The Cost-Effectiveness of Nurse Staffing and Skill Mix on Nurse Sensitive Outcomes

A Report for The National Institute for Health and Care Excellence

Professor Graham Cookson

Dr Andrew McGovern

University of Surrey

April, 2014

University of Surrey

One of the UK's leading professional, scientific and technological universities in the UK, the University of Surrey ranked 39th in the prestigious Top 100 List of the world's most international universities, part of the Times Higher Education (THE) World University rankings for 2013-14. Actively involved in successive research collaborations with industrial and research partners across Europe since the Fourth Framework Programme, the University of Surrey received funding for 160 projects in the Seventh Framework Programme, including 26 Marie Curie fellowships.

Department of Health Care Management & Policy

The Department of Health Care Management and Policy (DHCMP) at the University of Surrey has been involved in quality improvement interventions over the last 15 years, primarily for long term conditions in the UK and internationally. Our interests are how to measure quality and health outcomes from routine data, quality improvement and technology trials, and integrating the use of the computer into the clinical consultation.

Despite being a small group, we have over 150 full length peer review scientific research publications; in addition to over 100 other peer review journal articles, letters or editorials and in excess of this number of conference abstracts. We have direct links with an excellent group of international collaborators; and links through the primary care informatics working groups of IMIA and EFMI (the International and European informatics organisations).

The Economics group in DHCMP has 10 members and is led by Professor Graham Cookson. The principal focus of the group is on the determinants of health care provider's productivity, the efficiency and effectiveness trade-off in health care, and the role of the health care workforce in this relationship.

Acknowledgements

In November 2013, National Institute for Health and Care Excellence (NICE) was asked by the Department of Health (DH) and NHS England to develop new guideline outputs which focus on safe staffing. In March 2014, NICE commissioned Professor Graham Cookson and his team at the University of Surrey to produce an economic evaluation of the effects of nurse staffing and skill mix on outcomes of care. This report is the result of that work. The authors would like to thank: the team at NICE but in particular Ian Rodrigues, Jasdeep Hayre and Lorraine Taylor; Keith Hurst (Keith Hurst Associates); Jeremy Brindley-Codd and Neil Webb (Guy's and St. Thomas' Trust); Barbara Fittall (NHS England); and the members of the NICE Safer Staffing Advisory Committee for helpful comments and insights into the production of this report.

Any errors or omissions remain our own.

Table of Contents

List of Figures

List of Tables

1 Executive Summary

In November 2013, National Institute for Health and Care Excellence (NICE) was asked by the Department of Health (DH) and NHS England to develop new guideline outputs which focus on safe staffing. In March 2014, NICE commissioned this report which aims to estimate the cost-effectiveness of altering nurse staffing and skill mix on outcomes of care in acute, adult inpatient wards. Following a systematic Evidence Review [4], the Safer Staffing Advisory Committee set the scope of this report to consider three outcomes (falls, medication errors and missed care) for which there appeared to be good evidence that they were sensitive to nursing care and were amenable to measurement.

There are only four existing economic studies of nurse staffing and patient outcomes including one cost-effectiveness study. None of the studies are from the UK nor do they use ward level data. It is therefore necessary to estimate these effects using UK ward level data, but the quality of the available data is poor. There are no current, nationally representative datasets that combine ward level staffing with patient level characteristics and outcomes. A small number of trusts are making significant inroads into collecting such datasets and this should be encouraged.

However, from this analysis there is still some evidence that nurse staffing levels and skill mix do impact patient outcomes, which is congruent with the extant literature. There is great uncertainty as to the strength of these effects and whether these effects are linear (i.e. the same at all levels of nurse staffing) or subject to either diminishing returns (i.e. a tailing off effect) or thresholds. In terms of making recommendations, it is important to also consider that based upon this analysis the size of ward (the number of occupied beds) and the layout of the ward may also influence nurse staffing decisions and therefore outcomes, but these relationships may be confounded by unobserved patient factors.

It was not possible to combine the benefits of the intervention into a common metric (e.g. QALYs) therefore it is impossible to ascertain the overall cost-effectiveness of changing nurse staffing or skill mix. Considering the two outcomes considered in this economic evaluation (falls and medication errors) separately: the Incremental Cost-Effectiveness Ratios of increasing the proportion of nurses in the skill mix were £1,412 per fall prevented and £127,293 per medication error prevented. No relationship between the outcomes and staffing levels (either at the individual nurse/HCA level or overall) was found and therefore no cost-effectiveness analysis performed. However, much of the literature points towards a relationship between registered nurse numbers and improved outcomes.

On the basis of this analysis, from an NHS perspective, changing skill mix to reduce drug errors is unlikely to be cost-effective. However, changing skill mix to reduce falls pay be if the willingness to pay for them exceeded £1,412 per fall. A number of alternative strategies for achieving these outcomes are identified from the literature, which are available for greater cost per unit. On this basis, this intervention may represent value-for-money.

However, despite the findings being based upon the best available evidence, great caution should be exercised when using these results as there is great uncertainty as to the benefits of staffing interventions due to limited data. Further, this evaluation considered only two possible benefits: a reduction in falls and a reduction in medication errors. The only other study to calculate the cost-effectiveness of increasing nursing levels (from Australia) found the intervention to be cost-effective largely because they quantified many more outcomes, including mortality and life years saved from reducing failure to rescue events.

2 Introduction

2.1 The Role of Economic Evaluation in the NICE process

The NHS has limited resources and almost endless uses of those resources. Therefore, when a new intervention or technology is adopted some amount of the existing health care provision will be displaced. This is what economists refer to as the 'opportunity cost' of an intervention. To maximise society's health gain from the NHS's limited budget, and to make decisions on whether to adopt new interventions in a coherent and transparent manner an economic evaluation is performed.

NICE plays a central role in the process by advising the NHS on the (clinical) effectiveness and cost-effectiveness of health care interventions and technologies. An intervention is cost-effective if it generates more health gain than it displaces as a result of the additional costs imposed on the system. Sometimes a new intervention dominates the existing best practice by being both cheaper and more effective, in which case the outcome is clear. More often the proposed intervention is more expensive and may be more effective.

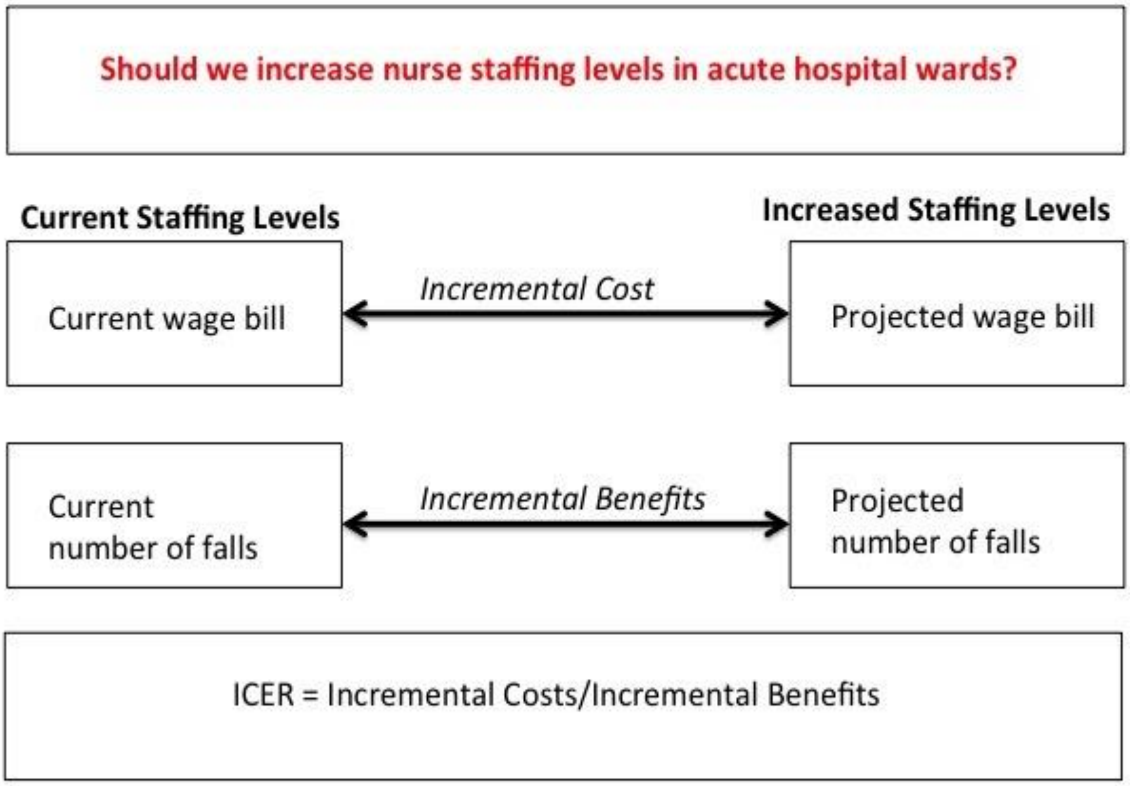
An economic analysis is usually required because the costs and/or benefits of a new intervention are uncertain. There are numerous reasons for this uncertainty. For example, there may be several small-scale studies reporting conflicting levels of effectiveness of a new treatment, or the context or population of these studies may not be wholly representative of the NHS patient population. Alternatively, widespread adoption of a new intervention may alter the market and therefore the price of the intervention. Frequently, the costs of an intervention are borne today but the benefits occur over several years into the future. All of these situations require careful modelling to enable a fair comparison of alternative outcomes. Inevitably, the economist must make assumptions about the most plausible values of the costs and benefits of an intervention based upon the best available evidence.

To illustrate the impact of these assumptions on the results of the economic analysis a sensitivity analysis is performed. This technique varies the main assumptions used to produce the base case to include plausible but extreme values of these assumptions. If varying these assumptions has little effect on the result of the economic analysis then we can be confident that the findings are robust and representative of the truth. If the results of the economic analysis vary considerably during the sensitivity analysis then additional research or evidence may be required to establish the truth, and less weight should be given to the economic evaluation in any decision making process.

NICE prefers that cost-effectiveness is reported as a cost per quality-adjusted life year (QALY) because this enables comparisons across different disease areas, populations or even between service level and disease-specific treatments to be made on a common metric. Additionally, it has the benefit of combining the multiple benefits of an intervention into a single outcome measure. QALYs are measured by estimating the health utility or value of being in different health states (where 1 is equivalent to a notional health state of perfect health and 0 is being dead) and are combined with the length of time spent in each of these health states as a result of the intervention. When it is not possible to measure QALYs, it is appropriate to report the benefits of the intervention in terms of some disease or topic specific outcome. For example, in terms of increasing ward level staffing the outcome may be the number of falls prevented.

Once the costs and benefits of an intervention have been measured, calculating the cost-effectiveness of the proposed intervention is straightforward as Figure 1 illustrates. It is usual to compare the new intervention with current or best practice. Dividing the incremental or additional costs by the incremental or additional benefits produces the Incremental Cost Effectiveness Ratio (ICER).

Figure 1 Simplified calculation of Incremental Cost-Effectiveness Ratios



As a concrete example, consider a hypothetical situation where the increase in staffing intervention was to add one additional nurse per ward at a cost of £31,867[1] per annum and in one year the only effect was to reduce the number of falls by 4. The ICER in this example would be £7,967 per averted fall.

If the new intervention is less effective *and* more costly than existing practice it is not cost-effective, and if it is more effective *and* cheaper than existing practice it is cost-effective. In these circumstances the outcome is straightforward. Usually however, the new intervention is either less effective but significantly cheaper, or more effective but also more expensive. In these circumstances the ICER is compared to the value of the interventions or treatments which are displaced if the new intervention is adopted: the opportunity cost. This is usually thought to be in the region of £20,000-£30,000 per QALY. There is little guidance available when the ICER is expressed in the original units of effects (e.g. falls prevented) and careful consideration needs to be given as to the value-for-money represented by the intervention in these situations.

2.2 Safer Staffing

Ensuring that staffing levels are sufficient to maximise patient safety and quality of care, whilst optimising the allocation of financial resources, is an important challenge for the NHS. The National Institute for Health and Care Excellence (NICE) has been asked by the Department of Health and NHS England to develop an evidence-based guideline on safe and cost-effective nurse staffing levels in acute adult inpatient wards. The definition of nursing includes both registered nurses and unregistered nurses, more commonly called health care assistants in the NHS.

A number of negative outcomes arising from low nurse staffing levels have been identified by systematic reviews [1-3] including the systematic review commissioned alongside this economic evaluation [4]. These reviews have demonstrated associations between many patient outcomes and nursing-to-patient ratios, and the skill mix of the nursing team. Studies demonstrate possible associations between nurse staffing level and infections, falls, pressure ulcers, medication errors, and missed care [4]. There is also some limited evidence to suggest that skill mix is important in that a higher proportion of health care assistants (HCAs) is associated with higher rates of adverse outcomes [4]. However there is a paucity of UK based studies and this is a substantial limitation of

[1] This figure calculated by adding the mean annual basic salary (excluding overtime) of an Agenda for Change Band 5 nurse of £25,744 to the mean on-costs of employing the nurse of £6,123 taken from the Personal Social Services Research Unit costings for July 2013-June 2013. It excludes overheads, capital costs, overtime, London weightings or training and qualification costs.

the existing literature [4]. Furthermore the differences between the UK health care system and those elsewhere result in severe limitations when applying data from the existing economic analyses to the UK health care system.

There is very limited evidence on the cost-effectiveness of increasing nurse staffing or altering the nursing skill mix. The Evidence Review [4] found four main papers that tackle the economics of nurse staffing, none of which come from the UK, and none of which adopt a ward level perspective. Only one of the papers produced a cost-effectiveness measure, the remaining three compared the costs of increasing nurse staffing with the cost savings from reducing length of staff and avoiding adverse outcomes.

Following the Evidence Review [4], the NICE Safer Staffing Advisory Committee identified that three main adverse outcomes – medication errors, falls and missed care – were both sensitive to nursing levels and skill mix and were measurable in the English NHS. While numerous other outcomes were considered the evidence was either weak or ambiguous, or the contribution of the nursing team relatively small.

As the existing economic evidence is not from the NHS and does not provide estimates for ward level interventions in staffing and skill mix, it is of limited use to this evaluation. The economic evaluation will have to use the best available evidence and data from the UK to determine the relationship between nursing levels and skill mix and nurse sensitive outcomes, before evaluating the cost-effectiveness of altering staffing or skill mix.

2.3 Purpose of this report

This report aims to assess the cost-effectiveness of altering nurse staffing levels and skill mix in the English NHS. It accompanies the Evidence Review produced by the University of Southampton [4].

3 Methods

3.1 Economic Model Scope

Following the systematic Evidence Review [4] performed by Griffiths et al. at the University of Southampton, the Safer Staffing Advisory Committee restricted the scope of the economic analysis to three main outcomes thought to be sensitive to nursing care and amenable to measurement: falls, medication errors and missed care.

The formal scope of the economic evaluation was agreed as:

Population:	Acute, adult, inpatient wards in English NHS Trusts
Interventions:	Increasing nurse staffing levels
	Altering nurse skill mix to increase the proportion of Registered Nurses
Comparators:	"Current" practice – where "current" is defined by the available datasets
Outcomes:	Incremental cost per fall prevented
	Incremental cost per drug error avoided
	~~Incremental cost per unit of missed care avoided~~[2]
Perspective:	National Health Service and Personal Social Services
Evaluation method:	Cost-Effectiveness Analysis (CEA)
Time:	One year. No discounting is required.
Valuing Benefits:	A utility measure (e.g. QALY) is not available nor appropriate in this setting.
Evidence Synthesis:	The results from the Evidence Review by Griffiths et al. [4] will inform the statistical and economic modelling.

3.2 CEA Methodology

The Cost-Effectiveness Analysis (CEA) will estimate the incremental cost of (i) averting one fall, and (ii) avoiding one drug error. Due to a lack of data on missed care, this outcome will not be considered in this report. The two outcomes will be considered separately due to a lack of common metric (e.g. QALYs or money). The unit of analysis will be a single, representative ward the characteristics of which will be determined by the statistical analysis.

[2] As will be made clear later in the report, missed care could not be considered as an outcome in this study due to a lack of data.

Table 1 lists the parameters used in the CEA and, taking falls as an example, the CEA uses them in the following steps:

Incremental Cost Effectiveness Ratio: Incremental cost/incremental benefit

Incremental benefit: effectiveness of intervention x exposure

Effectiveness: change in the fall rate (falls per 1,000 occupied bed days)

Exposure: 1,000s of bed days per ward per year

As the intervention is an increase in nurse numbers or a change in the proportion of registered to unregistered nurses, it will only be necessary to calculate the incremental cost and not the baseline cost as the remainder of the cost is still incurred after the intervention. For example, if we consider increasing registered nurse staffing by 10% from 10 to 11 FTE, then the incremental cost is the wage of 1 FTE nurse because both the current practice and intervention will incur the cost of the other 10 FTE nurses.

The following assumptions are also made:

- That the data used in the statistical analysis is representative of wards in English NHS trusts i.e. that there is no selection bias
- That there has been accurate recording of the outcomes
- That the relationship between staffing levels, skill mix and nurse sensitive outcomes is common across time (i.e. it hasn't changed in the past decade) and across trusts.
- That any unobserved patient, ward or trust level characteristics do not confound the results
- That the relationships are linear
- That any effectiveness results adopted from the literature are applicable to an NHS, ward level setting

The importance of these assumptions for the validity of the findings and the likelihood that they hold are discussed in Section 5. The impact of these assumptions on the CEA cannot be modelled through sensitivity analysis. However, the findings are also compared across the different datasets, alternative model specifications, alternative regression methods, and with the extant literature in Section 5 to consider the plausibility of the results. The computer code used to generate the statistical results and the CEA calculations have been checked by both authors, and another colleague from the Department of Health Care Management & Policy. Finally, a sensitivity analysis is performed in Section 4 to determine the sensitivity of the findings and conclusions to the values chosen for the parameters in Table 1.

