Predicting Nurse Staffing Requirements From Routinely Collected Data (PREDICT-NURSE) protocol

Version	Date	Main changes	Editor
1.0	16th Feb 2025	Formatting including adding version control table. This version is pending ethical approval. Added research team details and funding disclaimer.	Christina Saville
2.0	30th April 2025	Added end of study date following HRA review.	Christina Saville

Research team

Name	Role	Organisation
Christina Saville	PI	University of Southampton
Peter Griffiths	Co-I	
Paul Meredith	Co-I	
Chiara Dall'Ora	Co-I	
Jamal Hossain	Co-I	
Sue Wierzbicki	Co-I	Portsmouth Hospital University NHS Trust
Tom Weeks	Co-I	
Sinead O'Neill	Co-I	NHSE
Mary Wells	Co-I	Imperial College Healthcare NHS Trust Imperial College London
James Bird	Co-I	
Erik Mayer	Co-I	
Ian Dickerson	Co-I	PPI representative

Funding disclaimer

This study/project is funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research Programme (ref NIHR166784). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

1. Summary of research

Background: Having enough nurses on hospital wards is vital for patient safety but planning for varying numbers and needs of patients is hard. Almost all acute NHS Trusts in England use the NICE-endorsed Safer Nursing Care Tool (SNCT) to guide staffing decisions. However, this approach is labour-intensive and necessitates the collection of data specifically to measure staffing requirements, not informed by data gathered for administration or care management.

Aim: Develop a method to measure demand for nursing staff on hospital wards using routine data to help plan establishments (number of ward employees), monitor staffing adequacy in real-time, and inform safe and efficient deployment of staff.

Design: A retrospective observational study across wards providing acute adult somatic (i.e. not mental health) inpatient care in 5 general hospital Trusts, predicting nurse staffing requirements from routinely collected data and validating these predictions against patient and staffing adequacy outcomes. Algorithms will be developed according to user-centred design and by engaging with patients to understand experiences of hospital nurse staffing and implications for developing algorithms.

Workstream (WS) 1 Objective: understand what does/does not work for nurses and managers when using staffing tools, and incorporate this into algorithm design. **Method:** User-centred design approach comprising i) a national survey of staffing matrons and Chief Nursing Information Officers to find out how staffing tools are used and patient data availability/quality, ii) workshops with nurses and nursing managers to understand staffing decision support needs at different timepoints, iii) workshops with this group plus NHS IT managers and roster companies to discuss algorithm design considerations.

WS2 Objective: develop statistical/machine learning algorithms to estimate nurse staffing requirements from routinely available patient data. **Method:** Since there is no "gold standard" for measuring nurse staffing requirements, we will first replicate measurements from the SNCT, a patient acuity/dependency classification tool. We will develop alternative algorithms including replicating individual patient acuity/dependency classifications and replicating the staffing requirements for a whole ward. We will consider staffing decisions at different timepoints. Our predictor variables will come from administrative and care plan data.

WS3 Objective: assess the validity of algorithms. **Method:** We will fit regression models to investigate the associations between actual under/over-staffing relative to each candidate measure of staffing requirements and multiple outcomes. For this, we will use routine data extracted from hospital IT systems and a micro-survey of nurses to understand perceptions of staffing adequacy. We will test whether as staffing increases relative to a measure of staffing requirements, the risk of poor patient outcomes and perceptions that staffing is inadequate decreases. We will compare model fit against models with staffing requirements measured by the SNCT.

Timelines: 2.5 years

Anticipated impact: A better match between staffing and workload on hospital wards, more efficient deployment of scarce resources and less time-consuming staffing assessments

Dissemination: Open-access journal articles, magazine articles for nurses and videos/posters for the public. We will share results with intended users through workshops and user groups.

2. Background and rationale

2.1. What is the problem being addressed?

Having the **right number of nurses** caring for patients on hospital wards is vital for patient safety but planning for varying numbers of patients with unknown deterioration/recovery trajectories is hard.(1, 2) Extensive research shows that care quality and safety are compromised when nurse staffing is low,(3-6) but deciding how to allocate nursing staff is challenging because demand for care fluctuates;(7) both patient numbers and their needs vary over time. Decisions arise at multiple points including determining how many staff to employ (the establishment) and how many to deploy on each shift.(8) Given current staffing shortages, being able to identify wards to prioritise and (potentially) identify any spare staffing capacity is especially important.

To help guide decisions, a range of **tools attempt to measure demand** for nursing care and determine the required number of staff.(9) In England, the NICE-endorsed Safer Nursing Care Tool (SNCT), now used in almost all acute NHS Trusts.(10) Designed for establishment planning, it requires that patients are classified into one of five categories based on an assessment of acuity and dependency on nursing care. The category 'multipliers' reflect the relative demand for care and formulae are provided to estimate the number of nurses to employ to safely staff the ward, allowing for variation in demand and staff absences. When used for establishment planning, trained nurse observers assess all patients on a ward daily for 30 days and the establishment is set based on an average.(11) Increasingly the acuity/dependency classification is carried out by clinical nurses, sometimes multiple times per day, and results are used for real-time monitoring of demand and to guide deployment of staff.(7, 12)

However, the SNCT approach is **labour-intensive** and necessitates the collection of data specifically to measure staffing requirements, not informed by data gathered for administration or care management. Because SNCT assessments are intermittent, there is limited ability to monitor demand and take actions to match it in real-time. Even when used for establishment setting, prescribed methods for sampling do not reflect typical demand on some types of ward.(7, 13) Furthermore, the use of a limited number of categories does not reflect the continuous variability of need.

Our project seeks to develop and test methods to estimate nurse staffing requirements using routinely collected data, initially using the existing method (SNCT) as a benchmark. If successful, nurses' time can be saved, because bespoke assessments can be avoided and algorithms can allow real-time monitoring of variation in demand.

2.2. Why is this research important and needed now?

In a time of nursing shortages and a focus on delivering high quality and safe care in the most efficient manner (14), Trusts are turning to information systems to assist with decisions around deployment and redeployment of staff. A tool designed for use in establishment decisions has been taken up by many Trusts for use in real-time decision-making too (10), but with associated increases in documentation time, as well as increased training and monitoring requirements, due to the wider group of people entering data. In our PPI work, patients highlighted the importance of nursing staff being visible and available rather than on a computer, so a further increase in documentation counteracts this. When nurses are not available to monitor patients they fail to spot deterioration, so freeing up nurses' time might have patient safety implications (15). With the increasing availability of patient, administrative and care data available electronically in information systems, now is the first time that it is feasible to consider using this data for predicting staffing requirements in real-time, thus potentially providing greater reliability of staffing requirement data, as well as removing administrative burdens.

The problems associated with having **insufficient nursing staff** as measured against the assessed patient need on a hospital ward are well-evidenced.(16, 17) Negative outcomes for patients when there are too few staff include negative experiences of care, omissions of necessary care, adverse events, delayed discharge and increased risk of death.(4-6) In our previous HS&DR study [14/194/21], we found that flexible staffing strategies (redeploying staff and use of temporary staff) have the potential to reduce understaffing rates and be cost-effective, provided that baseline staffing establishments are sufficient.(7)

But in order to make decisions about redeploying staff or hiring temporary staff, **accurate information** about ward staffing requirements needs to be available in a **timely** way, which is currently not the case across the NHS. Providing up-to-date, objective and accurate information about demand for nursing care has large potential benefits, allowing areas at risk of short-staffing to be identified and prioritised for the deployment of staff.(8) Trusts are already using software for deployment decisions (generally using an acuity dependency measure based on the SNCT), which demonstrates this is seen as necessary, but the SNCT was neither designed nor validated for this purpose.

Currently the NHS faces significant and enduring **staff shortages**,(14) making it especially important to deploy the available staff efficiently and effectively. Reducing cases of severe or sustained understaffing on a ward has potential benefits for staff as well as the organisation more generally; prolonged exposure to understaffing is a risk factor for staff burnout as well as staff turnover and sickness absence.(18, NIHR 128056 under review)

Additional **benefits of workforce planning tools** and technologies include promoting the patient safety agenda within a Trust and helping learn 'what works' in effective staffing by comparison across wards.(19) Although there are problems with the translation from research to practice,(1) there is evidence that when a user-centred design approach is used, tools can be successfully designed and implemented by embedding in existing software.(20)

Minimising the **administrative burden** on staff of gathering data to monitor staffing requirements is also important. While assessing an individual patient to determine the SNCT category is not unduly time-consuming, when multiplied across patients and multiple daily assessments (as is increasingly being done), the workload is considerable. Our observational data (unpublished) suggests that assessing all patients on a thirty-bedded ward three times per day could occupy a senior nurse for up to 30 minutes per day, plus extra time for data entry. Additionally, to ensure reliable and accurate assessments, considerable time is required to train and assess staff competence, with high staff turnover and reliance on temporary staff creating further challenges.

