 patients increase
Twigg *et al.* 2013[79]	Australia	General med/surg	Retrospective observational study	Implementation of new staffing levels	Hospital	3	214,261	Cost-effectiveness of implementing a NHPPD method to guide nurse staffing
Van den Heede *et al.* 2010[80]	Belgium	Cardiac surgery	Secondary modelling (data from cross-sectional studies)[e]	Simulated change	Ward	28	9054	Cost-effectiveness of increasing nurse staffing to the 75th centile
Weiss *et al.* 2011[76]	USA	General med/surg	Prospective cross-sectional observational	Natural variation	Ward	4	1892	Cost–consequences (disaggregated) of increasing non-overtime RN staffing
Yakusheva *et al.* 2014[77]	USA	General med/surg	Retrospective observational study	Simulated change	Patient stay	1	8526	Cost–consequences (disaggregated) of increasing the proportion of BSN-qualified RN staffing

LPN, Licensed Practical Nurse; Med, Medical; N/A, not applicable; NHPPD, nursing hours per patient day; Surg, Surgical; BSN, Bachelors of Science in Nursing; UK, United Kingdom; USA, United States of America.

a Needleman 2001, 2002.[166,167]
b Griffiths 2018.[19]
c Aiken *et al.* 2002.[168]
d Kane *et al.* 2007.[39]
e Van den Heede *et al.* 2009.[169]

Note
Study providing effectiveness estimates.

TABLE 29 Risk of bias in economic studies

Study	Eligible population[a]	Selected participants[b]	External validity[c]	Design[d]	Sufficient power[e]	Control for confounding[f]	Internal validity[c]	Overall[c]
Behner *et al.* 1990[64]	**–**	**++**	–	++	**–**	**–**	–	–
Clark *et al.* 2014[65]	+	++	+	–	++	–	–	–
Cookson *et al.* 2014[66]	++	++	++	–	++	++	–	–
Dall *et al.* 2009[67]	++	++	++	–	++	++	–	–
Griffiths *et al.* 2018[19]	+	++	+	++	++	++	++	+
Griffiths *et al.* 2020[15,78]	+	++	+	++	++	++	++	+
Kim *et al.* 2016[60]	+	++	+	–	++	++	–	–
Lasater *et al.* 2021[68]	+	+	+	–	++	++	–	–
Lasater *et al.* 2021[69]	+	+	+	+	++	++	+	+
Lasater *et al.* 2021[70]	+	+	+	+	++	++	+	+
Li *et al.* 2011[71]	++	++	++	–	++	++	–	–
Li *et al.* 2016[61]	+	++	+	+	++	++	+	+
Martsolf *et al.* 2014[72]	++	++	++	–	++	+	–	–
McHugh *et al.* 2021[81]	++	+	+	+	++	++	+	+
Needleman *et al.* 2006[73]	++	++	++	–	++	++	–	–
Pang *et al.* 2019[62]	+	++	+	–	+	+	–	–
Ross *et al.* 2021[63]	+	++	+	–	++	++	–	–
Rothberg *et al.* 2005[74]	+	+	+	–	++	+	–	–
Shamliyan *et al.* 2009[75]	++	++	++	+	++	+	+	+
Twigg *et al.* 2013[79]	+	++	+	+	++	++	+	+
Van den Heede *et al.* 2010[80]	++	++	++	–	+	++	–	–

continued

TABLE 29 Risk of bias in economic studies (*continued*)

Study	Eligible population[a]	Selected participants[b]	External validity[c]	Design[d]	Sufficient power[e]	Control for confounding[f]	Internal validity[c]	Overall[c]
Weiss *et al.* 2011[76]	+	+	+	–	+	++	–	–
Yakusheva *et al.* 2014[77]	+	++	+	++	+	++	+	+

a Target population is patients in service for which staffing is measured. If the patient group included represents the target population and the sample of hospitals is 10 or more rate as strong.
b If staffing data are complete and the target patient population is fully sampled (e.g. administrative data) rate as strong.
c Summary rating based on lowest score in preceding section.
d Cross-sectional designs rated as weak unless accompanied by additional features such as matching/propensity (stronger). Longitudinal studies and studies with using individual patient exposure measures rated as strong.
e Based on sample size alone as no power calcs given – small (weak) < 1000 < medium (stronger) < 10,000 large (strong) 10,000 + patients.
f No adjustment for patient risk factors (weak), partial (stronger) or extensive (strong).
–, high risk of bias (weak); +, moderate (stronger); ++, low risk of bias (strong).