Parameter	Definition	Source and value
Exposure	1,000s bed days (adjusted beds x 365 days)	Audit Commission Dataset (2004).
Effectiveness	Change in the number of (i) falls and (ii) drug errors per 1,000 occupied bed days	Results Section 4.1.3.
Nurses	WTE registered nurses per occupied bed day	Audit Commission Dataset (2004).
HCAs[3]	WTE HCAs per occupied bed day	Audit Commission Dataset (2004).
Cost	The cost per WTE (i) nurses and (ii) HCAs	Public and Social Services Research Unit (2013). Section 3.3

Table 1 Economic Model Parameters and Sources

3.3 Costs

From an NHS perspective, only direct costs are considered. As this is a nurse staffing intervention this is understood to be the wage plus the on-costs (employer's national insurance and pension contributions). Overtime, training costs, and capital costs are excluded. Costs are taken from PSSRU's Unit Cost of Health and Social Care 2013 report [20] and are national averages in UK pounds for the period July 2012 to June 2013. The employment costs which are reported in Table 2 can be weighted for London trust by multiplying by a factor of 1.19 or reduced for trusts outside London by multiplying by a factor of 0.96.

Table 2: NHS Employment Costs – Source: PSSRU 2013.

Grade	AfC Band	Wage/Salary	On-Costs	Total Cost
Sister/Ward manager	7	£38,057	£9,546	£47,603
Charge nurse/Deputy Ward manager	6	£31,752	£7,794	£39,546
Staff nurse	5	£25,744	£6,123	£31,867
Health Care Assistant	2	£16,193	£3,468	£19,661

[3] The skill-mix of the ward is derived from these measures calculating the proportion of total staff WTE per bed that are HCAs.

The NPSA [16] report the estimated NHS cost of different types of fall (e.g. no harm) as well as the proportion of falls falling into each category. These can be used to produce a weighted average fall cost of £79.22, see Table 3, which uses current reference cost data to estimate the cost involved in treated different severity of injuries resulting from harm. The no harm cost is an hour of staff time to comfort the patient and return them to bed. The low harm is the same as no harm but with an additional 30 minutes to administer first aid and £5 to cover dressings. The moderate harm is the same as low harm plus the lowest cost of non-admitted accident and emergency care (e.g. sprains). The severe harm is an average of the reference costs for hip and knee fractures. There are no similar NHS costs available for medication errors.

Type of Harm	**% of Cases**	**NHS Cost per Case**
No Harm	67%	£41.00
Low Harm	30%	£65.50
Moderate Harm	3%	£324.00
Severe Harm	1%	£3,135.00
Death	0.1%	£0.00
Average		**£79.22**

Table 3 NHS Costs of Falls – Source: NPSA, 2007

3.4 Evidence of effectiveness of interventions

There are few existing economic evaluations of interventions to alter nurse staffing levels and/or skill mix. A complete review of these studies was provided in the Evidence Review [4], and this will not be replicated here. Five studies [7-11] were identified in the Evidence Review [4] and none of them provide robust evidence of the cost-effectiveness of staffing interventions.

The reasons for excluding these studies are outlined briefly. One of them [8] failed to measure costs at all. The remaining studies were not from the UK but from Australia [11] and the USA [7-10]. This is a problem due to the nature in which nursing care is provided but also the way in which the costs of adverse nurse sensitive outcomes are calculated. Further, none the studies used ward level data or relationships in their analysis, but relied on a mixture of patient level administrative data and hospital level staffing data. Therefore there are limited inferences that can be drawn about the effect of changing ward level staffing on patient outcomes from these studies.

Overall, there was little consistency across the four studies [8-11] due to the different data, methods, settings, perspective, base years, or range of effects considered. Unsurprisingly the net costs are positive when deducting the reduced expenditure (from reducing mortality, failure to rescue, length of stay and adverse events) and comparing this to the costs of increased nursing input. However, what is required for resource allocation decision making is convincing evidence of the cost-effectiveness of improving these outcomes relative to the next best alternative method of achieving these outcomes.

Only Twigg et al. [11] attempted to produce an incremental cost-effectiveness measure which was AUS$8907 per life year gained (c. GBP£5,500[4]) following the introduction of mandated minimum Nurse Hours Per Patient Day (NHPPD) in Western Australia. While this isn't expressed as a cost per QALY, unless the health status of those affected was extreme this value would compare favourably with the £20,000-£30,000 opportunity cost figure typically used in NICE economic evaluations. The approach used in Twigg et al. [11] is to compare the before and after adverse events (and their associated hospital costs) across 3 hospitals in Western Australia following the introduction of minimum NHPPD. They use a risk adjustment model to predict the expected number of adverse events that should have occurred in hospitals in the post intervention period in the absence of the intervention (the counterfactual) and ascribe any divergence between the predicted and actual rates of incidence to the increase in staffing. This is therefore a simple difference approach at a macro level. As no policy shift is observed in the UK, such an approach cannot be used. Further, as the minimum staffing number is variable across different wards dependent upon the "complexity" of the patients it isn't possible to estimate the relationship between the effectiveness of the intervention at a micro level.

The remaining papers [8-10] use effectiveness results from existing papers, mainly Needleman's seminal work [12], to establish the relationship between nurse staffing and a rangle of nurse sensitive outcomes. None of the studies report effects for missed care however. Further, only Dahl [8] measures the effect of nursing levels and skill mix on medication errors and falls. The elasticity or responsiveness of the outcomes to a 1% increase in Registered NHPPD (from a median of 7.8 NHPPD) is reported in Table 1 of [8] as -0.71% for falls and -0.06% for medication errors which are considered to be "strong" and "moderate" effects respectively according to the authors. The source for the falls effect is [19] which finds a 10% increase in the number of licenced nurses[5] reduces the number of falls by 3% per annum. The source for the medication error effect is [7] which finds that a 1% increase in NHPDD reduces medication errors by 0.06%, although this effect is not statistically

[4] Using the average 2013 exchange rate of 0.619 it was equivalent to £5513.

[5] Defined as the combination of Registered Nurses and Licensed Practical Nurses.

significant. Both of these effects are also derived from US data and are based on aggregated, hospital level staffing data.

Given the limited relevance of the existing literature, alongside the poor quality of the results, it will be necessary to generate effectiveness measures before the cost-effectiveness can proceed. The next section details the data sets and methods used to determine the effects of altering staffing levels and skill mix on nurse sensitive outcomes using ward level data.

3.5 Effectiveness of Staffing and Skill Mix on Outcomes

There are no ward level studies of the relationship between staffing level, skill mix and nurse sensitive outcomes which can be used as measures of the effectiveness of any policy which would alter these factors. A statistical analysis was therefore performed on three ward level datasets that were provided by NICE to determine the association between nurse staffing and nurse sensitive outcomes. These comprise the following; data for the Audit Commission (AC) investigation of staffing levels in 2004, UK Nursing Database of nursing quality audit data collected and provided by Keith Hurst Associates (UKND), and staffing level and outcome data provided by two large NHS foundation trusts (NFT). All the data contained in these datasets are recorded from UK hospitals, although only data from English trusts were included in the analysis. These dataset were analysed independently. They are subsequently referred to by their associated acronyms; AC, UKND, and NFT.

3.5.1 Audit Commission ward data

The AC datasets were collected in 2004 from hospitals across England. The following ward information was collected; ward primary specialty, total number of beds in primary specialty, number of those which are closed at weekends, number of those which are day-beds, ward secondary specialty, total number of beds in secondary specialty, number of those which are closed at weekends, number of those which are day beds, number of single rooms, and ward type (bays or Nightingale). Ward level data were collected on several measures of staffing; WTE (whole time equivalent) worked by staff band (bands A-H; before the adoption of the Agenda for Change system [5]), staff overtime and leave hours. Some patient factors were recorded; number of days with high dependency patients (defined as patients requiring one to one nursing), and number of admissions, discharges, transfers in and transfers out of each ward. Some outcome measures were also collected at ward level; patient falls, other incidents to patients (composite outcome), drug errors (medication administration errors), needlestick injuries, and other incidents to staff (composite outcome).

The AC dataset was collected via distribution of a national survey, whereby chief nurses and financial directors were responsible for data collection in individual hospitals. Varying data collection methods were used in each hospital. These methods were not recorded in the dataset provided. All ward data were collected over a four week period. Outcome values reported are totals over this four week period e.g. falls per ward equates to total number of falls per ward during a four week period. Number of admissions, discharges, and transfers are also reported for this time period. All these variables were converted into number of events per 1000 adjusted bed days (e.g. number of falls per 1000 bed days or number of discharges per 1000 bed days).

This dataset is used to describe nursing trends across wards, by ward type (medical, surgical, or mixed medical and surgical wards), across trusts, and within hospitals. Staffing associations with the two outcome measures were also studied; patient falls and drug errors after adjusting for hospital, ward, and patient factors. Wards were included in the analysis if they had complete data on staffing levels, bed numbers, and ward type.

A complete list of the variables included in the analysis is presented in Table 4. Admissions, discharges, and transfers were input into the statistical analysis as separate variables (as opposed to generating a single patient turnover variable) to explore the individual effects of each component of patient movement. In practice patient transfers require a substantially different nursing approach than patient discharges or admissions.

Variable name	Description
Ward speciality	All included wards were categorised as medical, surgical, or mixed (medical and surgical)
Ward size	Total number of adjusted beds (see full definition below).
Ward layout	Ward layout categorised as bays, Nightingale, or not recorded
Proportion of single rooms	The proportion of beds which were single rooms
Proportion of days with high dependency patients	The proportion of days (during a four week period) during which the ward reported having one or more high dependency patient (requiring 1 to 1 nursing care)
Admissions per bed	Number of new admissions per 1000 adjusted bed days
Discharges per bed	Number of discharges per 1000 adjusted bed days
Transfers in per bed	Number of patient transfers on to the ward per 1000 adjusted bed days
Transfers out per bed	Number of patient transfers out of the per 1000 adjusted bed days
Total WTE per bed	Total whole time equivalent of nursing and support staff (health care assistants) per adjusted bed
Senior nurse proportion	Proportion of the total nursing and support staff who are senior nurses
Charge nurse proportion	Proportion of the total nursing and support staff who are charge nurses
Staff nurse proportion	Proportion of the total nursing and support staff who are staff nurses
HCA proportion	Proportion of the total nursing and support staff who are health care assistants

Table 4 A complete list of the variables included for statistical modelling from the data provided in the AC dataset.

A measure of adjusted bed days was derived. This was the average number of beds available during a 24 hour period per ward. This measure allows correction for beds which were closed at weekends and for beds which were only open during normal working hours (day beds). A mathematical derivation of this measure was as follows: The total number of beds (B_{TOT}), the number beds closed at weekends (B_{WE}), the number of days weekend beds were closed (d_{WE}), and the number of day beds (B_{day}) were recorded in the AC dataset. The number of adjusted bed days (B_{adj}) was calculated as the total number of beds excluding the proportion of bed days lost due to beds being closed at weekends (B_{WE}^{lost}) and the proportion of beds lost by beds being closed out of normal working out hours (B_{day}^{lost}):

$$B_{adj} = B_{TOT} - B_{WE}^{lost} - B_{day}^{lost} \qquad (1)$$

The number of beds lost at weekends is given by:

$$B_{WE}^{lost} = \frac{d_{WE} \times B_{WE}}{7} + \frac{d_{WE} \times B_{day}}{7} = \frac{d_{WE}(B_{WE} + B_{day})}{7} \qquad (2)$$

Here the factor of seven occurs because each weekend day the bed is closed equates to 1/7th of a day lost on average per week. It is assumed that day beds are also closed at weekends. The number of beds lost by beds being closed out of normal working out hours is given by:

$$B_{day}^{lost} = \frac{2}{3} \times \frac{B_{day}}{7} \times (7 - d_{WE}) = \frac{2(7 - d_{WE})B_{day}}{21} \qquad (3)$$

The factor of 2/3 accounts for the closure of day-beds for two-thirds of the 24 hour period. The factor $(7 - d_{WE})$ gives the number of days which day-beds are open (either 6 or 7 depending on the number of days the ward is closed over the weekend). Substituting equations 2 and 3 into equation 1 gives the total adjusted beds days from our recorded parameters:

$$\begin{aligned} B_{adj} &= B_{TOT} - \frac{d_{WE}(B_{WE} + B_{day})}{7} - \frac{2(7 - d_{WE})B_{day}}{21} \\ &= B_{TOT} - \frac{d_{WE}(3B_{WE} + B_{day})}{21} - \frac{2B_{day}}{3} \end{aligned} \qquad (4)$$

A linear regression analysis was performed to identify associations with the outcome variables. Additional methods of regression were used for comparison and to test the robustness of the model findings (quasi-Poisson regression and negative binomial regression; see appendix 1).

3.5.2 UK Nursing Database

The UKND dataset has been collected over an extended period of time; from 1985 to 2014, but with greatly increased sampling frequency in more recent years. Data have been collected from hospitals across England, Ireland, Scotland, and Wales although only data from England have been used in this analysis. Data on time of year were collected for each ward sampled, hospital type (teaching hospital, district general hospital, or community hospital), and ward type. Ward type was recorded as bays, bays and side rooms, side rooms only, Nightingale, Nightingale and bays, racetrack, split-site (on two or more floors) and other. For the analysis presented here these were re-categorised to include mixed wards (bays and side rooms; and Nightingale and bays) in the 'bays' category as the number of these mixed wards was too small for subgroup analysis. The dataset included wards from a variety of medical specialities (e.g. paediatrics) but only those relating to acute, adult inpatient wards were analysed.

Patient factors recorded comprise; total ward occupancy, and patient dependency by dependency bands (low, low-medium, medium, and high). Detailed staffing data were available including; WTE per occupied bed by staff grade and for bank staff, and time spent on different nursing and administration tasks. Six outcome measures of nursing quality were calculated; how well patients' needs were assessed, how well care addressed patients' needs, extent to which care plan was implemented, how well patients' response to care and progress was recorded, and ward fabric and resources score. These six outcome measures were composite outcomes generated by compiling several nursing standards [6]. An overall ward quality measure was generated by summing these six outcomes.

Each ward was assessed by an independent auditor over a period of one month. Patient dependency scores were measured for all patients admitted during this period to gain a representative sample of the ward population. Nursing activities were observed in 10 minute cycles by trained non-participatory observers. Nursing activities were categorised and used to calculate care quality outcome measures. Ward staff were observed during two early shifts, two late shifts, and two night shifts on each ward. Additional data collection was used to compile nursing quality scores including interviewing patients, relatives and staff, analysis of nursing records, and inspection of the ward environment and equipment. The full data collection methodology utilised has been described elsewhere [6].

We used this dataset to corroborate staffing trends identified using the AC dataset. This dataset also enabled us to look at trends in overall staffing levels over time. We investigated associations with staffing and the composite nursing outcome measures provided adjusting for hospital, ward and patient factors where available. This was performed using linear regression models. The variables included in these models are listed in Table 5.

Variable name	Description
Hospital type	All included hospitals were categorised as teaching, district general, or community hospitals
Ward layout	Ward layout categorised as bays, racetrack, single rooms, Nightingale, split site, or other (see text)
Ward occupancy	The mean number of patients present on each ward
Patient dependency	The proportion of patients by dependency group (low, low-medium, medium, and high) calculated based on nursing requirements
Senior nurse proportion	Proportion of the total nursing and support staff who are senior nurses
Staff nurse proportion	Proportion of the total nursing and support staff who are staff nurses
Support staff proportion	Proportion of the total nursing and support staff who are support staff (primarily HCAs)
Bank staff proportion	Proportion of the total nursing and support staff who were bank staff

Table 5 A complete list of the variables included for statistical modelling from the UKND database.

Wards from the UKND dataset with complete data on ward type, patient dependency and staffing levels were included in the analysis. Wards were excluded if they were not adult wards, or were palliative care, intensive care, high dependency, or respite wards. Linear regression analysis was used to identify variables associated with outcome measures. Additional regression models (e.g. logistic) were used for comparison and are reported in the appendices.

3.5.3 NHS Foundation Trust data

[Additional information to be added to this section following the consultation period of the draft guideline]

3.5.4 Generic statistical methods across datasets

For all the datasets analysed exploration for non-linear associations with staffing variables was performed by dividing these variables into quintiles and by using squared and cubic forms of the variables. All regression analysis was performed using the statistical package R. Statistical significance is assumed where p values are less than 0.05.

4 Results

4.1 Audit Commission dataset

This dataset comprises data from 7,522 wards collected between April and August 2004. Data for each ward were collected over four consecutive weeks. For this analysis wards were excluded if data were incomplete (N=605) or the ward was not an adult medical, surgical, or mixed medical and surgical ward (psychiatric, paediatric, ITU, respite, rehabilitation, etc ward excluded; N=3,209). The remaining wards (N=3,708) were sampled from 428 distinct sites across 211 of the then existing NHS trusts. These wards comprised 2,211 medical wards, 1,365 surgical wards, and 132 mixed medical and surgical wards. A few of these wards had more detailed information about the ward speciality; medical admissions - 351, surgical admissions - 112, stroke wards - 101. The majority (n=3,144) had no additional information about ward speciality.

4.1.1 Ward and trust level staffing trends

Information on total bed numbers, number of beds closed at weekends, number of day beds, and number of days the ward was open, were available for all the included wards. The total number of beds included was 92,568 with an average of 25.0 (SD ±7.3) beds per ward. This included 3,792 beds closed at weekends of which 1,372 were day beds. The adjusted number of beds available per day (adjusted bed days) was calculated for each ward to compensate for day beds and beds closed at weekends. A total of 91,052 adjusted beds were included with an average of 24.6 (SD ±7.5) beds per ward. No data on bed occupancy was available (this was removed from the original dataset due to poor recording). Data were available for all included wards on; total number of admissions, discharges and transfers in and out during the four month study period.