2.3. Review of existing evidence

There is a clear continued demand for **nurse workforce planning tools** and systems to help with nurse staffing decisions, as demonstrated by the large volume of papers on tool development, dating back to the 1920s and still proliferating today.(9, 21) These tools include volume-based methods (such as patient-to-nurse ratios), patient prototype/ classification, professional judgement and timed-task approaches. Multiple review papers have found that despite the continual development of new tools, there is little evidence for their effectiveness in recommending appropriate staffing levels, as most papers focus on tool development rather than validation.(9, 22-24) In a survey with responses from 91 of the 148 acute care Trusts in England, 80% reported using the SNCT alongside their professional judgement for establishment reviews.(10) They also reported reviewing the adequacy of nurse staffing at the start of each shift with 75% of Trusts using professional judgement and 69% of Trusts using patient acuity/dependency systems.(10) A previous HS&DR-funded ethnographic study also found evidence for the usefulness of tools, with identified "programme theories" including that NHS managers need to combine local knowledge and professional judgement with data from workforce planning and deployment tools and technologies for effective staffing decisions.(19)

Despite its widespread use, in common with other nurse staffing tools, until recently there has been little **evidence about the SNCT**'s ability to accurately determine staffing requirements.(9) The SNCT multipliers are based on timed observations of patient care in over 2800 "high-quality" wards, although detailed study reports are lacking.(7) Recent research has increased confidence in the tool as a useful measure of staffing requirements; shortfalls from required daily staffing levels derived from the SNCT are associated with nurses' professional judgements of inadequate staffing.(13) In an observational study of over 130,000 patients in one English hospital we found that the hazard of death was increased by 9% when patients were exposed to staffing below the SNCT-recommended level.(25) These findings are consistent with other studies that show increased risk of death when staffing falls below requirements assessed using staffing tools.(26, 27) However there is currently a lack of clear evidence of the burden of data collection, and limited understanding of how current real-time systems are being used, but the time and workload is likely to be considerable.(12)

We searched the literature for **algorithmic tools** that use routine data (as opposed to bespoke assessments) to estimate nurse staffing requirements in general wards but found no evidence of approaches validated against existing widely-used approaches like the SNCT. The Rafaela system is one of the few nurse staffing methods with comparable evidence to the SNCT. This system has more complex multifactorial assessments(28) but evaluation studies have identified problems with compliance and reliability in practice, leading researchers to recommend against implementation.(29) Systems that attempt to directly quantify staffing requirements from the care plan are limited by the comprehensiveness of factors included and the difficulties in gaining reliable empirical estimates of time, often relying on professional judgement.(9, 30-32) Several studies have demonstrated strong correlations between nursing workload and clinical assessments,(30, 33) data available in the patient record,(30) or ward activity data e.g. transfers.(30) We found one example of an algorithmic model predicting general ward patients' acuity, although the paper is not transparent about the predictors used.(34) However, we found no studies assessing how well existing data replicates commonly-used measures of staffing requirements, which is the research gap we will address in this study.

3. Aims and objectives

This study aims to develop a method to measure demand for nursing staff on hospital wards using routine data, to facilitate establishment planning, to monitor staffing adequacy/safety in real-time and to inform safe and efficient deployment of staff. To achieve this we have a number of objectives:

1) understand what works and does not work for nurses and nurse managers when using staffing tools, and incorporating this into the design of an algorithm for estimating staffing requirements

2) develop statistical/machine learning algorithms to estimate nurse staffing requirements from routinely available patient data, starting with replicating Safer Nursing Care Tool assessments

3) assess the validity of algorithms by determining whether staffing below the estimated staffing requirements is associated with adverse patient outcomes/ nurse perceptions of inadequate staffing.

Throughout the study we will engage with patients and the public to understand their experiences of hospital nurse staffing and implications for our algorithm development.

4. Research plan / methods

4.1. Design

This is a **retrospective observational** study across multiple hospital Trusts, making use of routinely collected data to predict nurse staffing requirements and validate these predictions against patient and staffing outcomes. The algorithms will be developed according to **user-centred design** through engagement with nurses and nurse managers at key phases at the start and during the project. **Engagement with patients** is threaded throughout the study period.

4.2. Scope and selecting research sites

The **scope** of our study is all patients staying on inpatient wards (including admission units) for adult acute somatic (i.e. not mental health) care, as this is the scope of the Safer Nursing Care Tool.

We aim to include a **diverse sample** of hospital wards across a variety of hospital sizes, specialties, and geographical locations. Specifically, we plan to analyse data from at least five NHS hospital Trusts in England (hospitals in other parts of the UK use different staffing tools), starting with our two partner hospital Trusts. To test generalisability, we will recruit at least three further hospitals through groups such as the Shelford Group (who we have preliminary support from*)* or users of particular Electronic Patient Record systems or roster services such as Oceansblue (who have worked with us on previous projects). We will purposefully over-recruit to allow for drop-off. By collaborating with hospitals serving diverse populations, we aim to reduce the risk of algorithm bias resulting from training on unrepresentative data sets. While it will not be possible to sample to achieve generalisation / representativeness in a statistical sense, selection from willing trusts will be guided by diversity of clinical populations, Trust type (e,g. teaching versus not) and available data systems. This will increase confidence that algorithms work in diverse settings or identify possible limits of generalisability. We will favour sites who are willing and able to run the micro-survey component.

We have confirmed involvement of **Portsmouth** Hospitals University NHS Trust (where we performed our previous HS&DR studies NIHR 13/114/17 & 128056) and **Imperial** College Healthcare NHS Trust. Portsmouth Trust employs approximately 9,100 whole-time equivalent staff and provides acute services to about 675,000 people across Portsmouth and south-east Hampshire.(35) Imperial Trust employs more than 15,000 staff across five sites, of which three are relevant to our study (Charing Cross, Hammersmith, St Marys).(36) It serves eight London boroughs in the North West London Integrated Care System. Portsmouth has 1103 and Imperial has 1035 general and acute beds.(37)

4.3. **Workstream 1:** Understanding staffing decision support needs and applying user-centred design

This workstream (WS1) will capture nurses' and managers' **needs and views around staffing decision tools**, because they will be the end users directly affected by the algorithms we develop. The ultimate aim of this workstream is to start the process to implement the algorithm in existing software by engaging the relevant groups. WS1 will inform and be informed by results in WS2 and 3, and will run concurrently to the other workstreams.

WS1 will be guided by **user-centred design** principles, specifically human-centred design of artificial intelligence(38). The main point of user-centred design is "to ask what the goals and needs of the users are, what tools they need, what kind of tasks they wish to perform, and what methods they would prefer to use." (Norman & Draper, 1986). This will increase the chance that any algorithms developed as a result of this work will be fit for purpose and serve nurses' and managers' needs.(19) According to ISO 9241-210:2019 (Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems) we will follow these principles: 1) The design is based upon an explicit understanding of users, tasks and environments; 2) Users are involved throughout the design and development; 3)The design is driven and refined by user-centred evaluation; 4) The process is iterative; 5) The design addresses the whole user experience; 6) Design is performed by multidisciplinary teams and from multidisciplinary perspectives.

We will run a **survey and workshops** at two stages in the study: 1) to understand current practices and needs, as well as end-user perceptions of the existing processes for making staffing decisions 2) to identify opportunities that might be created by a routine data algorithm and/or constraints, including how timely and accurate estimates of staffing requirements need to be, how and when these estimates are used, and what variables should inform the algorithm. This has implications for the possible sources of patient data we might use, so it is important that end providers of software products and those managing the data are also involved as they will understand opportunities/constraints for linking up data from different sources. In the survey, we will sample staffing matrons and Chief Nursing Information Officers nationally, and in the early workshops, we will talk to nurses in charge of completing the SNCT assessments and nurse managers in charge of making staffing decisions. In the later workshops we will also include NHS IT managers and e-roster software product developers.

4.3.1. Sampling participants

Between 1 and 6 months into the study, we will launch a national online survey open to all **staffing matrons and Chief Nursing Information Officers (CNIOs)** through the Shelford Group, RLDatix user communities and NHS England Safe Staffing programmes, which we have links with. Staffing matrons have knowledge of how staffing decisions are made in their organisation while CNIOs have knowledge of IT systems.

Between 6-9 months into the study, we will run two workshops to understand current issues around decision support needs – one with **nurses** and one with **managers** who are involved in making staffing decisions. We will run separate groups for clinical nurses and nurse managers since they have different staffing responsibilities. A designated clinical nurse is in charge of monitoring staffing on their ward on a shift, while nursing managers (often safe staffing leads/matrons) have to balance staffing needs between wards and make business cases to finance directors. We aim to recruit between 5-10 clinical nurses through our study sites (at least one from each Trust and from a mix of relevant ward types) and 5 nursing managers (one from each Trust). Given the demands on nurses' time, this seems a realistic sample size to aim for. The identification of nurses' and managers' needs will guide the analyses carried out in WS2, at the end of which we will have an initial understanding of what the decision support tool(s) (based on our algorithms) might comprise of.