TABLE 30 Assessment of economic study quality

Study	Economic perspective	Discounting/ time horizon?	Staff cost	Consequential	Treatment cost	Post-discharge cost	Societal costs	Sensitivity analysis	Range of costs[a]	Level of economic analysis[b]	Risk of bias in underlying study
Behner *et al.* 1990[64]	Hospital	n/i	Salary	Hospital stay	AE			No	Moderate	Moderate	High
Clark *et al.* 2014[65]	Hospital	n/i	Salary					No	Limited	Moderate	High
Cookson *et al.* 2014[66]	Hospital	n/i	Employment					Yes	Limited	Moderate	High
Dall *et al.* 2009[67]	Societal	y/life	Employment	Hospital stay	AE	Follow-up care	Productive value	No	Extensive	Moderate	High
Griffiths *et al.* 2018[19]	Hospital	n/i	Employment	Hospital stay				Yes	Moderate	Moderate	Moderate
Griffiths *et al.* 2020[15,78]	Hospital	n/i	Employment	Hospital stay				Yes	Moderate	Moderate	Moderate
Kim *et al.* 2016[60]	Hospital	n/i	Charges	Hospital stay				yes	Moderate	Moderate	High
Lasater *et al.* 2021[68]	Hospital/ patient	n/i	Hospital costs			Re-admission		No	Moderate	Moderate	High

TABLE 30 Assessment of economic study quality (*continued*)

Study	Economic perspective	Discounting/ time horizon?	Staff cost	Consequential	Treatment cost	Post-discharge cost	Societal costs	Sensitivity analysis	Range of costs[a]	Level of economic analysis[b]	Risk of bias in underlying study
Lasater *et al.* 2021[69]	Hospital/ patient	n/i	Employment	Hospital stay		Re-admission		No	Moderate	Moderate	Moderate
Lasater *et al.* 2021[70]	Hospital/ patient	n/i	Employment	Hospital stay		Re-admission		Yes	Moderate	Moderate	Moderate
Li *et al.* 2011[61]	Hospital	n/i	Hospital costs					Yes	Limited	Low	High
Li *et al.* 2016[71]	Hospital	n/i	Hospital costs	Hospital stay	AE			Yes	Moderate	Moderate	Moderate
Martsolf *et al.* 2014[72]	Hospital	n/i	Hospital costs	Hospital stay	AE			Yes	Moderate	Moderate	High
McHugh *et al.* 2021[81]	Hospital	n/i	Salary	Hospital stay		Re-admission		No	Moderate	Moderate	Moderate
Needleman *et al.* 2006[73]	Hospital	n/i	Salary	Hospital stay	AE			No	Moderate	Moderate	High
Pang *et al.* 2019[62]	Hospital	n/i	Employment					No	Limited	Moderate	High
Ross *et al.* 2021[63]	Hospital	n/i	Hospital costs	Hospital stay				No	Moderate	Moderate	High
Rothberg *et al.* 2005[74]	Hospital	n/i	Salary	Hospital stay				Yes	Moderate	Moderate	High
Shamliyan *et al.* 2009[75]	Societal	y/life	Employment	Hospital stay	AE		Future earnings	Yes	Extensive	High	Moderate
Twigg *et al.* 2013[79]	Hospital	y/life	Employment		AE			Yes	Moderate	Moderate	Moderate
Van den Heede *et al.* 2010[80]	Hospital	y/life	Salary					Yes	Limited	Moderate	High
Weiss *et al.* 2011[76]	Hospital/ payer	n/i	Employment			Re-admission		No	Moderate	Moderate	High
Yakusheva *et al.* 2014[77]	Hospital	n/i	Salary			Re-admission		Yes	Moderate	Moderate	Moderate

AE, adverse event; i, immediate; n, no; y, yes.

a Limited – 1 cost domain, moderate – 2 or 3 cost domains, extensive – 4 or 5.

b Low – cost only study, moderate – cost–consequences or cost-effectiveness, high – cost–utility or cost–benefit.

Appendix 2 Additional methods

Staffing groups included in the national cross-sectional study

Our high-level groupings are based on those applied by NHS digital.

Medically qualified staff (three groups)

- General medicine specialist doctors were grouped with clinical oncology doctors to form the general medicine group.
- Surgeons were grouped with obstetricians and gynaecologists to form the surgical group.
- All other medical specialties were grouped together to form the other medical group (i.e. radiologists, emergency medicine, anaesthetists, and pathologists).
- We excluded psychiatrists and paediatricians.

Registered nurses

- The nurses' group was composed of nurses in adult services only; we did not include nurses from paediatric services.

Nurse support group

- Support to adult and general nurses, and nurses in training. Nurses in training include apprentices providing service (such as trainee nurse associates) but not supernumerary nursing students.

Allied health professionals

- Dietetics, music art and drama therapy, occupational therapy, orthoptics and optics, physiotherapy, podiatry and chiropody, radiography speech and language therapy were included in the AHP group.