The average WTE per adjusted bed day was 0.991 (SD ±0.424) for medical wards, 0.976 (SD ±0.496) for surgical wards, and 0.972 (SD ±0.320) for mixed wards. The ward total WTE was correlated with the ward total adjusted bed days (Figure 2). Note that 1 WTE per adjusted bed day equates to approximately one staff member per 4.5 beds as staff only work 37.5 hours per week. 0.991 WTE per adjusted bed day therefore equates to approximately 5.29 staff Hours Per Patient Day (HPPD) for medical wards (5.2 HPPD for surgical wards and 5.18 HPPD for mixed wards).

The distribution of WTE per adjusted bed day was tightly clustered around the mean value with a few outlying wards with very high staffing ratios (Figure 3). These outlying wards were all small wards with 9 beds or fewer and listed as medical or surgical wards with no additional ward details reported.

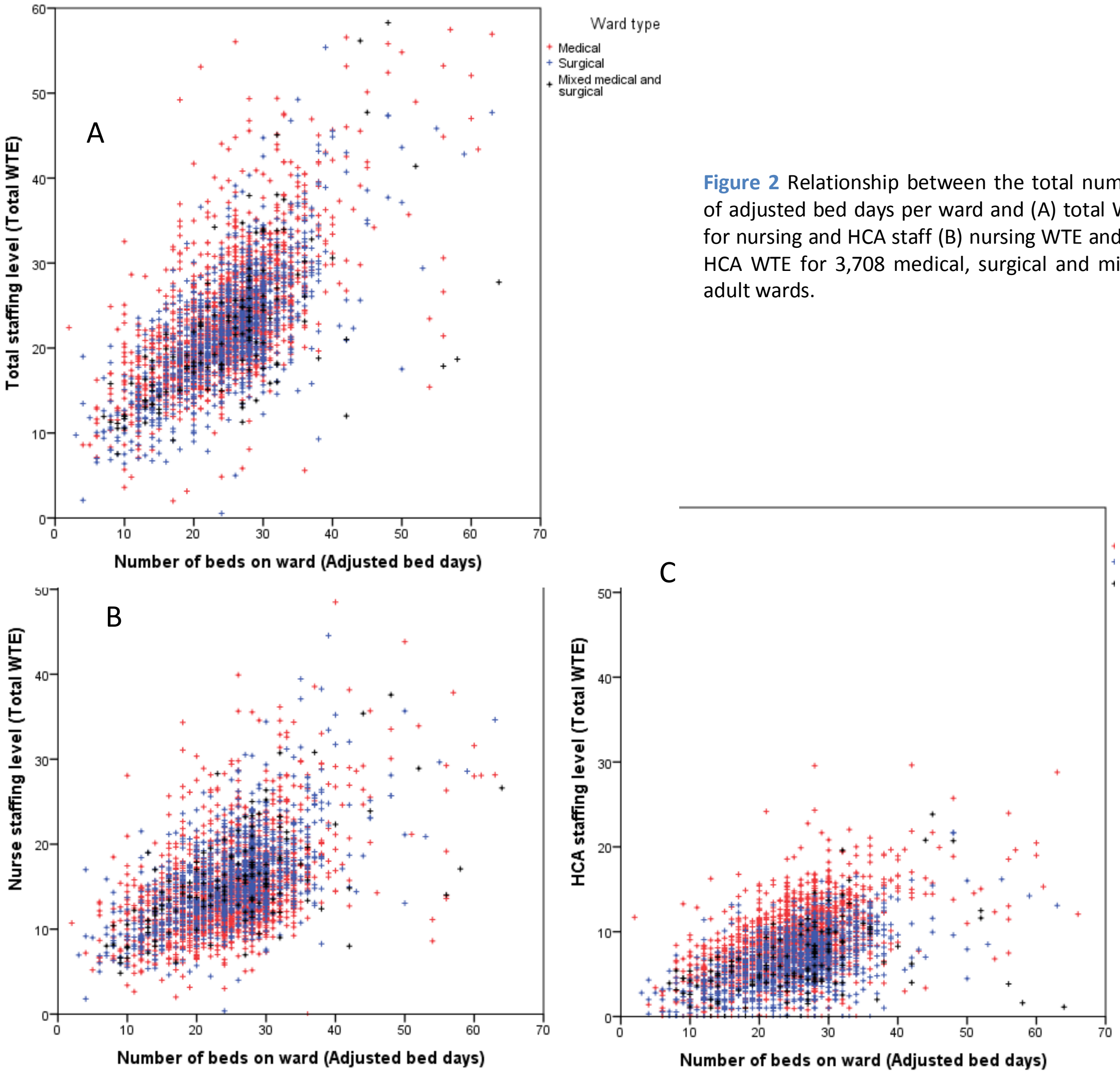
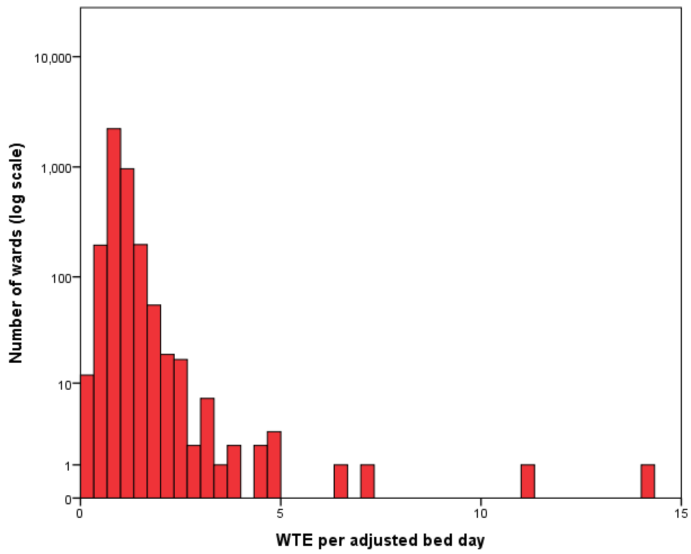


Figure 2 Relationship between the total number of adjusted bed days per ward and (A) total WTE for nursing and HCA staff (B) nursing WTE and (C) HCA WTE for 3,708 medical, surgical and mixed adult wards.

The average nursing WTE per adjusted bed day was 0.632 (SD ±0.324) for medical wards, 0.695 (SD ±0.374) for surgical wards, and 0.7086 (SD ±0.268) for mixed wards. This equates to an average of 3.37 nursing HPPD for medical wards, 3.84 HPPD for surgical wards, and 3.78 HPPD for mixed wards. This can be compared with the mandated minimum nurse HPPD in Western Australia of 5.0 NHPPD for moderately complex wards, 5.75 NHPPD for highly complex acute and 6.0 for high complexity wards[6].

Figure 3 Distribution of staffing levels across 3,708 medical, surgical, and mixed adult ward



Total WTE per adjusted bed was dependant on ward size. Larger wards had fewer staff per bed with a substantial increase in the number of staff per bed for the smallest wards; 10-12 beds or fewer (Figure 4).

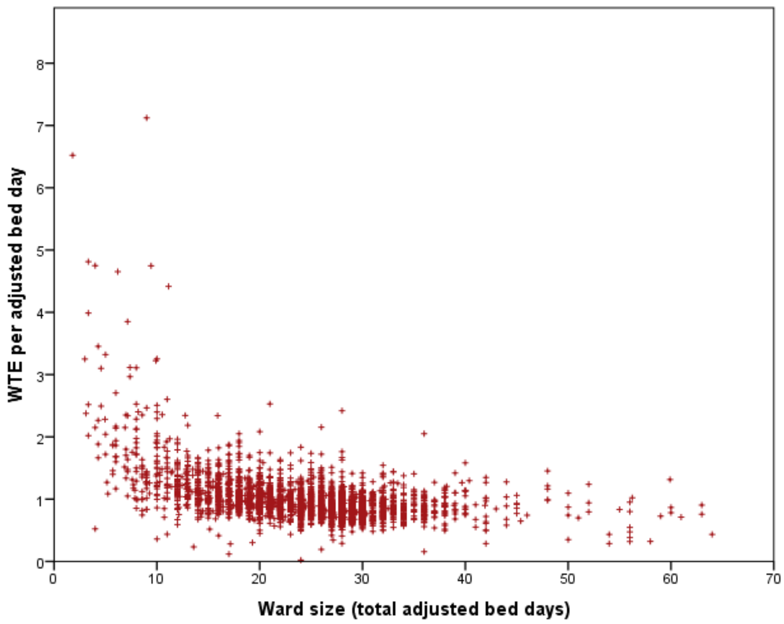


Figure 4 The relationship between staffing levels per bed and ward size.

[6] Moderately/highly complex acute wards are defined as having 35% patient turnover and 40/50% respectively emergency admissions. In comparison a high complexity ward would have high patient turnover at 50%, frequent (30 minute) monitoring and occasional 1:1 nursing.

The average number of medical, surgical or mixed wards sampled per trust was 17.6 (SD ±11.2; range 1 to 69). 188 trusts (of 211) had five or more wards sampled. Across these 188 trusts the mean ward size was 24.7 adjusted bed days (SD ± 3.5). There was little variation across trusts from the mean WTE per adjusted bed day of 0.97 (SD ±0.14). The distribution of mean WTE per bed day across trusts is shown in Figure 5. Excluding trusts with fewer than five wards sampled there was also little variation in the mean nursing WTE per adjusted bed day across trusts; 0.61 (SD ±0.21).

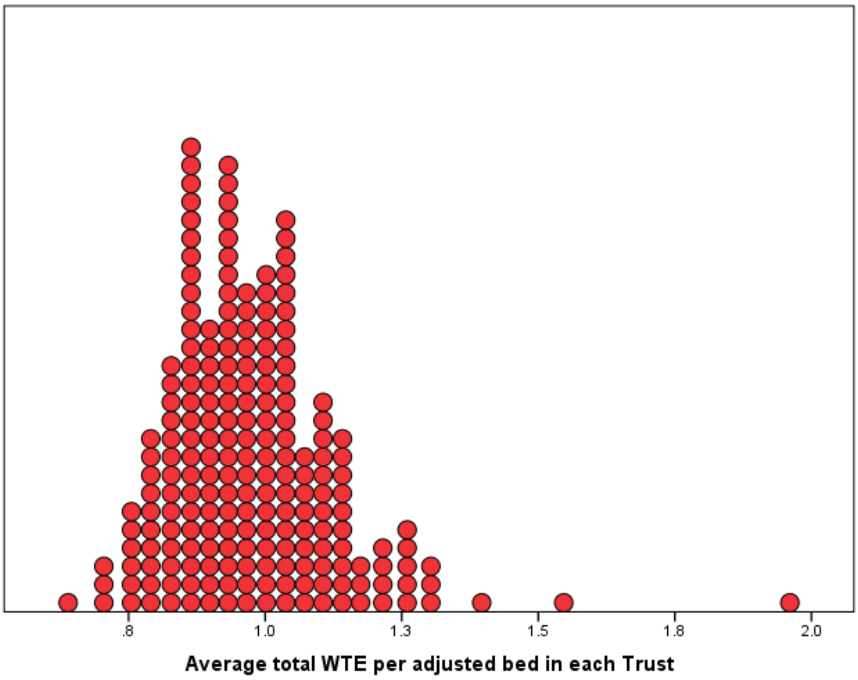


Figure 5 The distribution of mean WTE per bed day across the 211 sampled trusts. Each trust is represented by a circle. The most outlying trust with 1.9 WTE per bed only had one ward sampled.

4.1.2 Patient and ward outcome data

Several outcomes were measured at ward level. These comprised number of patient falls during the assessment period, other incidents to patients, medication errors, needlestick injuries, and other staff incidents. Ward outcomes were reported for between 77.9 and 88.8% of the included wards (

Table 6). The mean fall rate of almost 15 falls per 1,000 adjusted bed days is high compared to the national average reported by the NPSA [16] of 4.8 per 1,000. The data are heavily skewed with a small number of wards heavily influencing the mean, but the median rate is still 10.35 per 1,000 adjusted bed days. The medication error rate at 2.8 per 1,000 bed days is likely to suffer from serious underreporting with the hospital administration drug error rate estimated at 3-8% [18]. Medication error reporting is known to be problematic with underreporting and inaccurate reporting a common feature of the data.

Outcome measure	Wards reported N (%)	Mean per 1,000 adjusted bed days (SEM)	Incident total
Falls	3,471 (93.6)	14.91 (±16.2)	38,897
Other incidents to patients	3,293 (88.8)	10.03 (±15.5)	21,675
Drug errors	3,140 (84.7)	2.84 (±5.0)	5,989
Needlestick injuries	2,890 (77.9)	0.66 (±1.2)	1,314
Other incidents to staff	3,217 (86.8)	3.71 (±5.4)	8,039

Table 6 The number of outcome events per ward during a four week period for a total of 3,708 wards.

Some differences existed between the wards which were excluded because of incomplete outcome data and those included (

Table 7). There was a higher proportion of surgical wards excluded because of missing data on drug errors than included. The proportion of nurses and overall staffing levels were higher in the sample with data on drug errors and therefore included in the analysis than those with missing data. This may bias the findings slightly for this outcome.

	Analysis of patient falls		Analysis of drug errors	
	Excluded wards N (%*; 95% CI) or mean (SEM)	Included wards N (%*; 95% CI) or mean (SEM)	Excluded wards N (%*; 95% CI) or mean (SEM)	Included wards N (%*; 95% CI) or mean (SEM)
Ward type:				
Medical	352 (57.9%; 53.9% - 62.0%)	1,882 (59.9%; 58.2% - 61.7%)	118 (43.5%; 37.1% - 49.8%)	2,108 (60.7%; 59.1% - 62.3%)
Surgical	245 (39.1%; 35.0% - 43.1%)	1,143 (36.4%; 34.7% - 38.1%)	135 (50.6%; 44.3% - 57.0%)	1,245 (35.9%; 34.3% - 37.5%)
Mixed	25 (3.0%; 1.8% - 4.4%)	115 (3.7%; 3.0% - 4.3%)	21 (5.9%; 3.0% - 8.9%)	118 (3.4%; 2.8% - 4.0%)
Ward size (x=number of beds):				
x < 12	51 (6.9%; 4.9% - 9.0%)	150 (4.8%; 4.0% - 5.5%)	37 (11.4%; 7.6% - 15.6%)	64 (1.8%; 1.4% - 2.3%)
12 < x ≤ 20	161 (24.9%; 21.3% - 28.3%)	622 (19.8%; 18.4% - 21.2%)	183 (71.7%; 65.8% - 77.2%)	3,013 (86.8%; 85.7% - 87.9%)
20 < x ≤ 26	206 (32.5%; 28.5% - 36.3%)	889 (28.3%; 26.8% - 29.9%)	52 (16.9%; 12.2% - 21.9%)	394 (11.4%; 10.3% - 12.4%)
26 < x ≤ 33	196 (30.7%; 26.9% - 34.5%)	1,199 (38.2%; 36.5% - 39.9%)	38 (11.8%; 8.0% - 16.0%)	161 (4.6%; 3.9% - 5.4%)
33 < x	40 (5.1%; 3.3% - 7.0%)	279 (8.9%; 7.9% - 9.9%)	99 (35.4%; 29.5% - 41.8%)	679 (19.6%; 18.2% - 20.9%)
Ward type:				
Not recorded	35 (4.4%; 2.8% - 6.2%)	66 (2.1%; 1.6% - 2.6%)	79 (27.4%; 21.9% - 33.3%)	1,008 (29.1%; 27.5% - 30.6%)
Bays	468 (79.0%; 75.7% - 82.4%)	2,734 (87.1%; 85.9% - 88.2%)	61 (20.7%; 15.6% - 25.7%)	1,324 (38.2%; 36.5% - 39.8%)
Nightingale	112 (16.5%; 13.6% - 19.7%)	340 (10.8%; 9.7% - 11.9%)	18 (4.6%; 2.1% - 7.6%)	297 (8.6%; 7.6% - 9.5%)
Proportion of single rooms	0.19 (±0.17)	0.18 (±0.18)	0.18 (±0.16)	0.22 (±0.26)
Admissions	4.34 (±19.18)	4.12 (±7.27)	4.23 (±18.36)	5.39 (±8.10)
Discharges	3.47 (±3.45)	3.37 (±3.35)	3.41 (±3.39)	4.10 (±4.05)
Transfers in	1.29 (±2.12)	1.10 (±1.35)	1.26 (±1.98)	1.23 (±2.56)
Transfers out	1.55 (±4.30)	1.50 (±5.16)	1.51 (±4.33)	1.98 (±5.92)
Proportion of senior nurses	0.05 (±0.02)	0.05 (±0.03)	0.04 (±0.02)	0.05 (±0.03)
Proportion of charge nurses	0.07 (±0.04)	0.07 (±0.06)	0.07 (±0.04)	0.09 (±0.07)
Proportion of staff nurses	0.55 (±0.11)	0.55 (±0.12)	0.54 (±0.11)	0.60 (±0.12)
Total WTE per bed	0.98 (±0.46)	1.00 (±0.38)	0.98 (±0.45)	1.08 (±0.44)

Table 7 Difference between included and excluded wards for patient falls and drug errors models[7].

[7] *percent values refer to the column percent for each variable e.g. ward type.

4.1.3 Associations with outcomes

Several regression models were constructed to investigate associations between the available outcomes and staffing levels and skill mix. Henceforth, where the p-value of a regression estimate is less than 5% ($p<0.05$) the text will claim an “association” between that variable and the outcome .

Patient falls

Table 8 reports the results of a linear regression analysis. It demonstrates that falls were significantly less likely to occur on surgical or mixed (medical and surgical) wards More falls occurred on wards between 20 and 33 beds in size compared to wards between 12 and 20 beds. Fewer falls occurred on wards with a Nightingale ward layout compared to wards set out in bays. Higher admission rates were associated with more falls, and higher discharge rates and transfers out with lower rates.