At approximately 18, 22 and 26 months into the project, we will run three more iterative workshops with **nurses, nurse managers, NHS IT managers and e-roster companies** (e.g. Oceansblue, RLDatix) to identify opportunities for and barriers to embedding the algorithms in existing software, as well as desired features. This will use algorithms developed and validated in WS2 and WS3. In total, we are aiming to recruit no more than 10 participants per workshop to ensure everyone has a fair chance to be heard and contribute meaningfully.

4.3.2. Data collection/ Workshops

Through the **survey,** we aim to understand which **staffing tools** are being used nationally, and how they are used, for example to set establishments, plan rosters and deploy staff on a daily and shift-by-shift basis. We also aim to understand the availability and quality of patient data and **IT systems** in use to further understand transferability of any proposed algorithms we develop. There is existing research (NIHR14/194/20) on NHS managers' use of staffing tools, but a national picture of how these tools are used (e.g. frequency of data collection and decision-making) is still missing. These survey findings will influence our decisions around variables to include in the algorithms, and what type of decision support algorithms to prioritise in WS2.

During the **first set of facilitated workshops**, participants will be prompted to discuss their use and perception of the existing SNCT system, in terms of workload, challenges faced, and trust in outputs. **Staffing decision support needs** will also be discussed, including staffing decision-making requirements, including timescale of decisions, and how an algorithm could support them, in terms of both the content of the output as well as its design or presentation method. This reflects the first phase of human-centred design - understanding and specifying the context of use, by defining users and stakeholders, their characteristics, goals, and tasks, as well as the overall environment of the system. This is a fundamental activity that should be undertaken before the design of any system or algorithm. An expert in user-centred design will facilitate the workshops (RI) and be supported by further qualitative expertise from CDO and MW. The workshops will be audio-recorded and the transcripts subjected to thematic analysis in line with Braun and Clarke.(39) The qualitative analysis of workshop results, including the themes identified and their prominence in the discussions, will inform how we conduct and **prioritise activities in WS2**. For example, if participants are most concerned with decision support for when to redeploy staff, we will focus more heavily on this, or if they want a way of tracking staffing requirements for the purpose of flagging when their establishment appears wrong, we will spend more time on addressing this aspect.

The **second set of workshops** will be focussed on **algorithm design** considerations. We will present the algorithms in an interactive manner with real-life scenarios to demonstrate how the algorithm could be used in decision support. We will ask for feedback on the usefulness and feasibility of the algorithms and use measurement tools such as the Van Der Laan Acceptance Scale(40) and the System Usability Scale(41). We will then develop the algorithms further before the next workshop. Between each workshop, participants will be given the opportunity to reflect on the algorithms and provide qualitative insight into usability, and potential improvements therein, will be gathered via short questionnaires comprising open-ended questions. We will also ask nurses and managers about the workload implications of the algorithms; electronic systems are often introduced with the assumption that they will improve workforce efficiency, but this is not always realised(32). Since our goal is to release nurses' time, it is imperative that workload implications of changes to existing tools are considered.

Building on this work, we will apply for further funding, for example a **Knowledge Transfer Partnership**, to work with software providers and end users to build prototype decision support tools based on our algorithms.

4.4. Workstream 2: Developing measures of nurse staffing requirements based on routine data

This workstream aims to **develop and test algorithms** to estimate nurse staffing requirements from routinely available patient data. Since there is no "gold standard" for measuring nurse staffing requirements, in the first instance we will aim to replicate measurements using the NICE-recommended Safer Nursing Care Tool, which is in use in most English hospitals. This tool bases its estimates on a categorisation of patients on the ward into acuity/dependency groupings, with estimates of direct care time for these patients as well as allowances for other work associated with patients (indirect care) and other administrative tasks. Adjustments to account for differences in this allowance are currently left to professional judgement. We will extend our approach to attempt to capture some of these workload aspects beyond acuity/dependency, while recognising that not all elements of work are measurable.

We will develop two broad approaches using data that is already recorded to replicate the results of SNCT assessments: i)replicate **individual** patient acuity/dependency classifications and ii) replicate the staffing requirements for a **whole ward**. For i), where Trusts record individual SNCT ratings we will seek to develop models based on rich individual data from electronic patient records and care plans. However, given that staffing decisions are generally made based on the whole ward as nurses care for several patients each, we will also aggregate these to a ward level (ii). It is possible that inaccuracies in individual categorisations may have little impact on overall estimates when aggregated to a ward level. Furthermore, some Trusts record only counts of patients per SNCT category, which cannot be linked to individuals, so for these Trusts we can only follow approach ii) and estimate ward staffing requirements directly.

We will consider staffing decisions at **different timepoints**, such as daily deployment and redeployment decisions and twice-yearly establishment review (as specified in NICE guidance). The priorities will be determined based on the national survey and user needs identified in the first set of workshops in workstream 1. Some information is not known in real-time, so we will use different predictor variables depending on the timing of the decision. We will assess the predictive accuracy of the algorithms we derive, as well as assessing the feasibility of applying algorithms to datasets from other hospitals and test their external validity.

4.4.1. Data sources

The **data sources** are a range of hospital systems which will differ by Trust. The Safer Nursing Care Tool data that is collected twice-yearly for establishment reviews will be sourced from the relevant spreadsheets/Trust records, and where available the SNCT data that is recorded for deployment decisions will be extracted from the electronic roster (RLDatix SafeCare system). Patient data will be sourced from various IT systems storing patient administration records, nursing care documentation, care plans, pharmacy and surgery records. We will supplement this with contextual information about wards sourced directly from our Trust nursing contacts.

4.4.2. Sample sizes

For **developing the algorithms**, we will use retrospectively collected data from our two partner Trusts. Since the new version of the SNCT launched in October 2023, we will extract data from this date onwards or from when the new tool was embedded in the Trusts. We aim to collect at least one year of data. This covers two staffing establishment reviews where acuity/dependency scores are recorded for at least 30 days, so there will be a minimum of 60 (30 x 2) days' data per ward, although in many wards they collect data much more frequently giving about 730 (2 times per day x 365) data points per ward. Scaled up across wards and two Trusts this will give us in the region of 4,800-58,400 ward shift data points for the ward-level approach (assuming 40 wards per Trust), and approximately 120,000-1,460,000 patient shift data points for the patient-level approach (assuming on average 25 patients per ward). For **external validation** of the algorithms, we will use data from at least three further Trusts, and more recent data from our two partner Trusts (data covering at least 1 year).

4.4.3. Variables

Our **outcome variables**, which we are trying to predict, are all based on the Safer Nursing Care Tool assessment of staffing requirements. This is the overall nurse staffing requirement for a ward, which will include registered nurses, nursing associates and support workers (although note job titles vary between wards and Trusts), but will exclude managers who are not providing any direct care to patients. For individual patient predictions of staffing needs, the outcome variable is the acuity/dependency category (or the related workload multiplier). This is a categorical variable with 7 possible values. For ward-level predictions of staffing needs, the outcome variable is whole-time equivalents (WTE) per patient. This is calculated by multiplying the number of patients in each category with the corresponding multiplier, adding together and dividing by the number of patients. This can be calculated for each shift or time period that the SNCT data is recorded. There are different multipliers dependent on ward type and ward configuration (e.g. separate multipliers for admissions units and wards with all single rooms). For deployment decisions we can remove the uplift (for leave) and convert WTE to care hours on a day/shift, to calculate a care hours per patient measure.(7)

In terms of **predictor variables**, we will derive a set of variables that are likely to relate to patient acuity/dependency categorisations as assessed by the SNCT. We will be guided by the SNCT acuity/dependency level descriptions alongside expert judgement (JB, SW, SO) to decide which characteristics are important to consider. Routinely available data in the patient electronic record includes demographics, information about the patient journey (admission, ward transfers, discharge) and diagnostic/procedure information. Where available, we will consider variables from care plans and care documentation. These include records of infusions, oxygen, specialing needs (one-to-one supervision), observations, and moving and handling assessments. JB has already mapped variables from care documentation onto acuity/dependency categories, so we will build on this work. We will also investigate the possibility and feasibility of including variables from pharmacy and surgery datasets such as numbers of medicines, anaesthesia type/time and surgery times. For the **ward-level approach,** either we will predict the individual's acuity/dependency category and then aggregate into a ward-level prediction, or where individual ratings are not available, we will transform patient-level variables into ward-level variables e.g. proportion of patients who are male, average age, etc.

Depending on which staffing decision our algorithm is supporting, we will consider the **timeliness** of data, and correspondingly, different variables will be used. For example, establishment reviews can make use of retrospective data, while decisions about redeployments at the start of a shift can only use information available in real-time. We will carefully consider, informed by discussion with our nursing and IT system team members (SW, TW, JB), which variables are likely to be available in near real-time versus retrospectively. Where possible we will obtain timestamps for when data were recorded to help with this. Some variables will change over the course of a patient's stay while others are fixed. Relevant information that is known at admission includes age group, gender, method of admission, long-term conditions and admission diagnosis. Information that changes over the course of the stay includes vital signs, treatments and therapies, length of stay, number of transfers and complications. Further variables that are available retrospectively include procedure codes (HRGs), discharge destination and overall length of stay. Information from prior admissions for the same patient could also be incorporated.