Scientific, therapeutic and technical

- Applied psychologist, dentists, multi-therapists, operating theatre staff, psychological therapist, pharmacists, social services and other.

Safety thermometer harm definitions

Harms included in the safety thermometer are:

- **Pressure ulcers** (presence or absence, not count, with grade of worst PU recorded using categories from European Pressure Ulcer Scale). Reports differentiate OLD (present on admission or developed within 72 hours) and NEW (developed 72 hours of more after admission to reporting organisation);
- **Falls** – any falls within previous 72 hours (to report date) experienced by patient in a care setting (not necessarily current setting). Categorised for severity using NRLS categories (no fall, no harm, low harm, moderate harm, severe harm, death);
- **Urinary tract infections (UTIs)** – count of patients with UTIs, distinguishing between OLD (treatment started prior to admission to reporting organisation) and NEW (treatment started during current admission to reporting organisation);
- **Urinary catheter** – records whether patient has indwelling urinary catheter (at any point in preceding 72 hours) and number of days patient has been catheterised (1–28 days, > 28 days or days not known);

TABLE 31 Items contributing to the themes in staff survey benchmark data set

Health and well-being
Q5h – 'The opportunities for flexible working patterns'.
Q11a – 'Does your organisation take positive action on health and well-being?'
Q11b – 'In the last 12 months have you experienced musculoskeletal problems (MSK) as a result of work activities?'
Q11c – 'During the last 12 months have you felt unwell as a result of work-related stress?'
Q11d – 'In the last three months have you ever come to work despite not feeling well enough to perform your duties?'
Morale (2018 onwards)
Q4c – 'I am involved in deciding on changes introduced that affect my work area/team/department'.
Q4j – 'I receive the respect I deserve from my colleagues at work'.
Q6a – 'I have unrealistic time pressures'.
Q6b – 'I have a choice in deciding how to do my work'.
Q6c – 'Relationships at work are strained'.
Q8a – 'My immediate manager encourages me at work'.
Q23a – 'I often think about leaving this organisation'.
Q23b – 'I will probably look for a job at a new organisation in the next 12 months'.
Q23c – 'As soon as I can find another job, I will leave this organisation'.
Quality of care
Q7a – 'I am satisfied with the quality of care I give to patients/service users'.
Q7b – 'I feel that my role makes a difference to patients/service users'.
Q7c – 'I am able to deliver the care I aspire to'.

- **Venous thromboembolism (VTE) –** count of patients being clinically treated for VTE of any type. Distinguishes between OLD (treatment started prior to admission to reporting organisation) and NEW (treatment started during current admission to reporting organisation).

The safety thermometer coding of harms is as follows:

- Any harms include old or new PUs, falls with harm (category > 2), catheters with UTI and new VTEs.
- New harms include new PUs, falls with harm (category > 2), catheters with new UTI and new VTEs.

Data set preparation and cleaning longitudinal study

Date of death was partially validated by checking it against the discharge date when the discharge was coded as a death. In a small number of cases, the date of death had not been recorded or was a date after discharge. The discharge date was assumed to be the correct date of death when a discharge method of death was recorded. Admissions with a stay of 180 days or more were trimmed from the data set for analysis of LOS as occurrences were more likely to indicate a late closure of the admission electronic care record than a genuine stay exceeding that interval.

The staff roster records had an implied rather than explicit shift location given by their name and these did not always align with individual patient wards in space and time. Mapping tables were constructed for each hospital trust to associate roster names with patient wards and time intervals. Sometimes a roster covered more than one ward, which meant changing areas of staffing 'exposure' from wards to a collection of wards (superwards). Alternatively, there might be a number of rosters for one ward, for example different rosters for different job types or bands. Additionally, when a service or specialty changed locations, the ward names in the patient records and in the roster names in the eRoster systems did not always change synchronously. The mapping was undertaken with the assistance of the hospital information and workforce managers, but recorded change histories were patchy and so sometimes they had to rely on their memory.

Weekly run charts of RN staffing levels were constructed for each location showing ward occupancy, absolute staffing levels and care HPPD. Ward occupancy time series had a leading edge because patient data were extracted by admission date, leading to incorrectly high HPPD staffing levels. This was addressed by removing the first 2 weeks from the analysis, using only data from 13 Monday April. Remaining discontinuities were investigated. Investigation revealed that there was duplication in staff data leading to a sudden doubling of apparent staffing levels for some wards for some periods. This arose because a rostering system migration had been undertaken ward by ward and roster by roster accompanied by a data import from the outgoing system covering a variable duration. This led to duplicate staffing data when aggregating the extracts from the two systems. The duplicated data were identified by patient area, day and job type and was removed.