		Falls per 1000 adjusted bed days		
Adj. R-squared = 0.22		**Estimate (95% confidence interval)**	**t value**	**p value**
Ward specialty:				
	Medical	0.00 (reference)		
	Mixed	-5.06 (-7.58 to -2.54)	-4.014	<0.001
	Surgical	-9.17 (-10.20 to -8.15)	-17.881	<0.001
Ward size (x=number of beds):				
	x < 12	-1.92 (-4.22 to 0.38)	-1.674	0.094
	12 < x ≤ 20	0.00 (reference)		
	20 < x ≤ 26	1.55 (0.25 to 2.85)	2.386	0.017
	26 < x ≤ 33	1.38 (0.07 to 2.70)	2.103	0.036
	33 < x	0.19 (-1.70 to 2.08)	0.201	0.840
Ward layout:				
	Bays	0.00 (reference)		
	Nightingale	-3.77 (-5.21 to -2.32)	-5.226	<0.001
	Not recorded	-6.12 (-9.33 to -2.91)	-3.81	<0.001
Proportion of single rooms		1.39 (-1.73 to 4.51)	0.892	0.372
Proportion of days with high dependency patients		0.30 (-1.18 to 1.77)	0.404	0.686
Admissions per bed		0.07 (0.04 to 0.10)	5.143	<0.001
Discharges per bed		-0.53 (-0.69 to -0.38)	-6.823	<0.001
Transfers in per bed		0.00 (-0.23 to 0.22)	-0.015	0.988
Transfers out per bed		-0.17 (-0.28 to -0.06)	-2.99	0.003
Total WTE per bed		0.75 (-0.54 to 2.05)	1.163	0.245
Proportion of senior nurses		-23.44 (-42.72 to -4.16)	-2.431	0.015
Proportion of charge nurses		-26.67 (-36.40 to -16.94)	-5.482	<0.001
Proportion of staff nurses		-23.65 (-27.96 to -19.34)	-10.981	<0.001

Table 8 Linear regression analysis of ward and staffing factors associated with number of patient falls per 1000 bed days.

Fewer falls were also associated with a higher proportion of staff nurses and charge nurses (relative to the total combined nursing and support staff numbers). There was no association between falls and the total ward staffing level.

Staffing levels were also divided into quintiles to look for non-linear associations (numerical modelling results in Appendix 1). No consistent trend was identified between a higher proportion of senior nurses and patient falls (Figure 6), as the confidence intervals overlap. A skill mix containing more than 10% charge nurses was associated with fewer falls with the exception of the highest charge nurse staffing levels at above 20% (Figure 7). The figures presented here include 95% confidence intervals. Black squares represent the median and reference category.

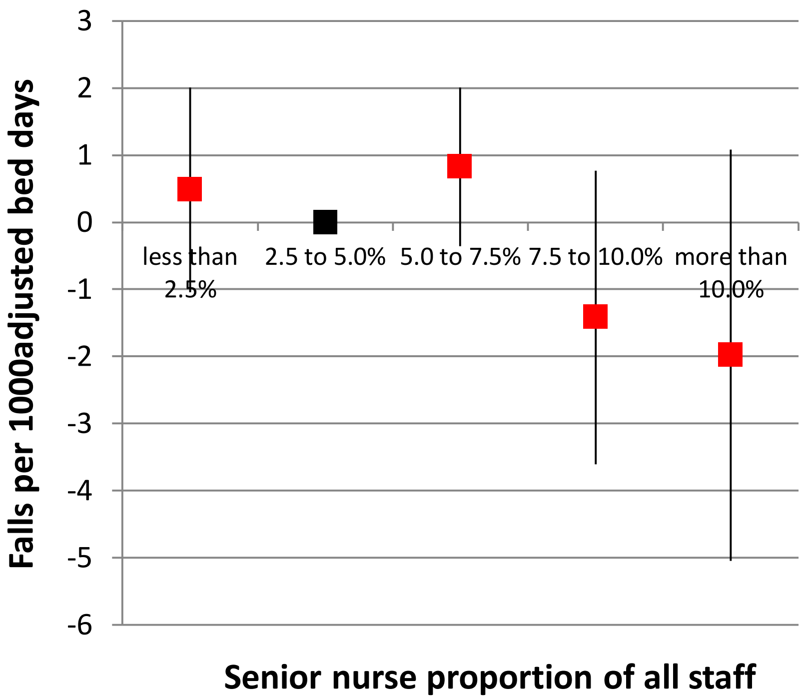


Figure 6 The relationship between the proportion of senior nurses per ward and number of patient falls (reference category 2.5-5.0% senior nurses)

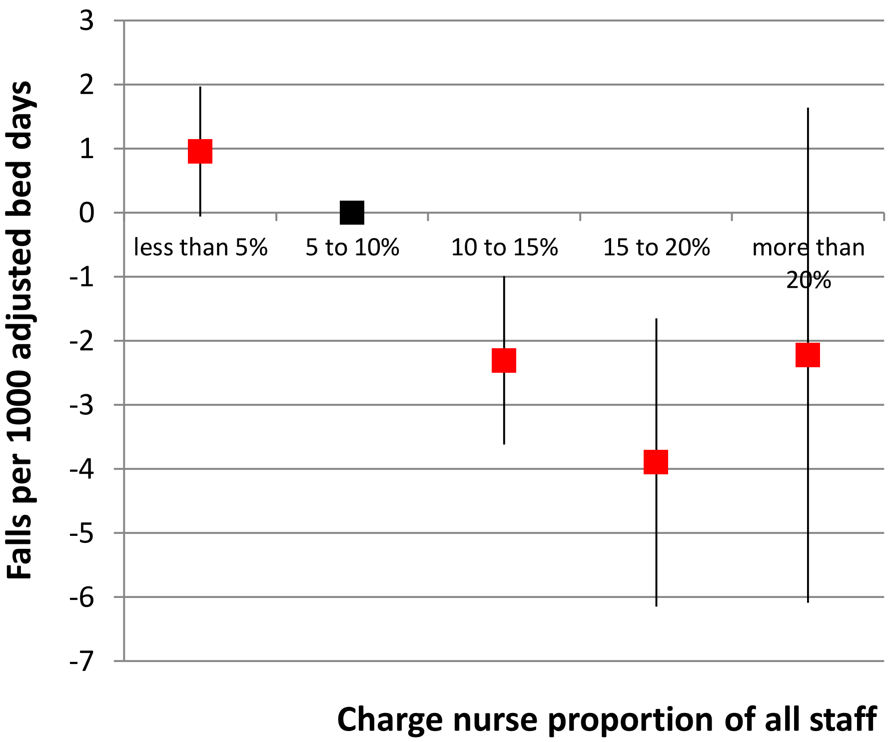


Figure 7 The relationship between the proportion of charge nurses and the number of patient falls (reference category 5-10% charge nurses).

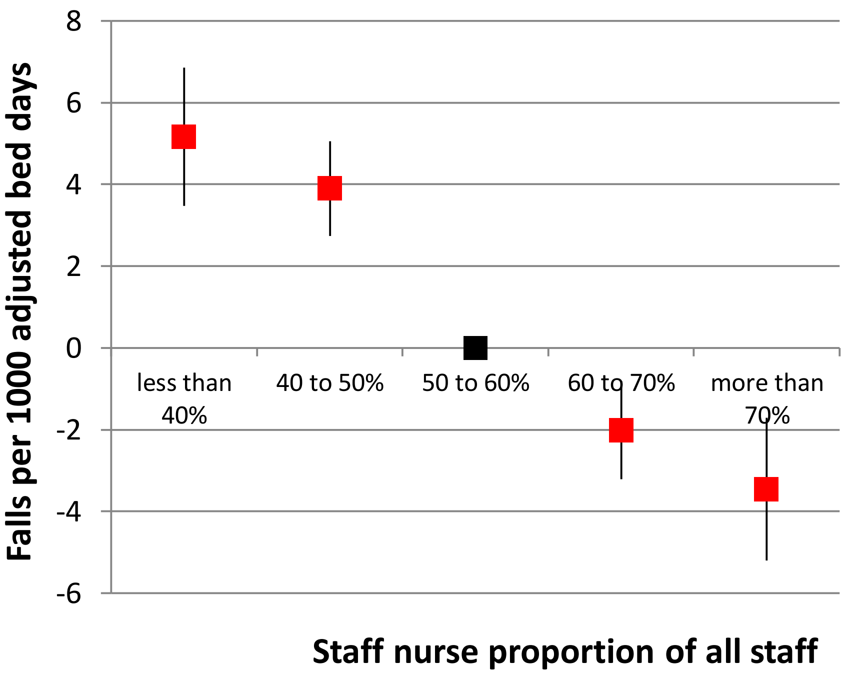


Figure 8 The relationship between the proportion of staff nurses and the number of patient falls (reference category 50-60%).

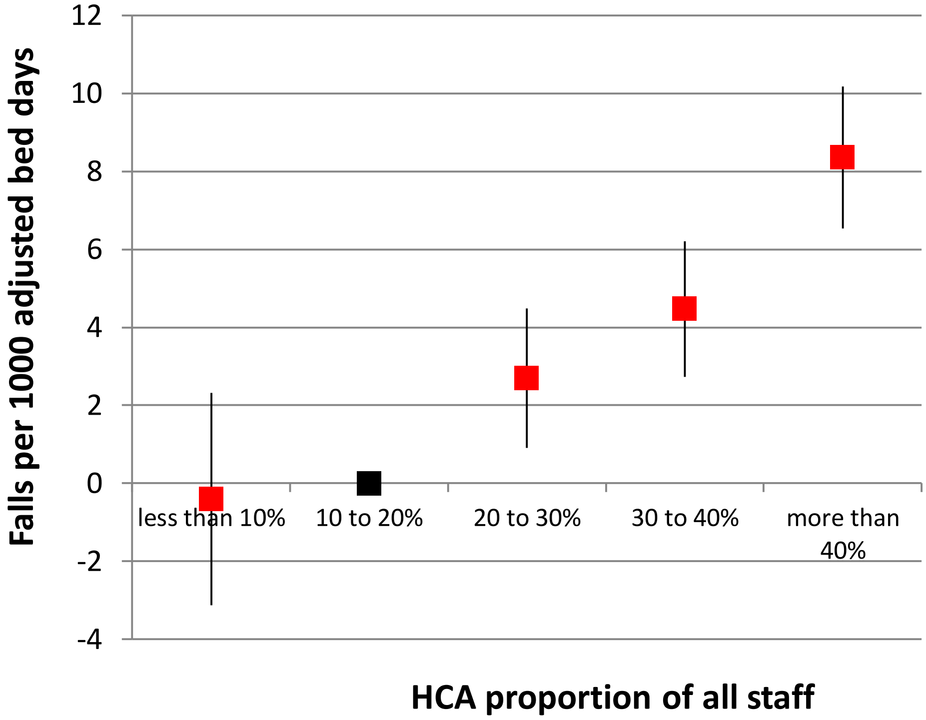


Figure 9 The relationship between the proportion of HCA and the number of patient falls (reference category 10-20%).

A skill mix containing a proportion of staff nurses greater than 60% was associated with fewer falls and less than 50% with more falls (compared with the median staff nurse category; 50 to 60% of all nursing care staff) (Figure 8). Additional modelling (replacing the nursing staffing variables with a single variable for the proportion of HCAs: Appendix 1) demonstrates that a skill mix with an HCA proportion greater than 20% of total nursing and support staff was associated with a higher number of patient falls (Figure 9).

Drug errors

Table 9 presents the results a regression analysis for drug errors. Surgical and mixed wards also had fewer drug errors than medical wards. Small wards (fewer than 12 beds) and Nightingale wards had significantly lower rates of drug errors. A higher rate of patient transfers onto a ward was associated with more drug errors, but admission rate had no association. A higher proportion of both charge nurses and staff nurses was associated with a higher rate of drug errors.

	Drug errors per 1000 adjusted bed days		
Adj. R-squared = 0.06	**Estimate (95% confidence interval)**	**t value**	**p value**
Ward specialty:			
Medical	0.00 (reference)		
Mixed	-0.78 (-1.65 to 0.09)	-1.802	0.072
Surgical	-0.54 (-0.89 to -0.18)	-3.045	0.002
Ward size (x=number of beds):			
x < 12	-1.55 (-2.34 to -0.77)	-3.948	<0.001
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	0.01 (-0.44 to 0.45)	0.04	0.968
26 < x ≤ 33	0.25 (-0.20 to 0.70)	1.107	0.268
33 < x	0.09 (-0.56 to 0.74)	0.276	0.782
Ward type:			
Bays	0.00 (reference)		
Nightingale	-0.35 (-0.84 to 0.15)	-1.408	0.159
Not recorded	0.57 (-0.53 to 1.67)	1.028	0.304
Proportion of single rooms	0.54 (-0.52 to 1.60)	1.025	0.306
Proportion of days with high dependency patients	0.32 (-0.18 to 0.83)	1.285	0.199
Admissions per bed	0.01 (0.00 to 0.01)	1.224	0.221
Discharges per bed	-0.04 (-0.09 to 0.02)	-1.407	0.160
Transfers in per bed	0.26 (0.19 to 0.34)	6.829	<0.001
Transfers out per bed	-0.02 (-0.06 to 0.02)	-0.917	0.359
Total WTE per bed	1.98 (1.58 to 2.39)	9.804	<0.001
Proportion of senior nurses	0.62 (-6.00 to 7.24)	0.186	0.852
Proportion of charge nurses	6.94 (3.60 to 10.27)	4.161	<0.001
Proportion of staff nurses	4.00 (2.52 to 5.48)	5.41	<0.001

Table 9 Linear regression analysis of ward and staffing factors associated with number of drug errors per 1000 bed days.

Dividing staffing skill mix variables into quintiles (proportion of senior nurses, charge nurses and staff nurses; see Appendix 2) demonstrated that increasing charge nurse proportion appears to be associated with more drug errors (Figure 10). A similar association was identified with the proportion of staff nurses (Figure 11). However, in both instances this is not true of all quintiles of these nursing groups because the confidence intervals for some of the quintiles are quite wide and therefore overlap. However, the overall trend is ambiguous from the figures.

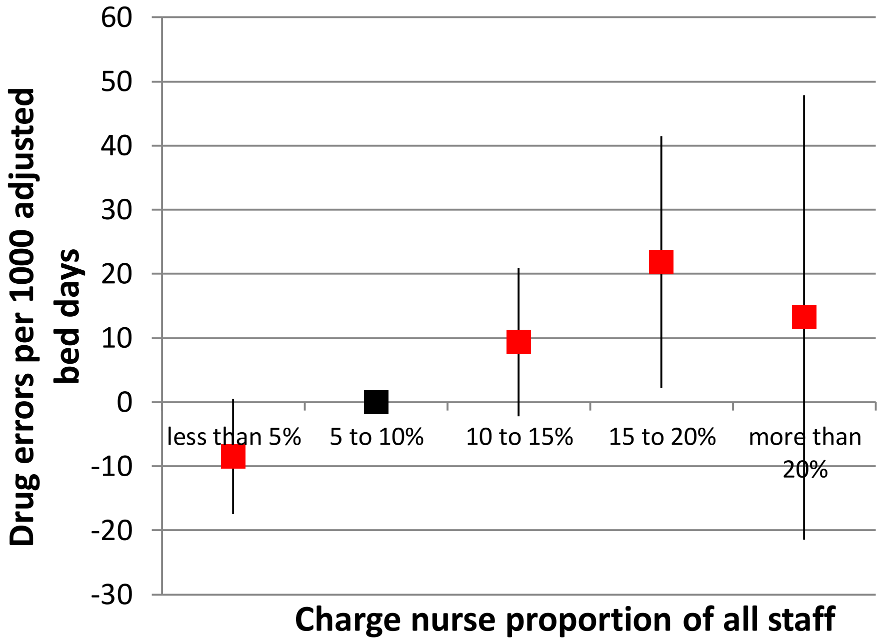


Figure 10 The relationship between the proportion of staff nurses and the number of patient falls (reference category 50-60%).

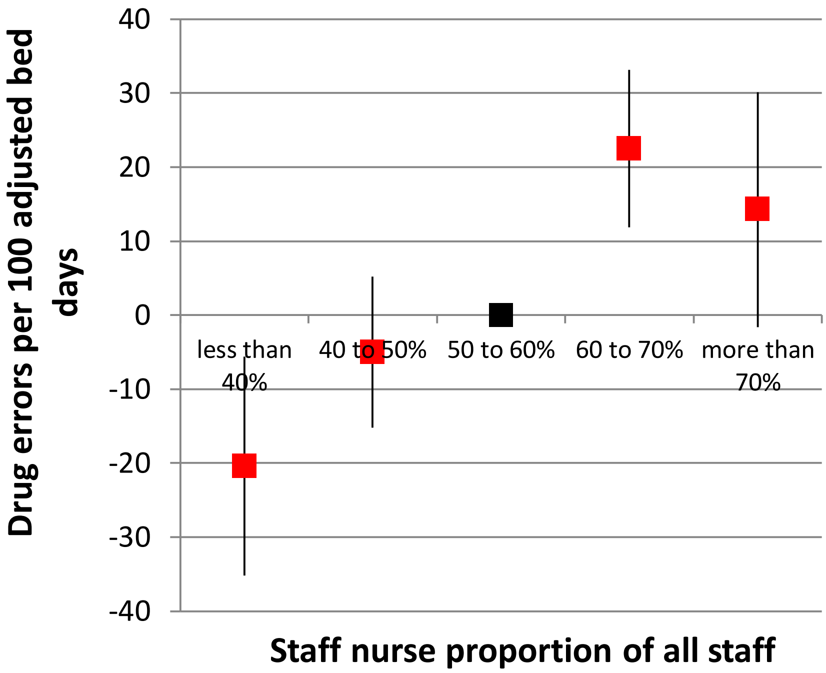


Figure 11 The relationship between the proportion of staff nurses and the number of patient falls (reference category 50-60%).

A multilevel analysis nesting wards within hospitals or within trusts was not possible due to the low number of wards in each category. Excluding the small number of wards with outlying high values of WTE (shown in

Figure 3) did not have any impact on which variables were statistically significant in the models presented. These data were therefore included in the results presented.