Based on discussions with nurses in WS1, we will explore whether we can **extend** our algorithms to address **skill mix**, if this is deemed an important issue. While the SNCT does not provide a recommendation for the skill mix, we will attempt to predict how many Registered Nurses versus how many Support Workers are needed. For this, as a starting point, we will look into segmenting our predictions into acuity versus dependency, with a starting assumption that wards/patients with higher acuity need a higher proportion of Registered Nurse time and wards/patients with higher dependency need a higher proportion of Support Worker time. We recognise the limitations of this assumption given that there are elements within the dependency categories that require Registered Nurses such as delivery of complex infusions, wound management, end-of-life care and complex discharge planning. We will also factor in Nursing Associates, in a subgroup analysis for wards where there is enough data to draw meaningful conclusions.

We will extend our measure to incorporate other aspects of nursing workload that are **unrelated to acuity or dependency**. This includes both those workload drivers that are currently left to "professional judgement" when using the SNCT, e.g. high rates of admissions/discharges or unusual ward layouts(42) as well as those that are captured by distinct sets of SNCT multipliers (e.g. Acute Admission Unit multipliers, single room multipliers).(43) Here we will work closely with our nursing co-Investigators and advisors, and review available literature to determine which candidate variables to include.(44, 45) We will make use of contextual ward information from ward managers to understand ward layouts including numbers of funded beds, rooms and single rooms.

4.4.4. Data analysis

The **data linkages** between datasets are complex but our team have prior experience with this (PM, CS, CDO, PG). Data will be processed and linked differently according to approach (patient- versus ward-level, establishment versus deployment decisions). For the patient-level approach, for example, all patient information pertaining to one admission will be linked together, and linked with the acuity/dependency category of the patient (repeated for each time the patient is assessed on the ward). Then these predictions will be aggregated to a ward level. In contrast, for the direct ward-level approach, the variables derived are case-mix variables. This will involve identifying which patients are staying on the ward and linking information about those patients together. To calculate the Safer Nursing Care Tool measure, for each assessment we will calculate the ward staffing requirement using the SNCT multipliers and the counts of patients per acuity/dependency category, and translate this into a per patient measure, whole-time-equivalents per patient (for establishment) or care hours per patient (for deployment). We will consider alternative approaches for how to deal with changing patient information, e.g. where a particular care activity is not recorded every day (**missing data**), whether to assume it is still needed or not.

In **machine learning** terminology, predicting individual acuity/dependency categories is a classification problem, since there are only limited choices, while predicting ward-level staffing requirements is a prediction problem, since the outcome is continuous. Thus different techniques are likely to be appropriate for these two problems. For the individual-level classification problem, we will begin by implementing ordered logistic regression. Additionally, we will explore using decision trees, random forests(46) and support vector machines, which are well-suited for handling categorical data with a limited number of outcomes. For ward-level regression prediction, we will start with simple approaches like linear regression before extending to other approaches if we meet challenges. We will consider GAM models,(47) random forests, support vector regression, gradient boosting, neural networks, etc., based on the suitability of the data.(48)

We will compare the **predictive performance** of these algorithms and use a split-sample (training/test dataset) approach, which avoids problems with overfitting (when a model is excessively complex, capturing noise rather than the underlying pattern in the training data).(49) By testing our model on a separate dataset, we ensure that our model generalises well beyond the training data. In the first instance, we are aiming to replicate the Safer Nursing Care Tool estimates of staffing requirements as closely as possible. Thus, we will assess how close our estimates/predictions are to the SNCT ones using Bland-Altman limits of agreement(50) and statistics such as root mean squared error. We will calculate these statistics for both the training and test datasets allowing us to monitor any discrepancies between training accuracy and real-world applicability. We will work with our nurse representatives to understand what "good enough" looks like and determine suitable thresholds. Other measures such as concordance and comparing ward rankings (i.e. highest to lowest estimated staffing requirement) between measures will also be considered. For establishment-setting decisions, we will take 30-day (the sample size recommended in current SNCT guidance) samples from the observed and predicted values and compare them. If identified as a priority in WS1, we will also look into whether tracking how ward acuity/dependency levels change over time using **statistical process control** charts could help identify when establishment reviews are needed.

We will assess if the algorithms fitted for the original hospitals are effective at predicting nurse staffing requirements more generally. For this we will perform **external validation**, assessing how well the algorithms perform on data from other hospital Trusts and from different time periods. Following development of algorithms for our partner Trusts, we will assess whether the same or similar variables can be derived from at least three other hospital Trusts' systems. We will learn to what extent it is necessary to adapt the method of derivations of variables for different Trusts or potentially remove some variables that cannot be calculated. Then we will test the performance of the algorithms on data from these other Trusts and again assess agreement with the SNCT estimates. This will highlight potential problems with the external validity and transferability of the algorithms. Similarly, we will assess the performance of the algorithms on more recent data from the same Trusts. Other measures of validity will be considered in workstream 3.

4.5. **Workstream 3**: **Validating measures of nurse staffing requirements against patient outcomes and nurse perceptions of inadequate staffing**

We will conduct a retrospective observational study to investigate the associations between actual under/over-staffing relative to each candidate measure of staffing requirements and a range of outcomes. One would expect that as staffing increases relative to a measure, the risk of poor patient outcomes and perceptions that staffing is inadequate decreases.(13, 26) In the absence of a gold standard, we will compare model fit against models with staffing requirements measured by the SNCT, as well as against the current ward establishments where possible. For this we will use routine data extracted from hospital IT systems along with an empirical (prospective) micro-survey of nurses (three tick-boxes) to understand perceptions of staffing adequacy.

4.5.1. Data sources

In addition to the **data sources** described in 4.4.1 above, for this workstream we will run a micro-survey of nurses consisting of three questions added to the existing Safer Nursing Care Tool data collection, as well as extracting roster data extracts.

Clinical nurses will answer this tick-box **micro-survey** at the same time that they are already recording the patient acuity/dependency levels, using existing data collection methods, e.g. functionality that is available in the roster systems. We developed this micro-survey for our previous study evaluating the Safer Nursing Care Tool (HS&DR 14/194/21), and it was viable with achieved completion rates of at least 85% per Trust.(13) In this survey we will assess perceived staffing adequacy, which although subjective, has been found to be strongly associated with patient, nurse and organisational outcomes.(51) A strength of these subjective assessments is that they include professional judgement for those aspects of workload that are hard or impossible to measure.(51)

Roster data is required to determine actual staffing, for determination of under/over-staffing compared to each measure of staffing requirements. We will extract both rosters for substantive staff as well as bank/agency rosters if held separately. Where possible we will also source data on the staffing establishment for each ward or estimate these from the roster data.

Patient outcomes will be sourced from **patient administrative systems**. We will consider adverse outcomes identified through diagnosis codes and other administrative data; in our previous study we found that incidents such as deep-vein thrombosis and pressure ulcers identified from diagnosis codes are associated with adverse staffing levels (NIHR 128056 under review). We will not use incident reports because in previous studies), it has been found that incident reports, e.g. from "datix" data, is of poor quality with higher staffed wards tending to record more incidents, possibly due to having more time for reporting.(26, NIHR 128056 under review)

Sample sizes

To estimate the sample size for the **perceived staffing adequacy outcomes**, we will assume a similar response rate (85%) as for our previous study. We are aiming to run the micro-survey for at least 9 months including a piloting/training phase of 3 months. We will aim for all eligible wards to complete the micro-survey but will work with Trusts as to what is feasible. In Trusts where SNCT data is only collected during establishment reviews, we will aim for the survey to cover 2 establishment reviews. Based on 20 wards per Trust (average number of wards with micro-survey data from previous study(13)), this would give us between about 1200 (30 day sample x 2 reviews x 20 wards) and 4653 (0.85 completion rate x 9 months x 20 wards) ward days of data per Trust for understanding the relationship between understaffing and perceived staffing adequacy. In total across Trusts this would give us between 6,000-23,265 data points. Based on 626,313 admissions in our previous study corresponding to 5 years' of data from 4 Trusts (NIHR 128056 under review), we estimate that across 5 Trusts in a 9-month period we would have in the region of 100,000 patient admissions for modelling the relationship between understaffing and **patient outcomes**.

4.5.2. Variables

We will consider a range of **patient outcome variables**. Our primary patient outcome is mortality within 30 days of admission, as it has a strong association with staffing levels.(5, 17) Other patient outcomes we will model include length of stay, readmissions within 30 days (which both have high strength of evidence for an association with staffing levels according to a recent review)(4) and healthcare-associated conditions such as pneumonia and pressure ulcers (which have moderate evidence).(4) We will limit the latter analyses to surgical patients since for medical patients it is harder to distinguish between complications and conditions present on admission.