Where discontinuities could not be resolved, a record was made of the ward and the days affected so that the staffing levels could be treated as missing values in subsequent analysis. Days were also treated as having missing staffing values if the RN staffing for the location was < 21 hours in total after excluding breaks, as in a general acute ward there is an expectation that a RN is always available and so we assumed either erroneous staffing data or that this was not an acute ward. Similarly, we treated days as having missing values if the calculated RN staffing level measured in hours per patient-day was < 0.5 or > 48 (where the nurse-to-patient ratio would exceed 2).

Weekly run charts of each ward's case mix were constructed showing measures of age, gender, method of admission and proportion of overnight stays along with monthly tables of the ward's modal specialty, top two modal SHMI diagnosis groups and number of admissions. These were used to identify significant changes in the use of a ward and where found the ward study period was split at the start of the change, in effect creating a new ward and giving it its own identifier. Each ward period constructed by the split was then treated separately when calculating the expected staffing levels. Mean daily staffing levels were calculated for each ward period by job group, that is the RN and NA care HPPD, along with mean patient movements per RN/NA hour.

TABLE 32 Candidate nurse-sensitive outcomes based on Blume *et al.* 2021[36]

Patient outcome	Include/ exclude	Strength of evidence (Blume *et al.* 2021)	Data source	Definition	Denominator
LOS	Include	High	PAS	Hospital discharge datetime minus hospital admission datetime.	All adults except maternity. Paediatrics as separate analysis.
Patient dissatisfaction	Exclude	High	Data unavailable		
Poor quality of nurse-delivered care	Exclude	High	Data unavailable		
Re-admission	Include	High	PAS	Re-admitted to the same hospital within 30 days of discharge. (hospital subsequent admission datetime minus hospital initial discharge date time ≤ 30 days) – non elective only	All adults except maternity. Paediatrics as separate analysis.
Failure to rescue	Exclude	Moderate	Coding in Datix not consistent between Trusts		
Medication error	Include	Moderate	Datix	**Datix:** 'Medication' Administration or supply of a medicine from a clinical area, Advice, Medication error during the prescription process, Monitoring or follow-up of medicine use, Other medication error, Patient's case notes or records, Patient's reaction to Medication, Preparation of medicines/dispensing in pharmacy, Supply or use of over-the-counter medicines	Incidents affecting patients in any inpatient ward that can be mapped to roster and PAS.
Mortality	Include	Moderate	PAS	Died within 30 days of admission (where admission day is day 1). Note we do not have time of death.	All adults except maternity. Paediatrics as separate analysis.
Pneumonia	Include	Moderate	PAS – diagnoses	Diagnosis code 'Influenza and pneumonia' 'J13', 'J14', 'J15', 'J16', 'J18'. For surgical admissions (medical admissions more likely to have this present on admission). Influenza excluded https://emedicine.medscape.com/article/234753-overview#a3 *JO9 Influenza due to certain identified influenza virus exclude* *J10 Influenza due to other identified influenza virus exclude* *J11 Influenza, virus not identified exclude* *J12 Viral pneumonia, not elsewhere classified exclude* *J17* Pneumonia in diseases classified elsewhere exclude*	All surgical admissions. Limited to adults and non-maternity wards?
Respiratory failure	Use proxy instead	Moderate	Definition unclear		

continued

TABLE 32 Candidate nurse-sensitive outcomes based on Blume *et al.* 2021[36] (*continued*)

Patient outcome	Include/ exclude	Strength of evidence (Blume *et al.* 2021)	Data source	Definition	Denominator
Central nervous system (CNS) complications	Exclude	Low	Definition unclear		
DVT	Include	Low	PAS –diagnoses	**Diagnosis code** 'I80.1 Phlebitis and thrombophlebitis of femoral vein, I80.2 Phlebitis and thrombophlebitis of other deep vessels of lower extremities' And I80.3 'Phlebitis and thrombophlebitis of lower extremities, unspecified' Matthews A, Bhaskaran K. *Clinical Code List – Deep Vein Thrombosis [Data Collection]*. London, United Kingdom: London School of Hygiene & Tropical Medicine; 2018. https://doi.org/10.17037/DATA.00000733	All surgical admissions. Limited to adults and non-maternity wards.
Emergency department visit	Exclude	Low	Data not consistently available		
Infection with multi-resistant germs	Exclude	Low	Definition unclear		
Missed discharge preparation	Exclude	Low	Coding in Datix not consistent between Trusts		
Patient falls	Include	Low	Datix	**Datix**: 'Slips, Trips and Falls' or 'Accident including slips, trips and falls' (T2), 'Patient Falls' (T4), 'Slips, trips, falls and collisions' (T3) and 'Fall' (TI)	Incidents affecting patients in any inpatient ward that can be mapped to roster and PAS.
Physiological/metabolic derangement	Exclude	Low	Definition unclear		
PU	Include	Low	Datix and PAS	**(A) Datix** 'Tissue damage' (T2), 'Pressure Damage/Moisture Lesion/Skin damage' (T4), 'Pressure sore/decubitus ulcer' or 'Skin' (T3) and 'Tissue Viability' (T1) (B) L89.2 Stage III decubitus ulcer or L89.3 Stage IV decubitus ulcer or L89.9 Decubitus ulcer and pressure area, unspecified, that is without mention of stage AHRQ indicator uses this definition (stage 3/4/unspecified).	(a) Incidents affecting patients in any inpatient ward that can be mapped to roster and PAS. (b) All surgical admissions. Limited to adults and non-maternity wards?
Sepsis	Use proxy instead	Low	Definition unclear	ICU admission from inpatient ward	All adults except maternity. Paediatrics as separate analysis
Shock or cardiac arrest	Use proxy instead	Low	Definition unclear		
Surgical wound infection	Exclude	Low	Definition unclear		