This linear regression analysis has some limitations when applied to the count-type outcomes of falls and drug errors. Linear regression assumes, amongst other things, that there is a continuous outcome variable (not a discrete/whole number count variable) and a normal distribution of the outcome values. We therefore applied additional model types to the same data to test how robust the relationships demonstrated here were under different modelling assumptions. These models are presented and discussed in Appendix 3. In summary, the associations found here for relationships with falls were not consistent across the model types but there were some consistencies for associations with drug errors.

The additional model types were inconsistent regarding the relationships between ward size and falls. They found no association between staffing levels or skill mix and patient falls and were inconsistent regarding the relationship with overall staffing levels and falls. These reported associations with numbers of patient falls should therefore be treated with caution.

Drug errors were found to be consistently lower on surgical and mixed wards in all model types. A higher proportion of transfers in was consistently associated with more drug errors. The relationship between higher a proportion of charge and staff nurses and more drug errors was also consistent across all the model types. These robust relationships are likely to reflect real associations. The models were inconsistent regarding the relationship with overall staffing levels and drug errors. The literature is also inconsistent in this respect, with roughly a third of the nine studies reviewed that included this variable finding a negative relationship, and an equal number finding a positive relationship.

4.2 UKND dataset

This dataset comprises 1,003 data samples from 799 wards collected between 1985 and 2014. Data were collected across all months of the year. For this analysis ward samples were excluded if data were incomplete (N=375) or the ward was not an adult medical, surgical, mixed medical and surgical, or gynaecology ward (paediatric, critical care, rehabilitation, etc wards excluded; N=171). Wards were then excluded if they were not in England (N=64). The remaining ward samples (N=393) were measured on 357 unique wards. The mean number of samples per ward was 1.10 (range 1-4). These ward samples comprised 266 medical wards, 111 surgical wards, 4 mixed medical and surgical wards, and 12 gynaecology wards.

4.2.1 Staffing trends

Information on number of patients by level of dependency, total ward occupancy, staffing levels by staff grade (WTE per occupied bed), and ward type were available for all the included wards. The total bed occupancy for the sample was 9,924 with a mean of 21.7 (SD ±7.5) occupied beds per ward.

The average WTE per occupied bed was 1.33 (SD ±0.68). Annual average WTE per occupied bed was observed to have increased over time (Figure 12). The ward total WTE was strongly correlated with the ward total adjusted bed days. WTE per occupied bed was again dependant on ward size with a substantial rise in WTE per occupied bed in wards with fewer than 12 occupied beds (

Figure 13).

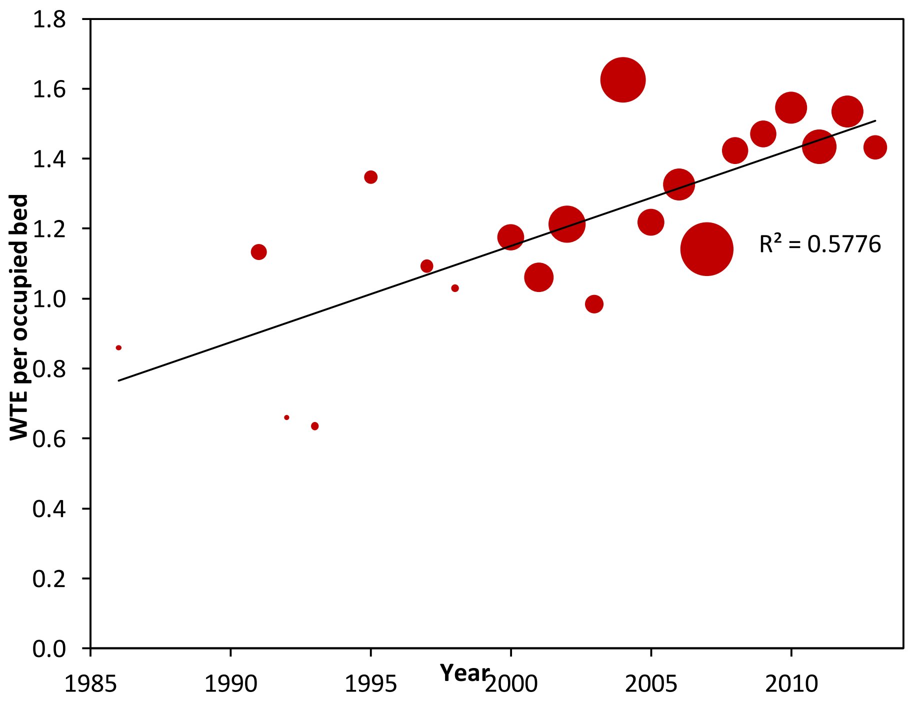


Figure 12 The increase in WTE per occupied bed over time. Data points are the mean values for all wards observed in each year. The area of each point represents the number of wards sampled in that year.

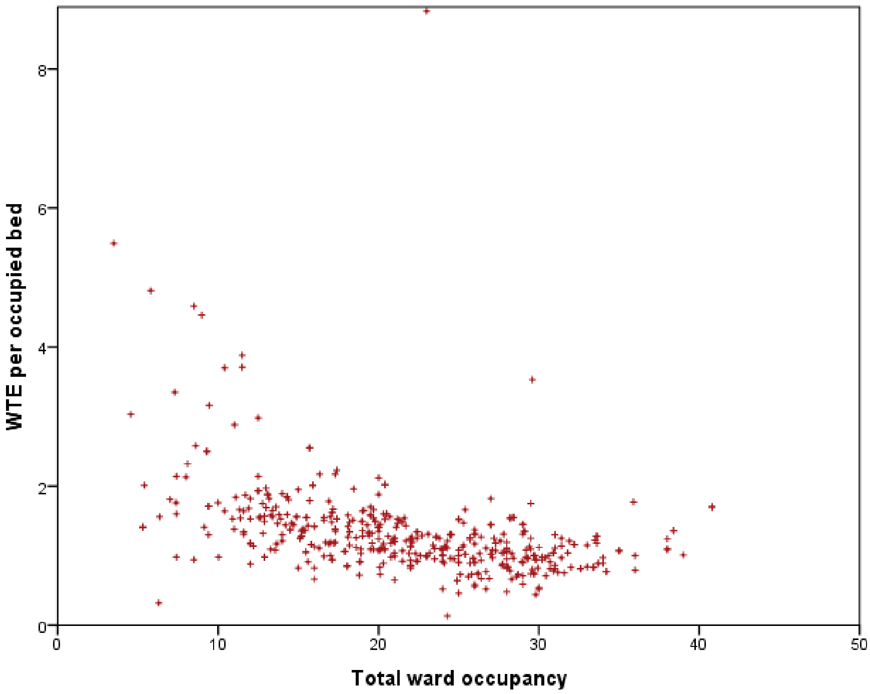


Figure 13 The relationship between staffing levels and total ward occupancy

4.2.2 Patient dependency

Patients had been categorised into four dependency categories; low, low-medium, medium, and high dependency. Patient dependency data were available for all 393 ward samples included. Patients with low-medium and medium dependencies were the most abundant (Table 10).

Patient dependency	Mean number per ward (SEM)	Mean proportion of ward (SEM)
Low	2.58 (±2.20)	8.63% (±11.19)
Low-medium	9.43 (±5.42)	42.36% (±19.68)
Medium	7.94 (±4.16)	37.02% (±17.66)
High	2.86 (±2.95)	11.99% (±13.62)

Table 10 Numbers and proportion of patients by dependency group

4.2.3 Patient and ward outcome data

A number of composite nursing quality outcomes were reported for most of the included ward samples at ward level. These comprise; how well patients' needs were assessed, how well care addressed patients' needs, extent to which the care plan was implemented, how well patients' response to care and progress was recorded, ward fabric and resources, and overall quality. Each score was reported as a percentage value (Table 11). The actual data on the component factors for each score was not available but the definitions used in the survey instrument were made available.

Outcome measure	N (%)	Mean score (SEM)	Range
How well patients' needs were assessed (%)	391 (99.5)	70.8 (±16.9)	22 - 100
How well care addressed patients' needs (%)	391 (99.5)	60.8 (±23.0)	0 - 100
Extent to which care plan was implemented (%)	391 (99.5)	88.6 (±8.7)	45 - 100
How well patients' response to care and progress was recorded (%)	390 (99.2)	67.3 (±20.3)	15 - 100
Ward fabric and resources score (%)	378 (96.2)	83.4 (±9.8)	53 - 100
Overall nursing quality score (%)	391 (99.5)	70.8 (±16.9)	22 - 100

Table 11 Nursing quality outcomes of 457 samples measured on 390 wards.

4.2.4 Correlations with outcomes

Linear regression analysis reported in Table 12, demonstrated an association with lower scores for assessments of patients' needs and overall nursing quality in district general when compared to teaching hospitals. Community hospitals were associated with lower scores for assessment of patients' needs and how well patients' progress was documented.

Higher ward occupancy (more than 33 patients) was associated with lower scores in all quality measures. Ward occupancy above 26 patients was associated with lower scores for how well care addressed patients' needs, the extent to which the care plan was implemented, and how well patients' progress was documented. Whilst consistent across the quality measures, the association between ward occupancy and quality measures was small and there was a high degree of variability in all the quality measures (Figure 15 demonstrates the high degree of variability in how well care addressed patients' needs as an example). Patient dependency had little impact on the quality of care with only a small association between higher numbers of low-medium dependency patient and lower scores for how well care addressed patients' needs, the extent to which the care plans were implemented, and overall nursing quality.

There was no correlation between total staffing levels per occupied bed and assessment of patients' needs, how well care addressed patients' needs, how well the care plan was implemented, or overall nursing score. Recording of patients' response to care and progress was the only outcome measure to demonstrate a significant improvement with higher staffing levels. When the overall staffing level was categorised into quintiles (Appendix 4) this effect was found to be non-linear (Figure 14).

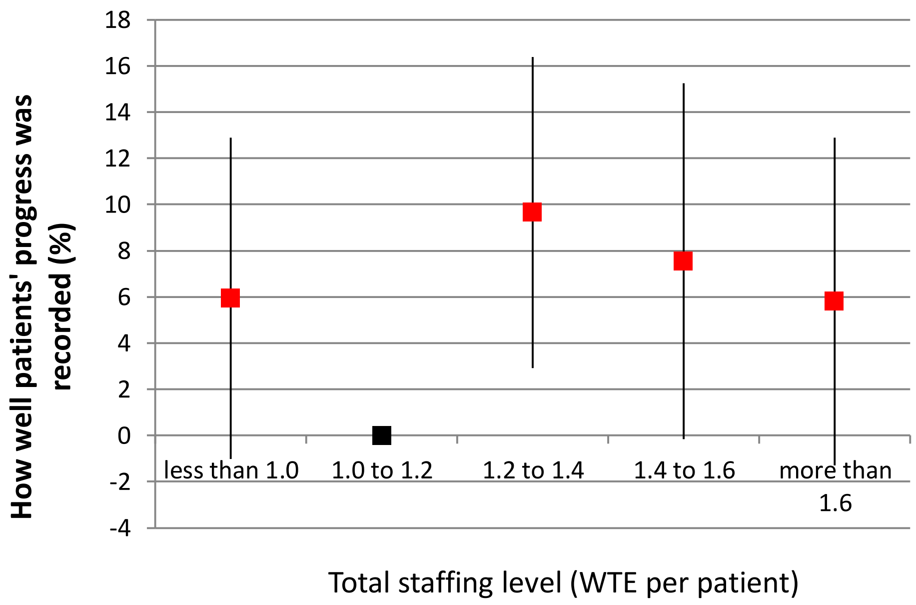


Figure 14 The effect of staffing levels on the quality of recording and documentation. The reference category is shown in black.

The impact of skill mix on quality was varied. A higher proportion of support staff was associated with better assessment of patients' needs, improved care in addressing needs, and better overall quality. A higher proportion of bank staff and ward sisters were associated with lower scores for assessment of patients' needs.

		How well patients' needs were assessed		How well care addressed patients' needs		Extent to which care plan was implemented		How well patients' progress was recorded		Overall nursing quality score	
	N (%) or mean (SD)	**Estimate (95% Confidence limits)**	**P value**	**Estimate (95% Confidence limits)**	**P value**	**Estimate (95% Confidence limits)**	**P value**	**Estimate (95% Confidence limits)**	**P value**	**Estimate (95% Confidence limits)**	**P value**
Adj. R-Squared		0.08		0.11		0.07		0.08		0.12	
Hospital type:											
Teaching hospital	216 (47.3%)	1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)	
Community hospital	121 (26.5%)	2.87 (-3.40 to 9.14)	0.361	-11.32 (-19.71 to -2.93)	0.007	0.88 (-2.36 to 4.11)	0.589	-8.57 (-16.13 to -1.01)	0.024	-1.01 (-5.08 to 3.05)	0.618
District general hospital	116 (25.4%)	-1.00 (-5.63 to 3.63)	0.666	-5.43 (-11.63 to 0.77)	0.081	0.53 (-1.86 to 2.92)	0.658	-0.85 (-6.43 to 4.72)	0.760	-1.57 (-4.57 to 1.43)	0.296
Ward layout:											
Bays	283 (61.9%)	1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)	
Racetrack	24 (5.3%)	-0.32 (-7.63 to 7.00)	0.931	4.02 (-5.77 to 13.81)	0.412	0.80 (-2.98 to 4.58)	0.672	-4.16 (-12.95 to 4.64)	0.345	0.18 (-4.57 to 4.94)	0.939
Single rooms	25 (5.5%)	5.96 (-1.61 to 13.54)	0.116	2.76 (-7.38 to 12.90)	0.587	1.31 (-2.61 to 5.22)	0.505	9.21 (0.09 to 18.32)	0.044	3.06 (-1.86 to 7.99)	0.214
Nightingale	10 (2.2%)	8.29 (-3.75 to 20.33)	0.169	-5.55 (-21.66 to 10.56)	0.491	-3.12 (-9.34 to 3.09)	0.316	-4.95 (-19.43 to 9.52)	0.494	0.50 (-7.32 to 8.32)	0.899
Split site	14 (3.1%)	3.54 (-5.73 to 12.80)	0.446	-11.18 (-23.58 to 1.22)	0.072	-8.91 (-13.70 to -4.13)	<0.001	-1.16 (-12.31 to 9.99)	0.836	-6.71 (-12.73 to -0.69)	0.026
Other	101 (22.1%)	-3.91 (-8.26 to 0.45)	0.074	-11.37 (-17.21 to -5.54)	<0.001	-4.23 (-6.48 to -1.98)	<0.001	-8.40 (-13.64 to -3.16)	0.001	-6.96 (-9.78 to -4.13)	<0.001
Ward occupancy (x=number of patients):											
x < 12	125 (27.4%)	-0.02 (-6.86 to 6.82)	0.996	3.29 (-5.87 to 12.45)	0.473	-2.62 (-6.15 to 0.92)	0.140	-2.76 (-10.99 to 5.47)	0.503	-1.14 (-5.58 to 3.31)	0.609
12 < x ≤ 20	44 (9.6%)	0.19 (-4.19 to 4.58)	0.929	1.00 (-4.87 to 6.87)	0.733	-0.03 (-2.29 to 2.23)	0.979	-0.51 (-5.78 to 4.76)	0.846	-0.04 (-2.89 to 2.81)	0.975
20 < x ≤ 26	139 (30.4%)	1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)		1.00 (reference)	
26 < x ≤ 33	119 (26.0%)	-2.03 (-7.12 to 3.06)	0.425	-8.59 (-15.40 to -1.78)	0.012	-2.44 (-5.07 to 0.19)	0.064	-8.82 (-14.94 to -2.70)	0.004	-3.09 (-6.38 to 0.19)	0.061
33 < x	30 (6.6%)	-8.78 (-16.41 to -1.15)	0.022	-15.78 (-25.99 to -5.56)	0.002	-4.81 (-8.75 to -0.87)	0.015	-11.13 (-20.45 to -1.81)	0.017	-7.72 (-12.68 to -2.77)	0.002
Patient dependency (% of all patients):											
Low dependency	8.63 (±11.19)	-0.04 (-0.26 to 0.19)	0.735	-0.09 (-0.39 to 0.20)	0.528	-0.03 (-0.15 to 0.08)	0.545	-0.27 (-0.54 to 0.00)	0.046	-0.04 (-0.18 to 0.11)	0.614
Low-medium dependency	42.36 (±19.68)	0.00 (-0.14 to 0.14)	0.990	-0.17 (-0.35 to 0.02)	0.075	-0.06 (-0.13 to 0.02)	0.117	-0.07 (-0.24 to 0.10)	0.429	-0.06 (-0.16 to 0.03)	0.160
Medium dependency	42.36 (±17.66)	-0.01 (-0.20 to 0.17)	0.906	-0.04 (-0.29 to 0.21)	0.728	-0.03 (-0.13 to 0.06)	0.507	-0.05 (-0.27 to 0.176)	0.674	-0.03 (-0.15 to 0.09)	0.669
Staff level (% of all staff):											
Ward sisters	8.63 (±11.19)	-0.23 (-0.45 to -0.01)	0.039	-0.23 (-0.53 to 0.07)	0.121	0.04 (-0.07 to 0.16)	0.443	-0.13 (-0.40 to 0.13)	0.317	-0.05 (-0.20 to 0.09)	0.472
Staff nurses	42.36 (±19.68)	0.07 (-0.04 to 0.18)	0.223	-0.18 (-0.34 to -0.03)	0.015	0.03 (-0.03 to 0.09)	0.288	0.08 (-0.05 to 0.22)	0.231	0.00 (-0.07 to 0.07)	0.968
Support staff	37.02 (±17.66)	0.10 (0.02 to 0.17)	0.014	0.19 (0.09 to 0.30)	<0.001	0.00 (-0.04 to 0.04)	0.913	0.06 (-0.03 to 0.15)	0.207	0.06 (0.01 to 0.11)	0.021
Bank staff	11.99 (±13.62)	-0.31 (-0.55 to -0.07)	0.010	-0.18 (-0.50 to 0.14)	0.257	-0.09 (-0.21 to 0.04)	0.166	-0.16 (-0.45 to 0.12)	0.255	-0.14 (-0.30 to 0.01)	0.067
Total staffing level (WTE per occupied bed)	1.41 (±0.71)	1.63 (-1.20 to 4.46)	0.249	-1.15 (-4.94 to 2.64)	0.544	0.10 (-1.36 to 1.56)	0.891	2.27 (0.57 to 3.97)	0.018	1.02 (-0.82 to 2.85)	0.267

Table 12 Associations between hospital, ward, patient and staff factors and nursing quality outcome measures. Overall quality score is summation of the nursing quality indicators shown and ward fabric and resources score.