Our other outcomes are measures of **perceived staffing adequacy** assessed through the prospective micro-survey**.** We will ask clinical nurses three questions as in our previous study: i) Were there enough staff for quality?, ii) Were breaks missed? and iii)Was care left undone?(13) These were derived following pilot work of a more extensive set of questions, and are aimed to be quick to answer and give a good indication of overall staffing adequacy on a ward. In our previous study these were all associated with deviation from SNCT-recommended staffing.(13)

Our **independent variables** are the deviation of actual staffing from measures of staffing required. In other words, under-/over-staffing compared to different measures of what is needed. Actual (achieved) staffing is measured as total care hours per patient day and includes all nursing staff on the ward: both Registered Nurses and Support Workers, both permanent and temporary (bank or agency) staff. Required staffing will be measured in different ways including: using the SNCT, current ward establishments, algorithmic measures (including the ward-level and patient-level approaches).

To start with we will model overall staffing but where possible will also consider **skill mix** by splitting out staff groups. We will use current ward establishments to calculate the planned skill mix and apply these to our alternative measures of staffing requirements. This will allow us, for example, to assess Registered Nurse understaffing and Support Worker understaffing separately, which is important since differing relationships have been found with patient outcomes.(5) We will use the pay band to categorise staff, given the differences in roles and job titles between Trusts.

4.5.3. Data analysis

This workstream will involve further **data linkages**. For each ward and shift or 24-hour period, we will calculate the actual staffing and the staffing requirements according to the most promising algorithms, as well as according to the SNCT. These will be converted to care hours per patient day. This requires working out which staff shifts and which patient stays overlap with each time period, as well as calculating the staffing requirements based on patients on the ward in each period. Then for each patient admission, we will link the understaffing/overstaffing they experienced during their hospital stay with their outcomes. Separately for each period we will link ward understaffing/overstaffing with measures of perceived staffing adequacy (from the micro-survey).

We will use **regression modelling** to assess the relationship between deviation from staffing requirement (measured in different ways) and outcomes. Depending on the nature of the outcome variable we will fit the appropriate model type (e.g. survival model for mortality, logistic regression for hospital-acquired conditions and perceived staffing adequacy). Our longitudinal data will allow us to explore how staffing over a patient's stay affects outcomes, since effects of low staffing may accumulate; we will consider approaches such as moving averages and exponentially weighted moving averages.

We will assess intra-cluster correlation coefficients to see whether there are unmeasured **differences between wards**. If appropriate, we will use multi-level regression models to assess how relationships differ between wards and hospitals, which can be used as an indicator of good/poor fit of a staffing measure, as in our previous work.(52) Specifically, random intercepts models can help us estimate, for each ward, the probability of the nurse in charge reporting enough staff for quality when staffing matches the measured staffing requirement. On the other hand, random slopes models can indicate for which wards relationships are in the same/opposite direction to what is expected.

We will compare how well different measures of staffing requirements appear to capture workload according to Akaike Information Criterion **model fit.** We will also compare the benefits of using alternative staffing requirement measures by using decision analytic methods like **decision analytic curves**, as deployed in research about the Rafaela staffing tool.(26) The advantage of this method is that it provides an estimate of net benefit without requiring additional data on costs or quality-adjusted life years.(53)

We will investigate the impact of changing our modelling approach and assumptions through **sensitivity analyses**, for example varying the exposure window and fitting alternative model types. We will also run subgroup analyses to understand the fit for particular patient groups.

4.5.4. Equality, Diversity & Inclusion

Equality, Diversity and Inclusion are key aspects that we will consider in each workstream. In particular, we will consider these when engaging staff and patients in workshops, and when developing and validating algorithms. We will gather information about protected characteristics from participants in workstream 1 workshops through an optional online anonymous form, with the results stored in an access-restricted secure electronic file-store area. For workstreams 2 and 3, since we are using routinely collected data, we will have access to information to derive only some of the protected characteristics (sex, age band and partial information on disability and pregnancy/maternity).

When working with staff and patients in workshops (workstream 1 and PPI), our main goal is to **remove barriers to access** to achieve meaningful engagement, and we plan to do so by:

1) Organising workshops in a choice of online or in-person formats, following best practice to ensure all participants get equal opportunities to co-create and give feedback: i) Accessing relevant technology e.g. “owl” microphones/cameras for those attending in-person; ii) Avoiding running meetings for longer than 90 minutes, iii) If group work is required, ensuring participants online can participate meaningfully in a breakout room and report back; iv) Ensuring all workshop materials and agenda are available in advance to all, especially online participants.
2) Planning workshops at times of the day that might fit our intended participants – e.g. active members of the nursing workforce might not be available during standard work hours. We will also reimburse patient-facing nurses for their time since they do not have research time in their role. We will also consider childcare commitments and pay for childcare to allow participants to attend. Similarly we will cover costs of carers or personal assistants in line with NIHR payment guidance.
3) Actively recruiting people from a variety of backgrounds e.g. ages, ethnicities and socio-economic backgrounds. We will do this by: emphasising that we want views from a diverse group of individuals when we advertise and potentially targeting particular groups who tend to be underrepresented, through support groups such as Equality 4 Black Nurses or the Imperial Health Charity, or cultural groups such as the African Women’s Forum in Portsmouth.

In workstreams 2 and 3 we are analysing data on all patients staying on inpatient wards, so there is no bias introduced from recruitment methods. However, there are other EDI considerations, particularly around potential **algorithm biases** which can arise at different points in the algorithm building cycle, for example when models are trained on unrepresentative data sets(54). We will consider potential biases at each stage of the algorithm building cycle.

For example, we will minimise the risk of **representation bias** arising by using data from hospitals serving diverse populations. Although in our datasets we will not have access to data on ethnicity or socio-economic status, so cannot specifically check how well algorithms work for particular groups, we aim for our datasets to cover diverse populations, for example the patient population served by Imperial hospital serves a relatively high proportion of people from ethnic minorities(55) and Portsmouth has a relatively high proportion of people of low socioeconomic status.(56) However, as well as making efforts to include data from under-served groups in our analysis, we shall specifically examine the performance of our algorithms in wards with high numbers of patients from under-served groups such as the over 75s, those with learning disabilities and those with mental health conditions. If the performance (predictive accuracy or validity when assessed against patient/staffing adequacy outcomes) is poorer in these wards then we shall investigate if there are important variables missing that could affect nurse staffing requirements. To understand this, we will engage with our patient and nurse representatives.

5. Dissemination, Outputs and anticipated impact

For an **academic audience**, we intend to produce at least three journal articles from our research, as well as presenting at two conferences. We will write at least one paper per workstream, covering: i) staffing decision support needs, ii)estimating staffing requirements from routine data and iii) validity of measures of staffing requirements. We will target relevant journals with open access options such as the International Journal of Nursing Studies, BMC Health Services Research and BMJ Open. We will present at a conference early in the project to gain feedback on preliminary work and plans (target is the Royal College of Nursing Conference), as well as sharing preliminary at another conference (target is the Health Services Research UK conference).

Throughout the study we will inform and engage **patients and NHS staff** about our work. We will share our results directly with the intended users of the tools we develop through workshops (see workstream 1) and existing links to the NHS Safe Staffing fellowship programme, RLDatix user group and the Shelford Group. We will attend meetings of such user groups to engage them about our work at key stages of the study. Since beneficiaries of the research are clinical nurses, nursing managers and patients, we will produce targeted resources for these groups. These will be summaries of our research, as we have done previously, in more accessible formats e.g. podcasts, blogs or videos(57) for patients, and Evidence Briefs published in the Nursing Times for nurses.(58, 59)

We are maximising the chance that our algorithms and the knowledge derived in this project will **enter the NHS** through early engagement with key stakeholders, as well as ongoing dialogue through an advisory board. We are engaging safe staffing leads and CNIOs through a national survey early in the study and will give them the option to receive study updates. We are engaging roster companies, with whom we have existing links, in the workstream 1 workshops to identify opportunities to embed algorithms in their systems. We have preliminary support from the Shelford Group who developed the Safer Nursing Care Tool, and are engaging with them as to how to reach Trusts with any new validated algorithm. Through our user-centred design process, we will design algorithms taking account of user needs, and identify opportunities for implementation as well as potential barriers within this project (WS1).

If this research is successful, and we are able to develop algorithms that are promising in terms of predictive validity, external validity and correspondence with patient and staffing adequacy outcomes, we will apply for **further funding**. This will likely take the form of a Knowledge Transfer Partnership e.g. through Innovate UK (UKRI), which would allow us to work with an industrial partner (likely RLDatix or Oceansblue who work in this space) on the next step of developing software prototypes that make use of our algorithm as decision support within existing roster systems.

Possible **barriers** for adoption and implementation of our algorithms include mismatch between algorithms and decision support needs, incompatibility with data systems, distrust in the algorithms. We aim to understand and mitigate against these barriers as well as identifying further barriers and solutions through our national survey and workshops with nurses, NHS IT managers and roster companies in workstream 1. We have applied for an SNCT researcher license to use the SNCT in our research; this IP is owned by Imperial Innovations. We have engaged the SNCT committee of the Shelford Group who will be reviewing our proposal; the chair has given preliminary approval of our study.