TABLE 32 Candidate nurse-sensitive outcomes based on Blume *et al.* 2021[36] (*continued*)

Patient outcome	Include/ exclude	Strength of evidence (Blume *et al.* 2021)	Data source	Definition	Denominator
Upper gastrointestinal bleeding	Exclude	Low	Definition unclear		
Urinary tract infection	Exclude	Low	Definition unclear		
Poor discharge status	Exclude	Based on expert interview in Blume's review – not studied yet	Data unavailable		
Central venous catheter occlusion	Exclude	Based on expert interview in Blume's review – not studied yet	Data unavailable		
Infection of vascular access site	Exclude	Based on expert interview in Blume's review – not studied yet	Definition unclear		
Mycosis	Exclude	Based on expert interview in Blume's review – not studied yet	Definition unclear	?Diagnosis code 'Mycoses' (B35–B49) where not listed in first consultant episode But there is probably a limit on what type of mycoses are plausibly caused by low staffing	

TABLE 33 Unit costs to be used in the analysis

Variable	Currency	Unit cost base case (£)	Low (£)	High (£)
***Staffing costs*[a]**				
RN Band 5	Per hour of staff time	43		
RN Band 6		53		
Assumed weighted average cost of replacement staff		45.35	42.70	53.31
RN Band 7		64		
NA		33.45	28.72	36.52
Hospital admission costs				
Re-admission[b]	Cost per re-admission	2906.50	1118.40	3468.43
Hospital stay[c]	Cost per day (excessbed-day)	270	248	309

a Unit Costs of Health and Social Care (2022).[146]
b National (NHS) reference costs for hospital stay.[145]
c Costs of re-admissions were based on the distribution of HRG codes.[130]

Notes
Base-case RN cost is based on mean salary and on-costs for band 5 and 6 (weighted 75% band 5 and 25% band 6 based on observed distribution of senior/junior staff). Low estimate is based on mean band 5 salary and on-costs, high estimate is based on mean band 6 salary with on-costs.
Base-case NA cost is based on mean salary and on-costs for band 4, low estimate based on the entry point for band 2, based on maximum annual salary for band 4 with on-costs.
Re-admission cost was estimated as the weighted average of HRG costs for re-admissions observed in study data set. Low value is the 25th percentile of this distribution, high value is the 75th percentile.
Hospital cost per day (for reduction in LOS) was estimated as the weighted average of excess bed costs reported in National Tariff for 21/22 (weighted by activity reported to generate costs). Low value is the 25th percentile of this distribution, high value is the 75th percentile.

TABLE 34 Abridged life table life expectancy estimates, discounted life expectancy, quality-of-life weights and DANQALE

Age group	Life expectancy		Discounted life expectancy		At least one condition	Discounted quality-adjusted life expectancy	
	F	M	F	M		F	M
0	82.66	78.93	27.85	27.61	0.855	23.8	23.6
1–4	81.94	78.26	27.81	27.57	0.855	23.8	23.6
5–9	77.98	74.31	27.55	27.28	0.855	23.6	23.3
10–14	73.01	69.33	27.17	26.85	0.855	23.2	23.0
15–19	68.03	64.36	26.72	26.34	0.855	22.8	22.5
20–24	63.09	59.46	26.20	25.75	0.855	22.4	22.0
25–29	58.15	54.61	25.57	25.05	0.855	21.9	21.4
30–34	53.23	49.78	24.83	24.24	0.814	20.2	19.7
35–39	48.35	44.99	23.97	23.28	0.809	19.4	18.8
40–44	43.52	40.26	22.95	22.17	0.763	17.5	16.9
45–49	38.75	35.62	21.77	20.89	0.757	16.5	15.8
50–54	34.08	31.09	20.41	19.42	0.739	15.1	14.4
55–59	29.49	26.69	18.85	17.76	0.740	13.9	13.1
60–64	25.03	22.44	17.07	15.90	0.726	12.4	11.5
65–69	20.75	18.41	15.09	13.87	0.741	11.2	10.3
70–74	16.67	14.67	12.90	11.72	0.718	9.3	8.4
75–79	12.83	11.19	10.55	9.45	0.697	7.4	6.6
80–84	9.40	8.18	8.17	7.25	0.638	5.2	4.6
85–89	6.48	5.68	5.90	5.25	0.595	3.5	3.1
90–94	4.16	3.78	3.94	3.60	0.595	2.3	2.1
95+	2.13	2.06	2.09	2.02	0.595	1.2	1.2