The observed associations with reported quality scores should be treated with caution as there was large variance in all of the outcome measures. As an example, Figure 15 shows this effect for scores of how well the care addressed the patients' needs plotted against ward occupancy (strong association demonstrated by the linear regression model). Whilst a general trend towards lower quality is associated with increasing ward occupancy there is a high degree of variability in the quality score. 95% confidence intervals for the trend-line are shown in red.

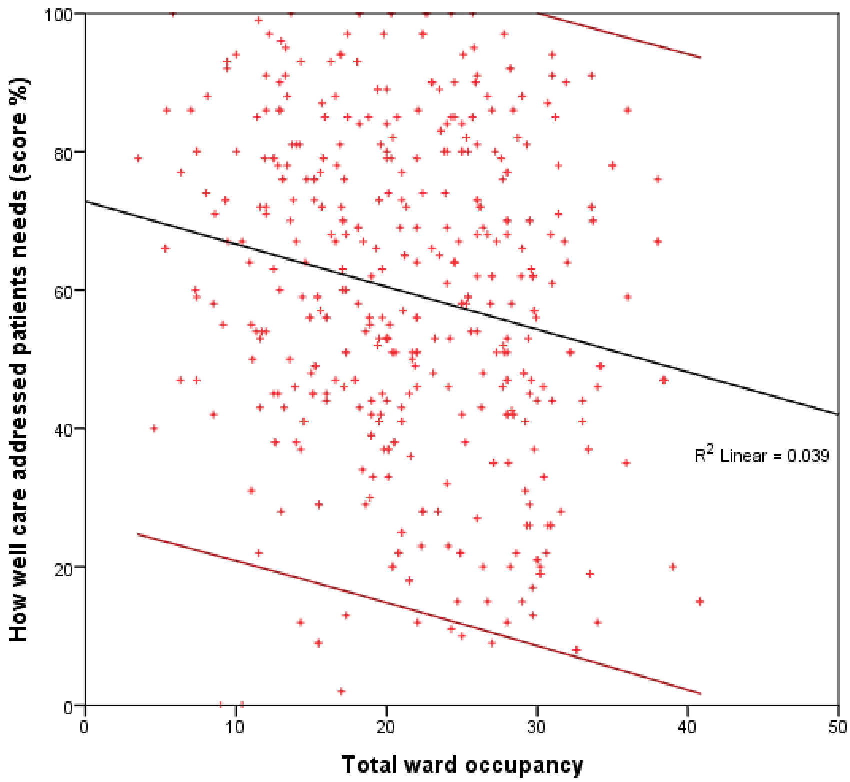


Figure 15 Total ward occupancy and the quality score for how well care addressed patients' needs for all included wards.

4.3 NFT dataset

[Additional information to be added to this section following the consultation period of the draft guideline]

4.4 Cost-Effectiveness Analysis

4.4.1 Economic Model Parameters

The results of the systematic literature review [4] identified four economic studies of nurse staffing and nurse sensitive outcomes. To reiterate, only one study reported findings for falls and medication errors but the setting of the study (US hospital level analysis) was determined to be unhelpful for this analysis. This study therefore used alternative sources of NHS ward level data to determine the effect of nurse staffing and skill mix on nurse sensitive outcomes. The results of this analysis can be summarised as:

- UKND provides no data for the outcomes of interest (falls and medication errors) and is a potentially biased, small subset of English wards. The AC dataset, whilst a decade old[8], is therefore preferred as it provides a more representative sample and reports data on falls and drug errors.
- No relationship was found between staffing numbers per adjusted bed (either by individual group or total staff) and the outcomes, despite using numerous alternative model specifications and regression models. Therefore, the intervention to increase staffing numbers will not be considered further as there is no evidence of its effectiveness, ceteris paribus, in these data. This is inconsistent with the literature review.
- There was consistent evidence of the relationship between a richer skill mix (i.e. a higher proportion of registered nurses per bed) for falls, but weaker and more inconsistent evidence of the effect of skill mix on drug errors. We will therefore use the results of the ward level analysis for falls in the CEA, but populate the economic model with the effectiveness results from a US study by Cho et al [7].
- From the quintile modelling reported in appendix 1 and 2, and visualised for example in Figure 9 and Figure 11 it is clear that a sensible skill-mix intervention would be a 10 percentage point increase (reduction) in the proportion of registered nurses (HCAs) from a ward average of 64% (36%) to 74% (26%).

Table 13 presents the parameter values that will be used as the base case for the analysis alongside the upper and lower values, which are derived from the descriptive and inferential statistical analysis reported in Section 4. The lower and upper values are set to twice the standard deviation or standard error of the base case, except for the costs of employing staff. For staff costs, the lower value is set to the bottom of the relevant Agenda for Change band (e.g. Band 2 or Band 5), and for the upper value the top of the relevant Agenda for Change band is used, multiplied by the London Weighting factor used by PSSRU of 1.19. In both cases the employer's on-costs (14% pension contributions and 13.8% national insurance contributions above £146 per week) are included to make this comparable to the national mean wages reported by PSSRU. The sensitivity of the results

[8] It is worth noting that the existing economic studies are largely based on relationships established in the late 1990's and early 2000's, with the exception of Twigg et al.

to the magnitude of the skill-mix intervention (fixed at 10 percentage points in the base case) is tested by varying this from 5 (lower value) to 15 (upper value) percentage points.

Parameter	Base Case	Lower Value	Upper Value	Source
Exposure	9,125	3,800	14,500	Sec. 4.1.1
Fall Effectiveness[9]	-0.15%	-0.12%	-0.18%	Table 8
Drug Effectiveness	-0.06%	+0.09%	-0.11%	Cho et al.
WTE Nurses	0.632	0.00	1.28	Sec. 4.1.1
WTE HCAs	0.359	0.143	0.559	Sec. 4.1.1
Cost of Nurses	£31,867	£26,401	£41,385	Table 8
Cost of HCAs	£19,661	£17,220	£25,453	Sec. 4.1.1
Ratio N/H	0.64	0	0.7	Table 2
Intervention	10	5	15	Sec. 4.1

Table 13 Parameters used in Sensitivity Analysis

4.4.2 Cost-Effectiveness and Skill Mix

The incremental cost of implementing the base case intervention is the same irrespective of the outcome under consideration. The ward level incremental cost is £30,907 as Table 14 illustrates.

	WTE	Total (WTE x 24.9[10])	Unit Cost	Total Cost
Before Intervention				
Nurses	0.63	15.79	£31,867.00	£503,222.71
HCAs	0.36	8.97	£19,661.00	£176,360.79
All Staff	0.99	24.76	£51,528.00	£679,583.50
Post Intervention				
Nurses	0.73	18.32	£31,867.00	£583,913.52
HCAs	0.26	6.44	£19,661.00	£126,576.94
All Staff	0.99	24.76	£51,528.00	£710,490.45

[9] The regression coefficients for Table 8 which report the change in the number of falls per 1,000 adjusted bed days from a 1 unit change in the staff proportion (e.g. moving from 0.25-0.26) have been converted to elasticities to make them comparable with the drug effectiveness reported in Cho et al. For clarity, from the regression coefficients a one percentage point increase in the proportion of nurses would generate a predicted 0.24, 0.19 and 0.28 reduction in the number of falls per 1,000 bed days for the base case, lower and upper values respectively.

[10] The exposure (the total number of adjusted bed days) is divided by 365 days to obtain the adjusted ward size for calculating the staffing costs.

Incremental Change	0.00	0.00	£0.00	£30,906.96

Table 14 Incremental Cost of Skill-Mix Intervention

The incremental benefit of increasing staffing by the base case level (of 10 percentage points) is estimated to be a reduction of 22 falls and less than half a drug error per ward per annum, as

	Rate per 1,000 Bed Days	Total Number
Before Intervention		
Falls	14.62	133.33
Drug Errors	2.84	25.90
Post Intervention		
Falls	12.22	111.45
Drug Errors	2.81	25.66
Incremental Change		
Falls	2.40	21.89
Drug Errors	0.03	0.24

Table 15 illustrates.

	Rate per 1,000 Bed Days	Total Number
Before Intervention		
Falls	14.62	133.33
Drug Errors	2.84	25.90
Post Intervention		
Falls	12.22	111.45
Drug Errors	2.81	25.66
Incremental Change		
Falls	2.40	21.89
Drug Errors	0.03	0.24

Table 15 Incremental Benefit of Skill Mix Intervention

The Incremental Cost Effectiveness Ratios are therefore £1,412 per fall averted and £128,779 per drug error avoided, this is in comparison to the current practice i.e. the current level of staffing and skill mix.

As we cannot express this ICER in a universal 'currency' such as QALYs, it is difficult to establish whether this represents value for money: whether it generates more health benefit than another intervention it may displace. One potential way to evaluate the falls ICER is to compare them to alternative interventions to reduce falls in hospital inpatients. For example, a UK community based fall prevention strategy reported an ICER of £3,320 cost per fall averted [17], and a US secondary care intervention using a medical vigilance aid had an ICER of £3,890 ($5895) per fall avoided [21].

Another way to compare the value for money of reducing falls through altering the staffing mix is to compare the ICER with the NHS costs incurred per fall estimated earlier. At £79 per fall this does not compare well, but this cost only includes the direct and immediate cost of providing care following a fall. It excludes increased length of stay, readmissions, rehabilitation, or other lasting health impacts on the patient. As a comparison, note that according to Dall [8] increasing staffing by 1 RN per annum would generate US$46,000 (in 2005) of avoided medical costs e.g. reduced falls, reduced medication errors, etc. Adjusted for 2013 prices (US$ 53,100[11]) and converted to £27,600[12]. Therefore, a 10% FTE increase in staffing as per the intervention considered in this study would generate £2,700 of benefit under these US figures. However that benefit is arrived at by using extremely high costs of medical care. For example the average cost of a fall in the US is $7,718 (in 2005 US dollars) or approximately £5,791 (in 2013 GBP) compared to an estimated NHS cost of £79.

4.4.3 Sensitivity Analysis

A one-way sensitivity analysis was performed whereby each of the parameters in Table 13 are varied from the base case to their upper and lower values. With the exception of employment costs, these upper and lower bounds are the base case ± twice the standard deviation or error. The sensitivity analysis demonstrates to what extent the results of the CEA are influenced by the assumptions that have been made and allow the uncertainty in the parameter values to be illustrated. Table 16 presents the results of the sensitivity analysis for the ICER for both outcomes. ICERs reported in red are positive amounts i.e. savings per unit of outcome.

The most significant variation in the results is due to staffing both in terms of the current WTE numbers of nurses and HCAs and their employment costs. For instance, in relation to the ICER for

[11] Using Consumer Price Index inflators

[12] Converted using the average exchange rate in 2013.

falls, compared to a base case estimate of £1,412 per fall averted, if the actual WTE HCA level was at its upper value (0.56 WTE compared to 0.36 WTE for the base case) the ICER rises to £5,622 per fall avoided. At the other extreme, if the actual WTE nurses level was at its upper value (1.28 WTE compared to 0.63 WTE for the base case) the ICER falls to -£936 per fall avoided. The ward would actually save money according to the sensitivity analysis.

	ICER - Falls			ICER - Drug Errors		
	Base	Lower	Upper	Base	Lower	Upper
Exposure	£1,412	£1,412	£1,412	£127,283	£127,283	£127,283
Effect	£1,412	£1,784	£1,210	£127,283	£203,653	£80,057
Nurse cost	£1,412	£780	£2,513	£127,283	£168,682	£142,484
HCA Cost	£1,412	£1,694	£742	£127,283	£366,571	£42,068
WTE Nurses	£1,412	£3,702	£936	£127,283	£800,808	£53,039
WTE HCAs	£1,412	£4,761	£5,622	£127,283	£1,029,974	£318,751
Ratio N/H	£1,412	£1,412	£1,412	£127,283	£305,480	£80,057
Intervention	£1,412	£1,443	£1,402	£127,283	£312,232	£79,467

Table 16 Sensitivity of ICER for Drug Errors and Falls to Underlying Parameter Assumptions

5 Discussion

We found staffing ratios were highest for very small wards (fewer than 12 patients), overall staffing levels per bed/patient had no correlation with most of the measured outcomes (except the quality of documentation), and the impact of skill mix on patient outcomes was varied but positive for falls. The Incremental Cost Effectiveness Ratios are therefore £1,412 per fall averted and £128,779 per drug error avoided, this is in comparison to the current practice i.e. the current level of staffing and skill mix.

5.1 Comparison of the datasets

The AC dataset does not report bed occupancy and therefore results are based on total bed numbers rather than bed occupancy. The UKND dataset reports bed occupancy but not total bed numbers. However the mean number of beds per ward from the AC dataset (24.6 ±7.5) is comparable to the mean bed occupancy per ward calculated from the UKND dataset (21.7 ±7.5). This suggests that bed occupancy is high (approximately 88%) and, in the absent of occupancy data, bed numbers may be a reasonable proxy for occupancy.

The average WTE per adjusted bed day was 0.991 (SD ±0.424) from the AC dataset. If ward occupancy is between 75-90% this equates to per occupied bed of 1.32-1.42 per occupied bed. This is consistent with the average value from the UKND dataset; WTE per occupied bed was 1.33 (SD ±0.68). Both AC and UKND datasets demonstrate that wards with fewer than 10-12 patients have substantially higher staffing to patient ratios than larger wards.

There was poor agreement in the categorisation of ward layout between the two datasets. The UKND dataset considered six categories of wary layout (excluding mixed layouts) whereas the AC dataset considered just two. In both cases Nightingale wards were associated with better outcomes for some of the outcome measures. This suggests that ward layout is an important factor in patient outcomes and a consensus on categorisation is needed before further investigation is undertaken.

Comparisons between outcome measures are difficult given the different outcomes available in the datasets. The effect sizes in both studies were small and confidence intervals were wide. This highlights the need for large scale, high quality data collection if these trends are to be investigated further.

5.2 Limitations

The outcome analyses presented here have the limitation present in all observational studies in that they do not test causal associations. We are therefore unable to conclude that alteration of staffing skill mix or any of the other predictor variables will have a beneficial (or detrimental) impact on patient outcomes. A cluster-randomised controlled trial may be required to identify causal associations and the impact of staffing changes. Many of the associations identified here are likely to represent underlying correlations; for example a higher proportion of HCAs may be required on wards with higher dependency patients; therefore an association between HCA numbers and falls is potentially spurious. Moreover, wards with higher staffing levels may have more resources available to document and record adverse outcomes (e.g. falls). Finally, there may be an endogeneity problem in that trusts with better patient outcomes may also have higher levels of staffing or richer skill mixes for another reason (e.g. high quality management) which is excluded from the models.

5.2.1 AC data limitations

This large dataset represents a subset of all wards in English hospitals and there is potential for selection bias in the wards which had complete denominator and outcome data, or even that were chosen by trusts to be returned to the Audit Commission at all. A further concern is the dataset contained self-reported information, which may lead to bias and inaccuracies as well as concerns about common understanding of the variable definitions (e.g. have all wards reported all falls irrespective of harm).

These data also only comprise a single time point. There was no data collection across different seasons (only data from April to August 2004 were collected). These data are also substantially dated, being collected 10 years ago; it is evident from the UKND dataset that there has been an increase in overall staffing levels since that time. This is supported by aggregate nurse staffing data published by the NHS Health and Social Care Information Centre.

Information about patient characteristics were limited to ward type and patient turnover variables (admissions, discharges, and transfers). Additionally the number of days the ward had high dependency patients, defined as patients requiring 1 to 1 nursing care was available but not used in the analysis. As a marker of patient dependency this variable is significantly limited. There is no distinction between patient needing almost no nursing care except perhaps six hourly observation (as is common on many surgical wards) and patient requiring a high level of nursing input not meeting the 1 to 1 nursing care criteria. Furthermore it does not account for the total number of patients with high dependency on the ward and therefore subject to bias by ward size; with larger wards more likely to have a high number of days with one or more patients with high dependency. No data on patient acuity, comorbidities, age, or gender were available. These variables are likely to be important factors in determining the number of drug errors and falls and may confound the findings, but they may also partially determine the number of staff required. Higher nursing numbers per bed was found to be correlated to a higher number of drug errors. It is unlikely that this association is causal. There are a number of plausible (but untested) explanations such as: underlying patient factors such as the level of comorbidities which are correlated with both increased nursing requirements and higher numbers of patient medications (and therefore drug errors); that wards with higher levels of staffing have the capacity or resource to record drug errors more accurately; or that hospitals with good management and high quality risk management processes have both high levels of staff and high levels of adverse outcome recording.

The available outcome measures were also limited in this dataset. No data on patient mortality, failure to rescue, pressure sore, or other important patient outcomes were available at ward level. We were therefore unable to analyse the effect of staffing on these important outcomes. Outcomes were also based on staff reporting. Staff documentation and recording was found to increase with staffing levels in the UKND observational data, an effect which may mask true correlations.