The expected **impacts** of this research are ultimately a better match between staffing and workload on hospital wards. Embedding algorithms to predict staffing requirements in roster systems would enable real-time monitoring of variation in demand, facilitate more efficient deployment of scarce resources and save nurses time in recording patient acuity/dependency classifications. Thus nursing staff working on hospital wards would benefit through better working conditions and reduced workload. Using the algorithm(s) to support better staffing decisions would likely benefit patients staying in hospital through safer and higher quality care and greater access to nurses, as well as patients on waiting lists through quicker admission to hospitals. More immediate impacts are greater knowledge about the usefulness of routine data for predicting nurse staffing requirements on hospital wards, greater understanding of nurses' staffing decision support needs and the required functionality of tools to help managers with planning staffing, which can feed into software product development.

We will **share** the progress and findings with nursing staff in the participating hospital sites through presentations (for example at standing meetings) and e-newsletter updates (also available to national survey participants). We will share the research with patients/the public by liaising with PALS services and will suggest displaying posters in the participating hospitals as well as uploading information on the hospital websites if possible.

6. Project / research timetable

Category	What	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
		01-Aug-24	01-Sep-24	01-Oct-24	01-Nov-24	01-Dec-24	01-Jan-25	01-Feb-25	01-Mar-25	01-Apr-25	01-May-25	01-Jun-25	01-Jul-25	01-Aug-25	01-Sep-25	01-Oct-25	01-Nov-25	01-Dec-25	01-Jan-26	01-Feb-26	01-Mar-26	01-Apr-26	01-May-26	01-Jun-26	01-Jul-26	01-Aug-26	01-Sep-26	01-Oct-26	01-Nov-26	01-Dec-26	01-Jan-27	01-Feb-27	01-Mar-27	01-Apr-27	01-May-27	01-Jun-27	01-Jul-27	01-Aug-27	01-Sep-27	01-Oct-27
Pre-study	Ethics, R&D and data governance approvals	■	■	■	■	■	■	■	■	■																														
	Research Fellow recruitment							■	■	■																														
	Additional Trust recruitment (3+)	■	■	■																																				
Workstream 1	National survey										■	■	■	■	■	■																								
	Analysing survey results																■	■	■																					
	Workshops (decision support needs)															■	■	■	■																					
	Thematic analysis of workshop results																			■	■	■	■																	
	Workshops (algorithm design considerations)																											■				■				■				
	Paper on staffing decision support needs																							■	■	■														
Workstream 2	Data extraction 1 (2 Trusts)										■	■																												
	Data linkage and processing for algorithms												■	■	■	■	■	■																						
	Fitting algorithms - ward-level approach																		■	■	■	■	■	■																
	Fitting algorithms - patient-level approach																		■	■	■	■	■	■																
	Fitting algorithms (extensions)																								■	■	■	■	■	■	■	■	■	■						
	Testing on other datasets																												■	■	■	■	■	■						
	Paper on estimating staffing requirements from routine data																																		■	■	■			
Workstream 3	Micro-survey set-up and training and pilot										■	■	■																											
	Micro-survey running													■	■	■	■	■	■	■	■	■																		
	Data extraction 2 (All Trusts)																						■	■	■															
	Data linkage and processing																								■	■	■	■												
	Regression modelling of patient outcomes																												■	■	■	■								
	Regression modelling of staffing adequacy																																■	■						
	Paper on validity of measures of staffing requirements																																		■	■	■			
PPIE	Recruiting patients to PPI group										■	■																												
	Workshops with patients/families												■	■	■	■																		■	■	■				
	Workshop with college students																							■	■															
	Science and Engineering fair																				■																			
	Podcast /blogs/videos															■	■											■	■							■	■			
	Posters for hospitals																																					■	■	
	Engage with user groups												■	■	■												■	■	■									■	■	■
	E-newsletter										■											■												■						■
Dissemination	Conferences														■										■															
	Write final report																																					■	■	■
	Magazine article for nurses																																					■		
Advisory group meetings												■						■						■						■						■				

7. Project management

The Principal Investigator (CS) will provide study oversight and have overall responsibility for project success, with mentoring support provided by PG, an experienced NIHR senior Investigator. Given the study complexity, we will recruit a project coordinator to support with study administration and communication activities, including tracking progress in obtaining approvals/data extraction for each Trust. The PI will regularly review budgets and timelines, as well as maintaining a risk register to track risks to project success and agreed mitigations. The individual workstreams will be managed day-to-day by named leads (CDO, PM) who will be in charge of prioritisation within workstream activities to complete on time. There will be regular weekly meetings of the core team (CS, PG, CDO, PM, Research Fellow) for analysts to share progress and for colleagues to provide expertise and support. Ad-hoc focussed meetings with other co-Is will be arranged as needed when their expertise or support is required. There will be quarterly meetings of all co-Investigators to discuss progress and plans. An advisory board, consisting of people with a range of relevant expertise or experience, will meet twice per year to provide guidance to the team members.

8. Ethics

Workstream 1 involves **staff as participants**, and workstreams 2 and 3 use **de-identified patient data.** The development and external validation of algorithms predicting individual patient acuity/dependency ratings requires granular patient data to be linked at an individual level so this part of the work will be completed within secure data environments. We will submit ethics applications to the HRA and to our university ERGO ethics system. We will begin the ethics approval process at an early stage, aiming for approvals to be in place by the start of the project.

Data will be pseudonymised by the Trusts' data analysts, who will destroy the key so there is no mechanism to **re-identify** using the original patient identifiers. Once the data leave the Trust, the pseudonymous identifiers could not be used to identify the source records by anyone outside. The large volume of records makes inadvertent identification of an individual by a member of the research team very unlikely. It could only occur if most or all of the information contained in the record were known to them already. The consequences of reidentification are unlikely to be harmful; nonetheless we will further reduce the risk for reidentification by aggregating variables (e.g. age groups rather than dates of birth), and restricting access to data to a small number of named individuals subject to contracts requiring adherence to confidentiality policies and proper treatment of the data. We will suppress small numbers when reporting results.

There are also ethical issues around **potential misuse of an algorithm**, which we will try to reduce as much as possible. We are involving the relevant groups in the design process to help identify risks of misuse and encourage an appropriate use; we are advocates for the use of measurement tools/algorithms to be starting points to be questioned rather than providing a number to be wholly relied upon, as demonstrated by our work developing a Professional Judgement Framework.(42) We are trying to improve on the current process of determining nurse staffing levels and believe that an algorithm could help automate part of this process and provide more reliable acuity/dependency measurements based on data that is already recorded routinely.

9. Project / research expertise

We are an **interdisciplinary research team** including academics with expertise in both the topic area and planned methodologies, as well as staff based in Trusts who have first-hand experience of nurse staffing processes. Our team has experience leading research groups and programmes of work (PG, MW, EM) as well as leading research studies (CDO, CS, PM). We have a long track record in safe staffing research (PM, PG, CDO, CS), including specific expertise in staffing tools. The team has strong data science skills including particular expertise in data linkage and database design (PM, CS), hierarchical statistical regression (PM, PG, CDO, CS), machine learning (JH, EM), computer science (PM), statistics (JH) and operational research (CS). We also have researchers with expertise in qualitative methods (CDO, RM, MW). The research fellow will be supported by regular meetings with the team and more frequent ad-hoc support from PM. In accordance with the Researcher Concordat, we have planned for at least 10 days of training each for the research fellow and PM, including costing for training courses and conferences.

Involving nurses and patients in the development of the algorithms is of the upmost importance to ensure they are useful and usable. **User-centred design** expertise and facilitation skills will be provided by RM. ID is our **PPI** representative and advisor, and CDO will lead the PPI work, as she has previous relevant experience including as the PPI link person for NIHR ARC Wessex. The **nursing perspective** on safe staffing and staffing tools will be provided by SW and JB; SW is safe staffing lead and within this role developed the Trust policy around nurse staffing so has great insight into the current processes, while JB is a CNIO with a deep understanding of care documentation and how this maps onto acuity/dependency. SO is Clinical Workforce Lead of the Chief Nursing Officer (CNO) Safer Staffing Faculty at NHS England and will advise on safe staffing and the Safer Nursing Care Tool, as well as supporting Trusts in nurse micro-survey set-up and training. TW in his role as Workforce Transformation Manager – Systems will provide insight into the **IT systems** and facilitate data extraction at Portsmouth. EM as Director of the iCARE Digital Collaboration Space at Imperial will facilitate work in the secure data environment and MW will facilitate access to data and systems, and advise on qualitative methods (WS1).