Appendix 3 Additional tables of results

TABLE 35 Sensitivity analysis – different 'exposure windows' for low staffing

	HR	*p*	LCL	UCL
Five-day exposure window				
SHMI	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.079	0.000	1.070	1.089
Days of low NA HPPD	1.072	0.000	1.062	1.081
Ward random effect (SD)	1.205		AIC	BIC
Model fit			788,843	790,453
Three-day exposure window				
SHMI	1.065	0.000	1.064	1.066
Days of low RN HPPD	1.068	0.000	1.054	1.083
Days of low NA HPPD	1.064	0.000	1.050	1.079
Ward random effect (SD)	1.216		AIC	BIC
Model fit			777,384	778,949
Ten-day exposure window				
SHMI	1.061	0.000	1.060	1.061
Days of low RN HPPD	1.080	0.000	1.074	1.086
Days of low NA HPPD	1.069	0.000	1.063	1.075
Ward random effect (SD)	1.228		AIC	BIC
Model fit			792,880.9	794,512.3

TABLE 36 Effects of staff mix on mortality – survival models

	HR	*p*	LCL	UCL
RN skill-mix				
SHMI	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.083	0.000	1.072	1.093
Days of low NA HPPD	1.067	0.000	1.057	1.078
Proportion of RNs (10%)	1.013	0.100	0.997	1.030
Ward random effect (SD)	1.206		AIC	BIC
Model fit			788,842	790,461
Grade mix				
SHMI	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.082	0.000	1.072	1.092
Days of low NA HPPD	1.070	0.000	1.061	1.080
Proportion of senior RNs (10%)	0.985	0.005	0.974	0.995
Proportion of senior NAs (10%)	0.979	0.087	0.955	1.003
Ward random effect (SD)	1.195		AIC	BIC
Model fit			788,814	790,439
Temporary staffing (bank)				
SHMI	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.081	0.000	1.071	1.091
Days of low NA HPPD	1.079	0.000	1.069	1.089
Proportion of bank RNs (10%)	1.024	0.003	1.008	1.040
Proportion of bank NAs (10%)	1.023	0.000	1.015	1.032
Ward random effect (SD)	1.211		**AIC**	**BIC**
Model fit			788,796	790,425
Temporary staffing (agency)				
SHMI	1.063	0.000	1.062	1.064
Days of low RN HPPD	1.081	0.000	1.071	1.090

continued

TABLE 36 Effects of staff mix on mortality – survival models[6] (*continued*)

	HR	*p*	LCL	UCL
Days of low NA HPPD	1.077	0.000	1.067	1.086
Proportion of agency RNs (10%)	1.032	0.000	1.019	1.044
Proportion of agency NAs (10%)	1.043	0.000	1.023	1.062
Ward random effect (SD)	1.201		AIC	BIC
Model fit			788,808	790,434
Model fit with no staff mix terms			*788,843*	*790,453*

NA, nursing assistant staff; LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

TABLE 37 Association between net hours of understaffing and outcomes, adjusted for risk

	Mortality				Re-admission				LOS			
Term	HR	*p*	LCL	UCL	HR	*p*	LCL	UCL	Ratio[a]	*p*	LCL	UCL
All staffing												
SHMI (risk)	1.064	0.000	1.064	1.065	1.006	0.000	1.005	1.007	1.033	0.000	1.032	1.033
RN HPPD (net)	1.002	0.192	0.999	1.005	1.001	0.359	0.999	1.003	1.007	0.000	1.006	1.008
RN HPPD (net)	1.006	0.007	1.001	1.010	0.995	0.000	0.993	0.997	1.004	0.000	1.002	1.005
Ward random effect (SD)	1.221		AIC	BIC	0.654		AIC	BIC	0.964		AIC	BIC
Model fit			790,123	791,739			2,426,566	2,428,480			7,092,972	7,093,040
Excluding ICU												
SHMI (risk)	1.066	0.000	1.065	1.067	1.006	0.000	1.005	1.007	1.032	0.000	1.032	1.033
Low RN hours (net)	1.009	0.000	1.005	1.013	1.001	0.518	0.999	1.003	1.007	0.000	1.006	1.008
Low NA hours (net)	1.004	0.091	0.999	1.008	0.995	0.000	0.993	0.997	1.004	0.000	1.002	1.005
Ward random effect (SD)	1.200		AIC	BIC	0.613		AIC	BIC	0.970		AIC	BIC
Model fit			702,983	7,045,434			2,253,269	2,255,122			6,520,622	6,520,690