5.2.2 UKND data limitations

This dataset contains a smaller sample of wards than the AC dataset (< 400) and therefore potential for selection bias is greater. Previous literature utilising this dataset has limited information on how the wards were selected for sampling [6] but it is understood that trusts request to be audited. This data sample is spread over a considerable number of years during which time staffing trends have changed. There are only 33 observations in the last 2.5 years and only 285 since 2004. Whilst the data are skewed towards more recent observations the results may be affected by the incorporation of some of the older ward samples. One of the great strengths of this dataset is that data were collected by independent observers and is therefore not subject to the limitations of data collected from patient notes (directly or indirectly). However there is potential for the observation method to influence working patterns and hence the measured outcome variables. For example observation may encourage more vigilant patient observation or data recording.

No patient outcomes (e.g. mortality, failure to rescue, falls, pressures sores) were reported in this study and we were therefore unable to use it to confirm or refute associations identified in previous studies. For those outcomes which were recorded important predictor variables were not available. For example there was no information available about patient factors such as age or case mix. It is therefore possible that the observed association between a higher proportion of HCAs and patient falls is the result of an unknown confounding factor; wards with a higher proportion of elderly patients are more likely to require a higher proportion of HCAs to meet their care needs. These ward will also have a higher number of falls. This effect is more evident in the drug errors data; a higher proportion of nurses is not likely to cause a higher number of drug errors. It is more plausible that there are patient factors requiring a higher level of nursing care which make more drug error more likely, such as higher levels of patient co-morbidities. Again these data were absent from the dataset.

Although the data had been recorded across several shifts including night shifts and weekends we were unable to analyse staffing trends across these periods because data from individual time periods was not available. Ward data provided were an amalgamation of all sampled time periods.

For all quality indicators the UKND dataset had considerable variation and therefore any apparent correlations with hospital, patient, and staffing factors should be treated with extreme caution. In addition we were unable to identify any published literature demonstrating validity of the tools used to collect this data. To the best of our knowledge there has been no demonstration of the inter-rater (rater-rater) and intra-rater reliability of the observational techniques used. Additional caution is therefore required in interpreting these data and reliability and validity of the data collection tool used should be confirmed.

There was no evidence in the statistical modelling of the UK datasets that there were non-linear relationships between the staffing levels and nurse sensitive outcomes. However, the data are poor. We therefore modelled under the assumption that these effects were linear and that wards that have high levels of nurses would gain the same benefits from increasing their staffing further as wards who had low levels of nurses.

5.2.3 NFT data limitations

This data are available from only a single trust. This trust is an exemplary trust in that it is already employing a safer staffing toolkit to prevent wards from falling below a minimum staffing level and has a policy of collecting staffing and outcome data for analysis. Conclusions drawn from this dataset may not therefore be generalisable to other UK hospitals.

[Additional information to be added to this section following the consultation period of the draft guideline]

5.2.4 Cost-Effectiveness Analysis limitations

There are a number of limitations in the economic analysis stemming largely from the lack of NHS data. No data were available for missed care, and the data available for falls and drug errors were a decade old. Nursing practices, the quality of recording of adverse events and staffing levels have changed considerably in the intervening period. This may limit the applicability of the findings, and the limitations of the underlying data have been discussed in the previous subsections. The cost-effectiveness analysis relied upon US evidence on the relationship between drug errors outcomes and skill-mix, and there are concerns over the appropriateness of this for the NHS context.

Further, when calculating the cost-effectiveness of the interventions we have omitted several other potential effects. For example, there is some evidence in the literature that increased nursing levels, especially registered nurses, reduced mortality, length of stay, failure to rescue, missed care and a much wider number of adverse events. However, this literature is either conflicting, the evidence thought to be weak, the effect sizes relatively minor or the contribution of nurses specifically to the effect uncertain. Therefore the Safer Staffing Advisory Committee left these outcomes out of scope for this report. Thus, while it is uncertain, the estimate of the cost-effectiveness could be an underestimate of the true benefit of increasing nurse staffing or skill mix changes.

5.2.5 Statistical Analysis Limitations

The only outcome which was found to improve significantly with increased staff to patient ratios from the observational data (UKND) was the quality of documentation. This has substantial implications for studies which are based on review of nursing quality using documented outcomes and not direct observation. Observed trends in studies of this type may reflect changes in the quality of the recording of information and not actual patient outcomes.

The impact of skill mix was found to be variable. A higher proportion of staff nurses and charge nurses were associated with a higher rate of drug errors. These and other results suggest there are missing variables which may be generating confounding effects. Careful measurement of these variables (e.g. patient characteristics) is important for any future investigation.

The regression modelling was difficult. The outcome measures had several properties (unequal and large variance) which meant that the assumptions of different regression model types limited their applicability to these data. Different model types provided inconsistent results. This may be an inherent problem with data of this type and multiple modelling methods may be required to check the robustness of results in all studies of this type. Studies of this type, which fail to comment on the robustness of their findings should be treated with caution.

5.3 Recommendations for Future Research

Overall, the quality of data utilised here has limited this analysis. We are unable to draw any reliable conclusions from this investigation about the impact of staffing levels on patient outcomes. Improved data collection and outcomes monitoring is needed. A high quality investigation of patient outcomes in relation to staffing levels within the NHS will require collection of a wide variety of outcome variables at ward level matched to accurate staffing data. These outcomes include, but are not limited to, patient mortality, failure to rescue, infection rates, incidence of bed sores, medication errors, falls and validated measures of nursing quality, patient and relative satisfaction. High quality patient, ward and hospital data are also required. These would include, but are not limited to, ward occupancy, size, type, and speciality; validated measures of patient dependence, acuity, and medication requirements; and temporal changes in these measures across days, weekends, holidays, and seasons. Without these data the presence of spurious observational associations cannot be reliably excluded. Moreover, the availability of patient level costing data would enable a clearer understanding of the cost implications of staffing changes and skill mix, and when combined with Hospital Episode Statistics this could provide further information on, for example length of stay and readmission rates.

5.4 Evidence Summary

This report complements the systematic Evidence Review [4] produced by Griffiths et al. and aimed to produce a cost-effectiveness analysis of increasing nurse staffing levels and altering skill mix on three patient outcomes: falls, medication errors and missed care. The evidence contained in this report can be summarised as follows:

- Due to a lack of data or any published economic studies, missed care could not be considered in this evaluation.
- Four existing economic studies of nurse staffing and skill-mix on nurse sensitive outcomes exist but none are from the NHS, nor were at ward levels and none provide reliable estimates of the likely cost-effectiveness of these staffing interventions. It was therefore necessary to undertake statistical analysis of NHS ward level data.
- Three ward level datasets have been made available and each has its limitations. The Audit Commission dataset includes an almost complete sample of NHS trusts with most reporting more than 5 wards but it is over a decade old and has no patient dependency data. The UK Nursing Database has a smaller sample size, may potentially be a biased sample due to self-selection into the auditing programme, and does not provide outcomes data relevant to the scope of this report. The NFT dataset is high quality, daily data on patient outcomes, dependency and staffing but the trust already employs a patient acuity tool which will limit the variability in the staffing levels and may limit the generalizability to other trusts. Overall the quality of the data mean the results must be used with caution.
- Staffing level was only associated with the quality of reporting and no other outcome measures. This may indicate that better staffed wards have better recording of outcomes data which may lead to confounding in the analysis due to the data being self reported.
- A richer skill-mix, containing a greater proportion of registered nurses, reduced the rate of falls but increased the drug error rate. This may be due to better recording or unobservable patient factors. The finding in relation to falls is more robust to alternative model specifications and regression techniques, whereas the finding in relation to drug errors varied considerably across alternative model specifications and was frequently statistically

insignificant. As a result, we used published effects of nurse skill mix on drug errors from a US study for parameterise the economic model.

- Ward size and ward layout were found to be statistically significantly associated with a range of outcomes. There was also a relationship between staffing per bed and ward size with smaller wards having less variability and smaller ratios than larger wards. This is likely due to the minimum number of staff that it is possible to deploy on a ward at any one time.
- The Incremental Cost-Effectiveness Ratios were £1,412 per fall avoided and £127,823 per drug error avoided. Decision making in the absence of a universal measure, such as QALYS, is difficult but it is highly unlikely that altering skill mix to reduce drug errors is cost-effective for the NHS based upon this analysis. It is more likely that changing skill mix to reduce falls is cost-effective and the ICER compares favourably with other interventions to prevent falls. Ultimately, the cost-effectiveness will depend upon the NHS's willingness to pay for fall reduction.

6 References

1. Kane, R.L., et al., *Nurse staffing and quality of patient care.* Evid Rep Technol Assess (Full Rep), 2007(151): p. 1-115.
2. Kane, R.L., et al., *The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis.* Med Care, 2007. **45**(12): p. 1195-204.
3. Shekelle, P.G., *Nurse-patient ratios as a patient safety strategy: a systematic review.* Ann Intern Med, 2013. **158**(5 Pt 2): p. 404-9.
4. Griffiths, P., et al., *The association between patient safety outcomes and nurse/healthcare assistant skill mix and staffing levels & factors that may influence staffing requirements*. 2014, University of Southampton.
5. NHS Employers. *How Agenda for Change works*. 2012; Available from: http://www.nhsemployers.org/PayAndContracts/AgendaForChange/Pages/Afc-AtAGlanceRP.aspx.
6. Hurst, K., *Relationships between patient dependency, nursing workload and quality.* Int J Nurs Stud, 2005. **42**(1): p. 75-84

7. CHO, S. H., KETEFIAN, S., BARKAUSKAS, V. H. & SMITH, D. G. 2003. The effects of nurse staffing on adverse events, morbidity, mortality, and medical costs. *Nurs Res,* 52**,** 71-9.

8. DALL, T., CHEN, Y., SEIFERT, R., MADDOX, P. & HOGAN, P. 2009. The economic value of professional nursing. *Medical Care,* 47**,** 97.
9. NEEDLEMAN, J., BUERHAUS, P., STEWART, M., ZELEVINSKY, K. & STEWART, M. 2006. Nurse Staffing in Hospitals: is there a business case for quality? *Health Affairs,* 25**,** 204-211.

10. SHAMLIYAN, T. A., KANE, R. L., MUELLER, C., DUVAL, S. & WILT, T. J. 2009. Cost savings associated with increased RN staffing in acute care hospitals: simulation exercise. *Nursing Economics,* 27**,** 302-314.

11. TWIGG, D. E., GEELHOED, E. A., BREMNER, A. P. & M. DUFFIELD, C. 2013. The economic benefits of increased levels of nursing care in the hospital setting. *Journal of Advanced Nursing***,** n/a-n/a..

12. J. Needleman et al., “Nurse-Staffing Levels and Quality of Care in Hospitals,” New England Journal of Medicine, May 30, 2002 346(22):1715–

13 .Sovie MD, Jawad AF. Hospital restructuring and its impact on outcomes: nursing staff regulations are premature. J Nurs Adm. 2001;31: 588–600.

14. Unruh L. Licensed nurse staffing and adverse events in hospitals. Med Care. 2003;41:142–152.
15. Oliver D et al. Prevention of falls in hospitals and care homes and in persons in those settings with cognitive impairment or dementia. Department of Health Accidental Injury Prevention Programme. (2005).
16. National Patient Safety Agency. 2007a Slips, trips and falls in hospital
17. Irvine L et al. 2010 Cost-effectiveness of a day hospital falls prevention programme for screened community-dwelling older people at high risk of falls Age Ageing (2010) 39 (6): 710-716 first published online September 10, 2010 doi:10.1093/ageing/afq108
18. National Patient Safety Agency. 2007b Medication Safety
20. PSSRU. 2013 *Unit Costs of Health and Social Care 2013*.
21. Spetz, J. Jacobs, J and Hatler, C. 2007. Cost Effectiveness of a Medical Vigilance System to Reduce Patient Falls, N*ursing Economics*,25(6):333-338

7 Appendices

Appendix 1

Linear regression model of associations with ward and nurse staffing levels (by quintile) and number of falls per 1000 bed days.

	Falls per 1000 adjusted bed days		
	Estimate (95% confidence interval)	**t value**	**p value**
Ward type:			
Medical	0.00 (reference)		
Mixed	-5.17 (-7.69 to -2.65)	-4.101	<0.001
Surgical	-8.96 (-9.99 to -7.93)	-17.399	<0.001
Ward size (x=number of beds):			
x < 12	-2.08 (-4.45 to 0.29)	-1.759	0.079
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	1.75 (0.42 to 3.07)	2.635	0.008
26 < x ≤ 33	1.81 (0.42 to 3.19)	2.606	0.009
33 < x	0.82 (-1.12 to 2.75)	0.841	0.401
Ward type:			
Bays	0.00 (reference)		
Nightingale	-3.80 (-5.24 to -2.36)	-5.274	<0.001
Not recorded	-6.28 (-9.49 to -3.07)	-3.909	<0.001
Proportion of single rooms	1.26 (-1.85 to 4.37)	0.809	0.419
Proportion of days with high dependency patients	0.39 (-1.08 to 1.87)	0.536	0.592
Admissions per bed	0.07 (0.04 to 0.09)	4.943	<0.001
Discharges per bed	-0.55 (-0.70 to -0.39)	-6.998	<0.001
Transfers in per bed	0.01 (-0.22 to 0.23)	0.073	0.942
Transfers out per bed	-0.15 (-0.26 to -0.04)	-2.757	0.006
Total WTE per bed	1.16 (-0.15 to 2.47)	1.765	0.078
Senior nurse proportion (x):			
x ≤0.025	0.48 (-1.05 to 2.01)	0.629	0.529
0.025< x ≤0.05	0.00 (reference)		
0.05< x ≤0.075	0.82 (-0.36 to 2.01)	1.392	0.164
0.075< x ≤0.1	-1.42 (-3.60 to 0.77)	-1.297	0.195
0.1< x	-1.98 (-5.04 to 1.09)	-1.291	0.197
Charge nurse proportion (x):			
x ≤0.05	0.95 (-0.06 to 1.97)	1.882	0.060
0.05< x ≤0.10	0.00 (reference)		
0.10< x ≤0.15	-2.31 (-3.62 to -0.99)	-3.503	<0.001
0.15< x ≤0.20	-3.90 (-6.15 to -1.65)	-3.472	<0.001
0.20< x	-2.23 (-6.09 to 1.63)	-1.153	0.249
Staff nurse proportion (x):			
x ≤0.4	5.16 (3.47 to 6.85)	6.105	<0.001
0.4< x ≤0.5	3.89 (2.74 to 5.05)	6.759	<0.001
0.5< x ≤0.6	0.00 (reference)		
0.6< x ≤0.7	-2.02 (-3.22 to -0.81)	-3.349	<0.001
0.7< x	-3.46 (-5.21 to -1.72)	-3.965	<0.001

Linear regression model of associations with ward and HCA staffing levels (by quintile) and number of falls per 1000 bed days.

		Falls per 1000 adjusted bed days		
		Estimate (95% confidence interval)	**t value**	**p value**
Ward type:				
	Medical	0.00 (reference)		
	Mixed	-5.21 (-7.73 to -2.68)	-4.127	<0.001
	Surgical	-9.21 (-10.24 to -8.19)	-18.006	<0.001
Ward size (x=number of beds):				
	x < 12	-2.17 (-4.48 to 0.13)	-1.887	0.059
	12 < x ≤ 20	0.00 (reference)		
	20 < x ≤ 26	1.58 (0.28 to 2.88)	2.433	0.015
	26 < x ≤ 33	1.46 (0.15 to 2.76)	2.227	0.026
	33 < x	0.32 (-1.55 to 2.19)	0.339	0.734
Ward type:				
	Bays	0.00 (reference)		
	Nightingale	-3.79 (-5.23 to -2.34)	-5.249	<0.001
	Not recorded	-6.18 (-9.40 to -2.97)	-3.846	<0.001
Proportion of single rooms		1.27 (-1.85 to 4.39)	0.812	0.417
Proportion of days with high dependency patients		0.27 (-1.21 to 1.74)	0.361	0.718
Admissions per bed		0.07 (0.04 to 0.10)	5.029	<0.001
Discharges per bed		-0.54 (-0.70 to -0.39)	-6.939	<0.001
Transfers in per bed		0.00 (-0.23 to 0.22)	-0.022	0.982
Transfers out per bed		-0.17 (-0.28 to -0.06)	-3.074	0.002
Total WTE per bed		0.86 (-0.44 to 2.16)	1.321	0.186
HCA proportion (x):				
	x ≤0.1	-0.41 (-3.13 to 2.31)	-0.301	0.764
	0.1< x ≤0.2	0.00 (reference)		
	0.2< x ≤0.3	2.69 (0.90 to 4.49)	3.004	0.003
	0.3< x ≤0.4	4.47 (2.73 to 6.21)	5.13	<0.001
	0.4< x	8.35 (6.53 to 10.17)	9.186	<0.001

Appendix 2

Linear regression model of associations with ward and nurse staffing levels (by quintile) and number of drug errors per 1000 bed days.