In addition to the co-Investigators, we have a **wider group of collaborators** who have expressed an interest in advising us on this research: Ann Casey (Senior Clinical Workforce Lead at NHS England and Head of CNO Safer Staffing Faculty), Christopher Morley (Shelford Chief Nurse representative), Andrew Worthington (Deputy Chief Nurse, Imperial College Healthcare NHS Trust), Nuno Pires (Lead nurse safe staffing, Imperial), Liz Rix (Chief Nurse, Portsmouth Hospitals University NHS Trust) and Karen Swinson (RLDatix). We will aim to recruit at least one clinical nurse onto the advisory group through our nurse co-applicants or study sites, although recognise the barriers to them playing an ongoing part in the study, given lack of research and development time in these roles. This group represent a variety of roles/positions so will give us a mix of perspectives on current use of the SNCT, factors affecting staffing requirements and decision support needs. We will extend our advisory group using our connections to e.g. nursing directorates in hospitals, roster companies and unions.

We will access specialist support as required, e.g. support with setting up contracts (Legal services team at University of Southampton), applying for ethics and governance approvals (Research and Innovation Services at University of Southampton) and implementation science support (Health Innovation Wessex).

10. Success criteria and barriers to proposed work

Success criteria	Potential barrier and risk mitigations
Approvals processes (Ethics, R&D and data governance) completed for at least five hospital Trusts	Barrier- Approvals processes are non-standardised and can vary in terms of requirements and timescales between Trusts. Mitigations- Our research team has previous experience with navigating these processes which may help speed this up. We will begin approvals processes at an early stage before the start of the study. We are planning a phased approach to the work by beginning developing algorithms on data from our two partner Trusts before external validation on data from all participating Trusts. We will over-recruit sites to allow for drop-offs.
Approval granted for set-up and use of a secure data environment	Barrier – Approvals process for use of secure data environments is another step with an unclear timescale. Mitigations – We have already started engaging with the secure data environment provider at our partner Trust where this is relevant. In the worst case, if approvals are not granted, the part of the work which does not require a secure data environment could still be completed.
High completion rate of micro-survey for assessing perceived staffing adequacy	Barrier – Clinical nurses may lack time or incentive to complete this. Mitigations – The micro-survey is very short (three tick-boxes) and is asking questions that nurses in charge would be considering as part of their routine work. We will provide training and run a pilot phase of the survey to spot any issues and address these, e.g. we will explain to nurses how our study has the potential to reduce work for them in the longer run.
Data extraction provided in a timely manner	Barrier - delays in data extraction, complexity of de-identification required or wrong data sent. Mitigations – We will provide detailed data specifications, communicate with data providers at an early stage to discuss the requirements and ask for sample datasets to check suitability before requesting the datasets for the full period. We will pay Trusts for the data on receipt to incentivise them to provide it. In the worst case, if we do not receive suitable data from one Trust in time, we will have data from the other Trusts so this will not prevent project completion.
Successful preparation of datasets	Barrier - Complexity of data linkage and pre-processing Mitigations - Similar work successfully completed by team in previous projects. We have supervision arrangements in place for the research fellow who will be well-supported.
Development of algorithms with high predictive accuracy and validity	Barrier – Routinely collected patient data may not be suitable for this purpose. However null results are still findings. Mitigations- We are requesting a wide range of data that is likely to pick up similar issues to the SNCT (regarding patient acuity and dependency) and more factors that the SNCT does not consider.
Algorithm implemented by software companies and Trusts (longer-term goal)	Barriers – mismatch between algorithms and decision support needs, incompatibility with data systems, distrust in the algorithms. Mitigations – Our user-centred design approach enables us to learn from potential end users and address these issues in the development of the algorithms rather than only finding out afterwards. We are engaging software companies at an early stage to understand opportunities and barriers to embedding algorithms in software.

References

1. Saville CE, Griffiths P, Ball JE, Monks T. How many nurses do we need? A review and discussion of operational research techniques applied to nurse staffing. International Journal of Nursing Studies. 2019;97(September):7-13. https://doi.org/10.1016/j.ijnurstu.2019.04.015.
2. Musy SN, Endrich O, Leichtle AB, Griffiths P, Nakas CT, Simon M. Longitudinal Study of the Variation in Patient Turnover and Patient-to-Nurse Ratio: Descriptive Analysis of a Swiss University Hospital. J Med Internet Res. 2020;22(4):e15554. https://doi.org/10.2196/15554.
3. Shin S, Park JH, Bae SH. Nurse staffing and hospital-acquired conditions: A systematic review. J Clin Nurs. 2019;28(23-24):4264-75. https://doi.org/10.1111/jocn.15046.
4. Blume KS, Dietermann K, Kirchner-Heklau U, Winter V, Fleischer S, Kreidl LM, et al. Staffing levels and nursing-sensitive patient outcomes: Umbrella review and qualitative study. Health Serv Res. 2021;56(5):885-907. https://doi.org/10.1111/1475-6773.13647
5. Dall'Ora C, Saville C, Rubbo B, Turner L, Jones J, Griffiths P. Nurse staffing levels and patient outcomes: A systematic review of longitudinal studies. Int J Nurs Stud. 2022;134:104311. https://doi.org/10.1016/j.ijnurstu.2022.104311.
6. Griffiths P, Recio-Saucedo A, Dall'Ora C, Briggs J, Maruotti A, Meredith P, et al. The association between nurse staffing and omissions in nursing care: A systematic review. Journal of Advanced Nursing. 2018;74(7):1474-87. https://doi.org/10.1111/jan.13564.
7. Griffiths P, Saville C, Ball JE, Chable R, Dimech A, Jones J, et al. The Safer Nursing Care Tool as a guide to nurse staffing requirements on hospital wards: observational and modelling study. Health Services and Delivery Research. 2020;8(16). https://doi.org/10.3310/hsdr08160.
8. Ernst AT, Jiang H, Krishnamoorthy M, Sier D. Staff scheduling and rostering: A review of applications, methods and models. Eur J Oper Res. 2004;153(1):3-27. https://doi.org/10.1016/s0377-2217(03)00095-x.
9. Griffiths P, Saville C, Ball J, Jones J, Pattison N, Monks T, et al. Nursing workload, Nurse Staffing Methodologies & Tools: a systematic scoping review & discussion. International Journal of Nursing Studies. 2020;103(March). https://doi.org/10.1016/j.ijnurstu.2019.103487.
10. Ball J, Barker H, Griffiths P, Jones J, Lawless JA, Burton C, et al. Implementation, Impact and Costs of Policies for Safe Staffing in Acute Trusts. The University of Southampton; 2019. https://doi.org/10.5258/SOTON/P0012.
11. InnovaHealthTec. Adult Inpatient Wards SNCT 2023 [Available from: https://www.innovahealthtec.com/adult-in-patient-ward-snct.]
12. Allen D, Jacob N, Strange H, Jones A, Burton C, Rafferty AM. “It's not just about the numbers”: Inside the black box of nurses' professional judgement in nurse staffing Systems in England and Wales: Insights from a qualitative cross-case comparative study. International Journal of Nursing Studies. 2023;147. https://doi.org/10.1016/j.ijnurstu.2023.104586.
13. Griffiths P, Saville C, Ball JE, Culliford D, Pattison N, Monks T. Performance of the Safer Nursing Care Tool to measure nurse staffing requirements in acute hospitals: a multi-centre observational study. BMJ Open. 2020;10(e035828). https://doi.org/10.1136/bmjopen-2019-035828.
14. Hoddinott S. The NHS crisis. Institute for Government. 2023. Available from: https://www.instituteforgovernment.org.uk/sites/default/files/2023-01/NHS-crisis-does-government-have-a-plan.pdf?ref=health-policy-insight
15. Smith G, Redfern O, Maruotti A, Recio-Saucedo A, Griffiths P. The association between nurse staffing levels and a failure to respond to patients with deranged physiology: A retrospective observational study in the UK. Resuscitation. 2020; 149. https://doi.org/10.1016/j.resuscitation.2020.01.001.
16. Junttila JK, Koivu A, Fagerström L, Haatainen K, Nykänen P. Hospital mortality and optimality of nursing workload: A study on the predictive validity of the RAFAELA Nursing Intensity and Staffing system. International Journal of Nursing Studies. 2016;60. https://doi.org/10.1016/j.ijnurstu.2016.03.008.
17. Griffiths P, Maruotti A, Recio Saucedo A, Redfern O, Ball J, Briggs J, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ Quality and Safety. 2019;28:609-17. https://doi.org/10.1136/bmjqs-2018-008043.
18. Shin S, Park JH, Bae SH. Nurse staffing and nurse outcomes: A systematic review and meta-analysis. Nurs Outlook. 2018;66(3):273-82. https://doi.org/10.1016/j.outlook.2017.12.002.
19. Burton CR, Rycroft-Malone J, Williams L, Davies S, McBride A, Hall B, et al. NHS managers' use of nursing workforce planning and deployment technologies: a realist synthesis. Health Services and Delivery Research. 2018;6(36). https://doi.org/10.3310/hsdr06360.
20. Finnish Institute of Occupational Health. Working Time Traffic Light Model improved working times and occupational safety in the social welfare and health care sector 2024 [Available from:

https://www.ttl.fi/en/topical/press-release/working-time-traffic-light-model-improved-working-times-and-occupational-safety-in-the-social.]
21. Lewinski-Corwin EH. The hospital nursing situation. American Journal of Nursing. 1922;22(8):603-6. https://journals.lww.com/ajnonline/citation/1922/05000/The-Hospital-NursingSituation.2.aspx.
22. Fasoli DR, Haddock KS. Results of an integrative review of patient classification systems. Annual Review of Nursing Research. 2010;28:295-316.
23. Edwardson SR, Giovannetti PB. Nursing Workload Measurement Systems. Annual Review of Nursing Research. 1994;12(1):95-123. https://doi.org/10.1891/0739-6686.12.1.95.
24. Aydelotte MK. Nurse Staffing Methodology, A Review and Critique of Selected Literature. Washington: Public Health Service (DHEW) , Washington, D.C. Div. of Nursing.; 1973. Contract No.: DHEW-NIH-73-433.
25. Griffiths P, Ball J, Bloor K, Bohning D, Briggs J, Dall'Ora C, et al. Nurse staffing levels, missed vital signs and mortality in hospitals: retrospective longitudinal observational study. Health Services and Delivery Research. 2018;6(38). https://doi.org/10.3310/hsdr06380.
26. Fagerstrom L, Kinnunen M, Saarela J. Nursing workload, patient safety incidents and mortality: an observational study from Finland. BMJ Open. 2018;8(4):e016367. https://doi.org/10.1136/bmjopen-2017-016367.
27. Needleman J, Buerhaus P, Pankratz VS, Leibson CL, Stevens SR, Harris M. Nurse staffing and inpatient hospital mortality. N Engl J Med. 2011;364(11):1037-45. https://www.nejm.org/doi/full/10.1056/nejmsa1001025.
28. Fagerström L, Lønning K, Andersen MH. The RAFAELA system: a workforce planning tool for nurse staffing and human resource management. Nursing Management. 2014;21(2):30-6. https://doi.org/10.7748/nm2014.04.21.2.30.e1199.
29. van Oostveen CJ, Ubbink DT, Mens MA, Pompe EA, Vermeulen H. Pre-implementation studies of a workforce planning tool for nurse staffing and human resource management in university hospitals. Journal of nursing management. 2016;24(2):184-91. https://doi.org/10.1111/jonm.12297.
30. Lee SH, Lee H, Yu S. Factors associated with nursing needs and nursing hours in acute care hospital settings: A cross-sectional study. J Nurs Manag. 2022;30(6):2005-14. https://onlinelibrary.wiley.com/doi/abs/10.1111/jonm.13634.
31. Myny D, De Bacquer D, Van Hecke A, Beeckman D, Verhaeghe S, Van Goubergen D. Validation of standard times and influencing factors during the development of the Workload Indicator for Nursing. J Adv Nurs. 2014;70(3):674-86. https://doi.org/10.1111/jan.12232.
32. Dall'Ora C, Griffiths P, Hope J, Briggs J, Jeremy J, Gerry S, et al. How long do nursing staff take to measure and record patients' vital signs observations in hospital? A time-and-motion study. Int J Nurs Stud. 2021;118:103921. https://doi.org/10.1016/j.ijnurstu.2021.103921.
33. Nieri A-S, Spithouraki E, Galanis P, Kaitelidou D, Matziou V, Giannakopoulou M. Clinical Severity as a Predictor of Nursing Workload in Pediatric Intensive Care Units: A Cross-Sectional Study. Connect: The World of Critical Care Nursing. 2019(4):175-84. https://doi.org/10.1891/WFCCN-D-19-00011.
34. Bekelaar JWR, Luime JJ, de Carvalho RM. Predicting Patient Care Acuity: An LSTM Approach for Days-to-day Prediction. In: Montali MS, A. Weidlich, M., editor. ICPM 2022. Cham: Springer; 2023. https://doi.org/10.1007/978-3-031-27815-0_28.
35. Portsmouth Hospitals University NHS Trust. About us 2024 [Available from: https://www.porthosp.nhs.uk/about-us.]
36. Imperial College Healthcare NHS Trust. About us 2024 [Available from: https://www.imperial.nhs.uk/about-us.]
37. NHS England. Bed Availability and Occupancy Data – Overnight. NHS organisations in England, Quarter 3, 2023-24 (revised 23052024)2024. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/.
38. Margetis G, Ntoa S, Antona M, Stephanidis C. Human-centered design of artificial intelligence. Handbook of human factors and ergonomics 2021. p 1085-106. https://doi.org/10.1002/9781119636113.ch42.
39. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology. 2006;3(2):77-101. https://doi.org/10.1191/1478088706qp063oa.
40. Van Der Laan JD, Heino A, De Waard D. A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research Part C: Emerging Technologies. 1997;5(1):1-10. https://doi.org/10.1016/S0968-090X(96)00025-3.

41. Brooke J. SUS: A “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability evaluation in industry. London: Taylor & Francis; 1996. p 189–94.
42. Saville C, Griffiths P, Casey A, Chable R, Chapman H, Radford M, et al. Professional Judgement Framework: A guide to applying professional judgement in nurse staffing reviews. 2023. Available from: https://eprints.soton.ac.uk/478317/.
43. The Shelford group. Safer Nursing Care Tool 2023 [Available from: https://shelfordgroup.org/safer-nursing-care-tool/.]
44. Myny D, Van Goubergen D, Gobert M, Vanderwee K, Van Hecke A, Defloor T. Non-direct patient care factors influencing nursing workload: a review of the literature. J Adv Nurs. 2011;67(10):2109-29. https://doi.org/10.1111/j.1365-2648.2011.05689.x.
45. Myny D, Van Hecke A, De Bacquer D, Verhaeghe S, Gobert M, Defloor T, et al. Determining a set of measurable and relevant factors affecting nursing workload in the acute care hospital setting: a cross-sectional study. Int J Nurs Stud. 2012;49(4):427-36. https://doi.org/10.1016/j.ijnurstu.2011.10.005.
46. Yakusheva O, Bang JT, Hughes RG, Bobay KL, Costa L, Weiss ME. Nonlinear association of nurse staffing and readmissions uncovered in machine learning analysis. Health Serv Res. 2022;57(2):311-21. https://doi.org/10.1111/1475-6773.13695.
47. Kaw N, Murray J, Lopez AJ, Mamdani MM. Nursing resource team capacity planning using forecasting and optimization methods: a case study. Journal of nursing management. 2020;28(2):229-38. https://doi.org/10.1111/jonm.12905.
48. Bell J. Machine learning : hands-on for developers and technical professionals. Hoboken: John Wiley & Sons; 2014. http://dx.doi.org/10.1002/9781119183464.
49. Ghavami P. Big Data Analytics Methods: Analytics Techniques in Data Mining, Deep Learning and Natural Language Processing. Berlin, Boston: De Gruyter; 2020.
50. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The lancet. 1986;327(8476):307-10. https://doi.org/10.1016/S0140-6736(86)90837-8.
51. van der Mark C, Vermeulen H, Hendriks PHJ, Oostveen CJV. Measuring perceived adequacy of staffing to incorporate nurses' judgement into hospital capacity management: a scoping review. BMJ Open. 2021;11(4):e045245. https://doi.org/10.1136/bmjopen-2020-045245.
52. Saville C, Griffiths P. Ward staffing guided by a patient classification system: A multi-criteria analysis of "fit" in three acute hospitals. Journal of nursing management. 2021;29(7):2260-9. https://doi.org/10.1111/jonm.13341.
53. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-74. https://doi.org/10.1177/0272989X06295361.
54. Suresh H, Guttag J. A framework for understanding sources of harm throughout the machine learning life cycle. In Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization 2021 (p1-9). https://doi.org/10.1145/3465416.3483305.
55. Office for National Statistics. Regional ethnic diversity 2022 [Available from: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest/.]
56. Portsmouth City Council. Portsmouth demography – JSNA report 2024 [Available from: https://www.portsmouth.gov.uk/services/health-and-care/health/joint-strategic-needs-assessment/portsmouth-demography/.]
57. Saville C, Monks T. Video of nurse staffing simulation 2019 [Available from: https://eprints.soton.ac.uk/id/eprint/430632.]
58. Saville C, Griffiths P, Ball J. How effective and cost-effective is the Safer Nursing Care Tool? Nursing Times. 2021;117(3):29-30. https://www.nursingtimes.net/clinical-archive/wellbeing-for-nurses/how-effective-and-cost-effective-is-the-safer-nursing-care-tool-08-02-2021/.
59. Saville C, Griffiths P, Ejebu Z, Fogg C. Nurse staffing decisions: the Professional Judgement Framework. Nursing Times. 2024;120(7). https://www.nursingtimes.net/clinical-archive/patient-safety/nurse-staffing-decisions-the-professional-judgement-framework-24-06-2024/.