LCL, lower 95% confidence limit; NA, nursing assistant staff; UCL, upper 95% confidence limit.
a Gamma regression exponentiated coefficient is a ratio and the exposure is the mean staffing relative to ward mean.

TABLE 38 Nonlinear and interaction effects in mortality – survival models

	All staffing			
	HR	***p***	**LCL**	**UCL**
Net hours of low staffing (relative to ward mean)				
Interactions				
SHMI	1.064	0.000	1.064	1.065
Low RN HPPD (net hours)	1.002	0.128	0.999	1.005
Low NA HPPD (net hours)	1.006	0.003	1.002	1.010
RN X NA HPPD interaction	1.000	0.000	1.000	1.000
Ward random effect (SD)	1.479		AIC	BIC
Model fit			790,114	791,771
Nonlinear effects				
SHMI	1.064	0.000	1.064	1.065
RN HPPD (net) linear	1.003	0.046	1.000	1.006
Quadratic	1.000	0.624	1.000	1.000
Cubic	1.000	0.276	1.000	1.000
NA HPPD (net) linear	1.003	0.186	0.999	1.007
Quadratic	1.004	0.000	1.003	1.005
Cubic	1.000	0.000	1.000	1.000
Ward random effect (SD)	1.248		AIC	BIC
Model fit			789,979	791,633
Fit for linear model with no interaction			790,119	791,767
Days of low staffing (relative to ward mean)				
Nonlinear effects				
SHMI	1.063	0.000	1.062	1.064
Low RN HPPD (day)	1.089	0.000	1.079	1.100
Low NA HPPD (day)	1.081	0.000	1.071	1.091

continued

TABLE 38 Nonlinear and interaction effects in mortality – survival models (*continued*)

	All staffing			
	HR	***p***	**LCL**	**UCL**
RN X NA HPPD interaction	0.987	0.000	0.982	0.993
Ward random effect (SD)	1.459		AIC	BIC
Model fit			788,822	790,473
Nonlinear effects				
SHMI	1.063	0.000	1.062	1.064
RN HPPD (net) linear	1.072	0.000	1.056	1.089
Quadratic	1.024	0.000	1.012	1.035
Cubic	0.995	0.032	0.991	1.000
NA HPPD (net) linear	1.078	0.000	1.061	1.094
Quadratic	1.029	0.000	1.017	1.041
Cubic	0.992	0.000	0.987	0.996
Ward random effect (SD)	1.208		AIC	BIC
Model fit			788,798	790,442
Fit for linear model with no interaction			788,838	790,482

NA, nursing assistant staff; LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

TABLE 39 Association between days of understaffing and outcomes for ward and patient subgroups

	Mortality				**Re-admission**				**LOS**			
Subgroup	**HR**	***p***	**LCL**	**UCL**	**HR**	***p***	**LCL**	**UCL**	**Ratio**[a]	***p***	**LCL**	**UCL**
All staffing												
SHMI (risk)	1.063	0.000	1.062	1.064	1.006	0.000	1.005	1.007	1.032	0.000	1.031	1.032
Days of low RN HPPD	1.079	0.000	1.070	1.089	1.010	0.000	1.005	1.016	1.687	0.000	1.666	1.707
Days of low NA HPPD	1.072	0.000	1.062	1.081	0.994	0.020	0.988	0.999	1.608	0.000	1.589	1.627
Ward random effect (SD)	1.205		AIC	BIC	0.657		AIC	BIC	0.902		AIC	BIC
Model fit			788,843	790,453			2,424,881	2,426,790			7,025,587	7,025,655
General (non ICU) wards												
SHMI (risk)	1.064	0.000	1.063	1.065	1.006	0.000	1.005	1.007	1.032	0.000	1.031	1.032
Days of low RN HPPD	1.088	0.000	1.078	1.098	1.011	0.000	1.006	1.017	1.664	0.000	1.644	1.684
Days of low NA HPPD	1.077	0.000	1.067	1.086	0.994	0.024	0.988	0.999	1.605	0.000	1.586	1.624
Ward random effect (SD)	1.209		AIC	BIC	0.641		AIC	BIC	0.916		AIC	BIC
Model fit			823,354	824,983			2,404,337	2,406,224			6,922,804	6,922,871
General (NEWS < 5) patients												
SHMI (risk)	1.064	0.000	1.063	1.065	1.006	0.000	1.005	1.006	1.030	0.000	1.029	1.030
Days of low RN HPPD	1.094	0.000	1.081	1.106	1.011	0.000	1.005	1.018	1.625	0.000	1.602	1.649
Days of low NA HPPD	1.087	0.000	1.075	1.099	0.995	0.110	0.989	1.001	1.567	0.000	1.545	1.590
Ward random effect (SD)	0.961		AIC	BIC	0.612		AIC	BIC	0.920	NA	AIC	BIC
Model fit			497,531	498,792			1,769,632	1,771,200			4,937,519	4,937,585
ICU												
SHMI (risk)	1.046	0.000	1.043	1.049	0.999	0.707	0.992	1.006	1.017	0.000	1.014	1.020
Days of low RN HPPD	0.999	0.974	0.962	1.039	0.970	0.248	0.921	1.021	1.078	0.007	1.021	1.138
Days of low NA HPPD	1.010	0.601	0.972	1.050	0.986	0.619	0.935	1.041	1.147	0.000	1.088	1.211
Ward random effect (SD)	1.027		AIC	BIC	0.378		AIC	BIC	0.583		AIC	BIC
Model fit			36,370	36,642			20,634	20,802			140,238	140,281