	Drug errors per 1000 adjusted bed days		
	Estimate (95% confidence interval)	**t value**	**p value**
Ward type:			
Medical	0.00 (reference)		
Mixed	-22.41 (-44.35 to -0.46)	-2.042	0.041
Surgical	-12.79 (-21.90 to -3.68)	-2.809	0.005
Ward size (x=number of beds):			
x < 12	-33.87 (-55.39 to -12.35)	-3.148	0.002
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	6.90 (-4.94 to 18.75)	1.166	0.244
26 < x ≤ 33	19.66 (7.45 to 31.88)	3.22	0.001
33 < x	34.29 (17.41 to 51.16)	4.064	<0.001
Ward type:			
Bays	0.00 (reference)		
Nightingale	-9.17 (-22.37 to 4.03)	-1.39	0.165
Not recorded	18.96 (-12.12 to 50.04)	1.22	0.223
Proportion of single rooms	-3.82 (-32.05 to 24.41)	-0.271	0.787
Proportion of days with high dependency patients	12.53 (-0.43 to 25.49)	1.934	0.053
Admissions per bed	0.06 (-0.16 to 0.29)	0.569	0.569
Discharges per bed	-1.20 (-2.57 to 0.17)	-1.755	0.079
Transfers in per bed	4.14 (2.22 to 6.05)	4.327	<0.001
Transfers out per bed	0.39 (-0.62 to 1.41)	0.77	0.441
Total WTE per bed	10.59 (0.20 to 20.98)	2.039	0.042
Senior nurse proportion (x):			
x ≤0.025	-2.17 (-15.96 to 11.62)	-0.314	0.753
0.025< x ≤0.05	0.00 (reference)		
0.05< x ≤0.075	-14.20 (-24.69 to -3.71)	-2.706	0.007
0.075< x ≤0.1	-7.59 (-27.41 to 12.24)	-0.766	0.444
0.1< x	-19.50 (-46.65 to 7.66)	-1.436	0.151
Charge nurse proportion (x):			
x ≤0.05	-8.50 (-17.49 to 0.49)	-1.89	0.059
0.05< x ≤0.10	0.00 (reference)		
0.10< x ≤0.15	9.38 (-2.22 to 20.99)	1.617	0.106
0.15< x ≤0.20	21.82 (2.15 to 41.49)	2.218	0.027
0.20< x	13.23 (-21.44 to 47.89)	0.763	0.445
Staff nurse proportion (x):			
x ≤0.4	-20.39 (-35.19 to -5.59)	-2.755	0.006
0.4< x ≤0.5	-5.01 (-15.19 to 5.17)	-0.984	0.325
0.5< x ≤0.6	0.00 (reference)		
0.6< x ≤0.7	22.54 (11.89 to 33.19)	4.233	<0.001
0.7< x	14.27 (-1.61 to 30.15)	1.797	0.072

Appendix 3

Here we present additional models using alternative modelling methods for identifying associations with falls and drug errors. Linear regression has limited applicability to count-type variables (such as number of falls and number of drug errors) as it assumes a continuous distribution of the outcome variable which is normally distributed. We therefore elected to use other model types to test the robustness of the correlations demonstrated by the linear regression analyses presented. We were unable to use Poisson regression as both outcome variables had substantial over dispersion. Quasi-Poisson and negative binomial regression analyses were used to allow correction for this over dispersion and are designed for use with count-type outcomes. These models assume a Poisson distribution and negative binomial distribution in the outcome variables respectively. In both these model types log of 1000 adjusted bed days was used as the exposure variable. The results are presented below.

Both the quasi-Poisson and negative binomial regression models demonstrate an inverse relationship with falls and ward size; larger wards were associated with fewer falls (Table A3.1 and Table A3.2). This is in contrast to the positive relationship identified with linear regression; smaller wards were associated with fewer falls. Fewer falls were associated with ward type not being recorded in both these models. This is suggestive of recording biases.

The negative binomial model demonstrates a significant relationship between increasing staff numbers and increasing falls. This was not identified in the linear regression or quasi-Poisson models and is therefore not robust. No associations with skill mix were identified in either of these additional models. Squared and cubic forms of all skill mix variables (proportion of senior nurses, proportion of charge nurses, proportion of staff nurses, and proportion of HCAs) were also found to have no association with falls in either model type [results not shown]. Dividing each of these skill mix variables into quintiles also failed to demonstrate any association with falls in either model type [results not shown].

Both quasi-Poisson and negative binomial regression analyses also demonstrate an inverse relationship between drug errors and ward size; larger wards are associated with fewer drug errors (Table A3.3 and Table A3.4). This is contradiction with the linear regression analysis (as with the association with falls and ward size described above). Surgical and mixed medical and surgical wards were associated with fewer drug errors in both models. This is consistent with the linear regression analysis. Patient transfers in per bed was consistently associated with a higher number of drug errors.

The quasi-Poisson regression model demonstrated a small association between increased overall staffing level and more drug errors. There was no association between overall staffing and drug errors in the negative binomial model. A higher proportion of charge nurses and staff nurses (relative to the total nursing and health care assistant staffing level) was associated with more drug errors in both types of model. This was consistent with the results of the linear regression analysis.

The contradictory results between the linear and other models regarding the association between ward size and number of patient falls appear to arise from increased variability of the number of falls per bed for the smaller wards. This, in turn, arises because as ward size decreases a single event (fall or drug error) has a larger impact on the number of events per bed. This leads to considerable heteroskedasticity in the data. This effect has differing impacts on the different model types used here. The results of all these models should be treated with extreme caution when considering the ward size variable.

	Number of falls		
	Estimate (95% confidence interval)	t value	p value
Ward specialty:			
Medical	0.00 (reference)		
Mixed	0.13 (-0.05 to 0.31)	1.47	0.142
Surgical	0.06 (-0.01 to 0.14)	1.644	0.100
Ward size (x=number of beds):			
x < 12	0.16 (-0.05 to 0.37)	1.493	0.136
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	-0.15 (-0.25 to -0.06)	-3.221	0.001
26 < x ≤ 33	-0.37 (-0.47 to -0.28)	-7.539	<0.001
33 < x	-0.51 (-0.64 to -0.38)	-7.689	<0.001
Ward type:			
Bays	0.00 (reference)		
Nightingale	0.08 (-0.02 to 0.18)	1.571	0.116
Not recorded	-0.36 (-0.63 to -0.08)	-2.569	0.010
Proportion of single rooms	0.12 (-0.11 to 0.36)	1.056	0.291
Proportion of days with high dependency patients	0.06 (-0.05 to 0.16)	1.09	0.276
Admissions per bed	0.00 (0.00 to 0.00)	-0.046	0.964
Discharges per bed	-0.01 (-0.03 to 0.00)	-1.739	0.082
Transfers in per bed	-0.01 (-0.03 to 0.01)	-0.725	0.469
Transfers out per bed	0.00 (0.00 to 0.01)	0.883	0.377
Total WTE per bed	0.08 (-0.06 to 0.21)	1.117	0.264
Proportion of senior nurses	-1.04 (-2.57 to 0.48)	-1.367	0.172
Proportion of charge nurses	-0.48 (-1.23 to 0.27)	-1.283	0.200
Proportion of staff nurses	-0.26 (-0.57 to 0.06)	-1.639	0.101

Table A3.1 **Quasi-Poison regression** analysis of ward, patient, staffing factors with number of **falls**.

	Number of falls		
	Estimate (95% confidence interval)	**t value**	**p value**
Ward specialty:			
Medical	0.00 (reference)		
Mixed	0.14 (-0.06 to 0.33)	1.411	0.158
Surgical	0.05 (-0.03 to 0.12)	1.15	0.250
Ward size (x=number of beds):			
x < 12	0.05 (-0.13 to 0.23)	0.554	0.579
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	-0.15 (-0.25 to -0.05)	-3.091	0.002
26 < x ≤ 33	-0.36 (-0.46 to -0.26)	-7.257	<0.001
33 < x	-0.47 (-0.62 to -0.33)	-6.544	<0.001
Ward type:			
Bays	0.00 (reference)		
Nightingale	0.08 (-0.03 to 0.19)	1.485	0.138
Not recorded	-0.36 (-0.61 to -0.12)	-2.933	0.003
Proportion of single rooms	0.19 (-0.04 to 0.43)	1.634	0.102
Proportion of days with high dependency patients	0.05 (-0.06 to 0.17)	0.977	0.329
Admissions per bed	-0.01 (-0.04 to 0.01)	-1.266	0.205
Discharges per bed	0.00 (-0.03 to 0.02)	-0.089	0.929
Transfers in per bed	-0.02 (-0.04 to 0.00)	-1.559	0.119
Transfers out per bed	0.02 (-0.01 to 0.04)	1.46	0.144
Total WTE per bed	0.14 (0.05 to 0.23)	2.998	0.003
Proportion of senior nurses	-1.26 (-2.75 to 0.23)	-1.689	0.091
Proportion of charge nurses	-0.51 (-1.26 to 0.23)	-1.375	0.169
Proportion of staff nurses	-0.31 (-0.64 to 0.02)	-1.89	0.059

Table A3.2 **Negative binomial regression** analysis of ward, patient, staffing factors with number of **falls**.

	Number of drug errors		
	Estimate (95% confidence interval)	t value	p value
Ward specialty:			
Medical	0.00 (reference)		
Mixed	-0.34 (-0.66 to -0.02)	-2.099	0.036
Surgical	-0.16 (-0.28 to -0.03)	-2.494	0.013
Ward size (x=number of beds):			
x < 12	-0.38 (-0.80 to 0.04)	-1.812	0.070
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	-0.12 (-0.29 to 0.05)	-1.448	0.148
26 < x ≤ 33	-0.10 (-0.27 to 0.07)	-1.214	0.225
33 < x	-0.21 (-0.42 to 0.00)	-1.974	0.049
Ward type:			
Bays	0.00 (reference)		
Nightingale	-0.15 (-0.35 to 0.05)	-1.505	0.133
Not recorded	0.30 (-0.12 to 0.71)	1.442	0.150
Proportion of single rooms	-0.14 (-0.54 to 0.27)	-0.677	0.499
Proportion of days with high dependency patients	0.13 (-0.04 to 0.29)	1.525	0.127
Admissions per bed	-0.03 (-0.06 to -0.01)	-2.377	0.018
Discharges per bed	0.02 (-0.01 to 0.06)	1.262	0.207
Transfers in per bed	0.03 (0.00 to 0.05)	1.959	0.050
Transfers out per bed	0.03 (0.00 to 0.07)	2.154	0.031
Total WTE per bed	0.40 (0.28 to 0.51)	6.795	<0.001
Proportion of senior nurses	-1.35 (-4.08 to 1.38)	-0.991	0.322
Proportion of charge nurses	2.61 (1.43 to 3.80)	4.419	<0.001
Proportion of staff nurses	1.58 (1.06 to 2.11)	6.025	<0.001

Table A3.3 **Quasi-Poison regression** analysis of ward, patient, staffing factors with number of **drug errors**.

	Number of drugs errors		
	Estimate (95% confidence interval)	**t value**	**p value**
Ward specialty:			
Medical	0.00 (reference)		
Mixed	-0.34 (-0.61 to -0.08)	-2.617	0.009
Surgical	-0.17 (-0.27 to -0.06)	-3.179	0.001
Ward size (x=number of beds):			
x < 12	-0.39 (-0.69 to -0.10)	-2.71	0.007
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	-0.06 (-0.20 to 0.08)	-0.807	0.420
26 < x ≤ 33	-0.01 (-0.15 to 0.13)	-0.112	0.911
33 < x	-0.13 (-0.32 to 0.05)	-1.414	0.157
Ward type:			
Bays	0.00 (reference)		
Nightingale	-0.11 (-0.27 to 0.05)	-1.403	0.161
Not recorded	0.40 (0.05 to 0.76)	2.254	0.024
Proportion of single rooms	-0.05 (-0.38 to 0.29)	-0.29	0.772
Proportion of days with high dependency patients	0.09 (-0.05 to 0.24)	1.289	0.197
Admissions per bed	-0.02 (-0.05 to 0.00)	-1.772	0.076
Discharges per bed	0.01 (-0.02 to 0.04)	0.606	0.545
Transfers in per bed	0.04 (0.01 to 0.06)	2.865	0.004
Transfers out per bed	0.03 (0.00 to 0.06)	1.806	0.071
Total WTE per bed	0.43 (0.32 to 0.54)	7.591	<0.001
Proportion of senior nurses	-0.93 (-3.15 to 1.28)	-0.841	0.400
Proportion of charge nurses	2.62 (1.59 to 3.66)	5.073	<0.001
Proportion of staff nurses	1.63 (1.18 to 2.08)	7.253	<0.001

Table A3.4 **Negative binomial regression** analysis of ward, patient, staffing factors with number of **drug errors**.

Appendix 4

Model of recording of patients' response to care and progress with overall staffing levels categorised into quintiles.

	How well patients' progress was recorded		
	N (%) or mean (SD)	**Odds Ratio (Confidence limits)**	**P value**
Hospital type:			
Teaching hospital	216 (47.3%)	1.00 (reference)	
Community hospital	121 (26.5%)	-9.15 (-16.70 to -1.59)	0.016
District general hospital	116 (25.4%)	0.43 (-5.24 to 6.09)	0.880
Ward layout:			
Bays	283 (61.9%)	1.00 (reference)	
Racetrack	24 (5.3%)	-4.80 (-13.68 to 4.09)	0.281
Single rooms	25 (5.5%)	10.19 (0.99 to 19.40)	0.027
Nightingale	10 (2.2%)	-6.53 (-20.98 to 7.92)	0.367
Split site	14 (3.1%)	0.61 (-10.84 to 12.05)	0.916
Other	101 (22.1%)	-7.44 (-12.71 to -2.17)	0.005
Ward occupancy (x=number of patients):			
$x < 12$	125 (27.4%)	-0.75 (-8.57 to 7.08)	0.849
$12 < x \leq 20$	44 (9.6%)	-0.60 (-6.04 to 4.84)	0.825
$20 < x \leq 26$	139 (30.4%)	1.00 (reference)	
$26 < x \leq 33$	119 (26.0%)	-8.41 (-15.11 to -1.71)	0.012
$33 < x$	30 (6.6%)	-11.49 (-21.22 to -1.77)	0.019
Patient dependency (% of all patients):			
Low dependency	8.63 (±11.19)	-0.35 (-0.62 to -0.08)	0.011
Low-medium dependency	42.36 (±19.68)	-0.08 (-0.25 to 0.09)	0.353
Medium dependency	42.36 (±17.66)	-0.06 (-0.28 to 0.165)	0.609
Staff level (% of all staff):			
Ward sisters	8.63 (±11.19)	-0.14 (-0.42 to 0.14)	0.323
Staff nurses	42.36 (±19.68)	0.09 (-0.05 to 0.23)	0.216
Support staff	37.02 (±17.66)	0.05 (-0.05 to 0.15)	0.300
Bank staff	11.99 (±13.62)	-0.17 (-0.46 to 0.12)	0.239
Total staffing level (x = WTE per occupied bed):			
$x < 1.0$	87 (22.1%)	5.94 (-1.03 to 12.90)	0.089
$1.0 \leq x < 1.2$	65 (16.5%)	1.00 (reference)	
$1.2 \leq x < 1.4$	78 (19.8%)	9.65 (2.92 to 16.37)	0.004
$1.4 \leq x < 1.6$	51 (13.0%)	7.54 (-0.17 to 15.26)	0.051
$1.6 \leq x$	112 (28.5%)	5.80 (-1.29 to 12.89)	0.103

Appendix 5

Linear regression models entering staffing as WTE per bed and skill mix as proportion of total WTE that are nurses.

Multiple R-squared: 0.2262,   Adjusted R-squared: 0.2226

	Falls per 1000 adjusted bed days		
	Estimate (95% confidence interval)	**t value**	**p value**
Ward type:			
Medical	0.00 (reference)		
Mixed	-5.01 (-7.54 to -2.49)	-3.977	<0.001
Surgical	-9.13 (-10.15 to -8.11)	-17.892	<0.001
Ward size (x=number of beds):			
x < 12	-2.03 (-4.33 to 0.28)	-1.756	0.079
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	1.52 (0.22 to 2.82)	2.346	0.019
26 < x ≤ 33	1.35 (0.04 to 2.65)	2.066	0.039
33 < x	0.11 (-1.77 to 1.98)	0.115	0.908
Ward type:			
Bays	0.00 (reference)		
Nightingale	-3.77 (-5.21 to -2.33)	-5.24	<0.001
Not recorded	-6.11 (-9.32 to -2.90)	-3.807	<0.001
Proportion of single rooms	1.43 (-1.69 to 4.55)	0.918	0.359
Proportion of days with high dependency patients	0.29 (-1.19 to 1.76)	0.39	0.696
Admissions per bed	0.07 (0.04 to 0.10)	4.91	<0.001
Discharges per bed	-0.53 (-0.69 to -0.37)	-6.806	<0.001
Transfers in per bed	0.00 (-0.23 to 0.22)	-0.006	0.995
Transfers out per bed	-0.17 (-0.28 to -0.06)	-3.025	0.003
Nurse WTE per bed	2.03 (-1.33 to 5.39)	1.209	0.227
HCA WTE per bed	-2.38 (-10.06 to 5.30)	-0.621	0.535
Proportion of nurses	-0.28 (-0.39 to -0.17)	-5.109	<0.001

Linear regression models entering staffing as WTE per bed and skill mix as proportion of total WTE that are nurses.

Multiple R-squared: 0.04209, Adjusted R-squared: 0.03685

	Drug errors per 1000 bed days		
Ward type:			
Medical	0.00 (reference)		
Mixed	-22.12 (-44.10 to -0.15)	-2.014	0.044
Surgical	-11.80 (-20.83 to -2.78)	-2.615	0.009
Ward size (x=number of beds):			
x < 12	-42.18 (-63.20 to -21.15)	-4.012	<0.001
12 < x ≤ 20	0.00 (reference)		
20 < x ≤ 26	10.99 (-0.67 to 22.64)	1.886	0.059
26 < x ≤ 33	26.15 (14.53 to 37.77)	4.501	<0.001
33 < x	42.71 (26.37 to 59.05)	5.228	<0.001
Ward type:			
Bays	0.00 (reference)		
Nightingale	-8.73 (-21.95 to 4.49)	-1.321	0.187
Not recorded	11.92 (-19.28 to 43.13)	0.764	0.445
Proportion of single rooms	-0.66 (-29.19 to 27.88)	-0.046	0.963
Proportion of days with high dependency patients	12.97 (0.00 to 25.93)	2.001	0.046
Admissions per bed	-0.02 (-0.24 to 0.21)	-0.135	0.892
Discharges per bed	-1.50 (-2.88 to -0.12)	-2.174	0.030
Transfers in per bed	4.11 (2.19 to 6.02)	4.295	<0.001
Transfers out per bed	0.68 (-0.33 to 1.69)	1.338	0.181
Nurse WTE per bed	37.15 (5.91 to 68.38)	2.378	0.017
HCA WTE per bed	-30.41 (-99.69 to 38.87)	-0.878	0.380
Proportion of nurses	0.39 (-0.61 to 1.40)	0.781	0.435
Ward type:			