continued

TABLE 39 Association between days of understaffing and outcomes for ward and patient subgroups (*continued*)

	Mortality				Re-admission				LOS			
Subgroup	**HR**	***p***	**LCL**	**UCL**	**HR**	***p***	**LCL**	**UCL**	**Ratio**[a]	***p***	**LCL**	**UCL**
Highly acute (NEWS 5+) patients												
SHMI (risk)	1.050	0.000	1.048	1.052	1.001	0.514	0.998	1.004	1.032	0.000	1.031	1.032
Days of low RN HPPD	1.036	0.001	1.014	1.059	0.988	0.244	0.968	1.008	1.664	0.000	1.644	1.684
Days of low NA HPPD	0.978	0.033	0.958	0.998	0.974	0.008	0.955	0.993	1.605	0.000	1.586	1.624
Ward random effect (SD)	0.645		AIC	BIC	0.391		AIC	BIC	0.916	NA	AIC	BIC
Model fit			107,720	108,462			103,017	103,669			6,922,804	6,922,871
Older people's wards												
SHMI (risk)	1.053	0.000	1.052	1.054	0.993	0.000	0.992	0.994	1.019	0.000	1.018	1.019
Days of low RN HPPD	1.048	0.000	1.037	1.060	0.999	0.798	0.990	1.007	1.422	0.000	1.396	1.449
Days of low NA HPPD	1.047	0.000	1.036	1.059	0.986	0.001	0.978	0.995	1.447	0.000	1.421	1.474
Ward random effect (SD)	0.569		AIC	BIC	0.481		AIC	BIC	0.695		AIC	BIC
Model fit			481,098	482,323			805,751	807,111			2,415,199	2,415,260
Older people > 75												
SHMI (risk)	1.055	0.000	1.053	1.056	0.989	0.000	0.987	0.990	1.012	0.000	1.011	1.012
Days of low RN HPPD	1.019	0.061	0.999	1.039	1.005	0.550	0.989	1.021	1.220	0.000	1.190	1.252
Days of low NA HPPD	1.009	0.377	0.990	1.028	0.969	0.000	0.954	0.984	1.247	0.000	1.217	1.278
Ward random effect (SD)	0.369		AIC	BIC	0.382		AIC	BIC	0.512		AIC	BIC
Model fit			171,307	171,641			249,476	249,912			824,108.7	824,163

[a] Gamma regression exponentiated coefficient is a ratio and the exposure is the proportion of days in the patient stay with understaffing.

TABLE 40 Recorded harms and levels of harm

Type of harm	Tissue damage		Falls		Medication		Other		All	
Level of harm	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
None	19,082	(46.1%)	15,409	(72.7%)	15,295	(86.9%)	38,306	(79.1%)	88,092	(68.5%)
Low	21,096	(51.0%)	5335	(25.2%)	2207	(12.5%)	9372	(19.3%)	38,010	(29.6%)
Significant	940	(2.3%)	260	(1.2%)	90	(0.5%)	593	(1.2%)	1883	(1.5%)
Severe	255	(0.6%)	186	(0.9%)	12	(0.1%)	182	(0.4%)	635	(0.5%)
All	41,373	(100%)	21,190	(100%)	17,604	(100%)	48,453	(100%)	128,620	(100%)
Incidents per ‘000 patient days	8.4		4.3		3.6		9.8		26.1	

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library