Improvement Analytics Unit

Protocol for an evaluation of a Wakefield care home intervention

Emma Vestesson, Therese Lloyd and Adam Steventon
January 2018

Addendum added April 2019

The Health Foundation

Contents

Glossary of acronyms

Abbreviation	Description
ANP	Advanced nurse practitioner
CCG	Clinical commissioning group
CQC	Care Quality Commission
EHCH	Enhanced health in care homes
GP	General practitioner
IAU	Improvement Analytics Unit
ID	Identifier (patient)
KPI	Key performance indicator
MCP	Multispecialty community provider
MDT	Multidisciplinary team
NHAIS	National Health Applications and Infrastructure Services
NCDR	National Commissioning Data Repository
PARR	Patients at risk of re-hospitalisation
SUS	Secondary Uses Service
WCH	Wakefield EHCH vanguard

Summary

Purpose of this document

This statistical analysis plan describes in detail all aspects of the proposed Improvement Analytics Unit (IAU) evaluation of the enhanced health in care homes (EHCH) vanguard in Wakefield, including the study design, statistical methods and variable definitions. In addition, it details the limitations of the analysis and how these should be considered when interpreting the results. This plan has been written before the analysis begins to ensure that all design and methods choices are made objectively and are not influenced by what is found in the data. In rare instances, it may be necessary to make changes to the design of the study at a later stage; if so, these changes will be mentioned and their rationale explained in the report or accompanying material.

This document has been agreed with the Wakefield vanguard to ensure clarity of purpose. It may also be of interest to others involved in evaluation and analysis of large data sets. The IAU welcomes comments and questions on this document.

This is a technical document written to guide analytical processes. The Summary section provides a more accessible overview of the proposed study.

At the completion of the study, this document will be appended to identify whether there was any deviation from the planned approach.

Purpose of this study

The IAU will be conducting an analytical study to feedback on the progress being made to improve secondary care outcomes for residents living in the 15 intervention care homes that were part of the Wakefield EHCH between February 2016 to March 2017 (referred to as phase 1).

What the study will look at

The study will examine whether there was an impact on secondary care outcomes for residents who lived in one of the 15 vanguard care homes between mid-February 2016 and mid-March 2017.

The IAU will compare the secondary outcomes of the residents in the vanguard care homes to a 'local matched control group' comprised of residents living in similar care homes in Wakefield. This 'local matched control group' will be chosen so that its residents are also as similar as possible to the vanguard care home residents in characteristics such as age, gender, co-morbidities and number of hospital admissions in the year prior to moving to a care home. The outcomes will be compared over the follow-up period, which can be up to 13 months, taking into account each person's length of time in a care home during the intervention period.

The primary outcomes examined will be:

- the number of emergency admissions per resident
- the proportion of total emergency bed days per resident, relative to the total number of days at risk (ie days in the follow-up period).

The secondary outcomes examined will be:

- the number of potentially avoidable emergency admissions per resident (defined below)
- the number of A&E attendances per resident
- the number of elective admissions per resident
- the number of outpatient attendances (excluding 'did not attends') per resident
- the proportion of total elective bed days per resident, relative to the total number of days at risk (ie days in the follow-up period)
- the proportion of deaths outside of hospital (used as a proxy for end-of-life care occurring in the patient's place of choice).

Outcomes in the intervention group will be compared with those of the local matched control group over a period of up to 13 months, taking into account each person's length of time in the study, and further adjusting for any remaining differences in care home or resident characteristics.

Subject to gaining timely access to local data, an additional analysis will try to estimate the impact of the multidisciplinary team (MDT) intervention compared with the base package of care in the vanguard care homes.

Data the IAU will use

The IAU will use pseudonymised patient-level national Secondary Uses Services (SUS) administrative hospital data for England from mid-August 2011 to mid-March 2017. The IAU will also use pseudonymised resident-level National Health Applications and Infrastructure Services (NHAIS) data to identify care home residents in Wakefield.

If the IAU can get access to local data, the IAU will also use a list of pseudonymised patient identifiers (IDs) for patients who were referred to MDTs from mid-November 2015 to mid-March 2017 and the date of the referral, supplied by Mid Yorkshire Hospitals NHS Trust.

'Pseudonymised' means that all direct IDs (eg name, address, date of birth, NHS number for patients) are removed from the data. Pseudonymisation reduces the risk that individual patients can be identified from the data.

Period and residents covered by the study

The study will look at all residents aged 65 or older living in any of the 15 intervention care homes in Wakefield Clinical Commissioning Group (CCG) at any point between mid-February 2016 and mid-March 2017. The IAU will look at impacts on hospital utilisation up until mid-March 2017.

The study will use hospital data for the two-year period before each person entered the study. This will enable the characteristics for each resident to be determined; these will be used to match residents to other similar residents.

Strengths and weaknesses of the study

The study will evaluate the impact of living in one of the intervention care homes on hospital utilisation, derived from national administrative hospital data (SUS data). It will not measure utilisation of other health and care services, impacts to quality of life, staff satisfaction or the quality of working relationships. The IAU evaluation should be viewed in conjunction with the qualitative research carried out by the local evaluation of the Wakefield EHCH.

The intervention and matched control groups will be similar on a range of observable characteristics both at care home and resident level, making the comparison more robust than without matching. The expected pool of potential controls is approximately twice the size of the intervention group. This is quite a low ratio of potential controls to intervention residents and can make it more difficult to make the intervention and matched control groups similar on all characteristics. Any remaining differences between the intervention and local matched control group will be adjusted for in the analysis.

There is a risk that the intervention and control groups are different in ways that cannot be observed by the IAU either at care home or resident level (for example, staffing ratios in care homes or a resident's level of social isolation). In particular, care homes were not selected randomly and may therefore be different to non-intervention care homes in Wakefield CCG, which can bias the results. However, by using local controls, instead of controls from outside Wakefield CCG, the risk of bias due to area- or hospital-level differences is mitigated, as both intervention and control groups will have access to the same hospital services and are also likely to be similar on factors that can influence outcomes but cannot be observed by the IAU.

It is estimated that approximately 1,200 residents will have lived in one of the intervention care homes during the period of our study. This may be too small a sample to detect a statistically significant difference between the intervention and control groups. This is particularly the case for any subgroup analysis.

The study will evaluate the effect of living in an intervention care home for a variable period of up to approximately 13 months. For many residents, the follow-up period will be much shorter. It is possible that the study period is not long enough to allow for the long-term impact of living in an intervention care home to become apparent.

The results are nonetheless expected to enable learning that, together with the local evaluation and other evidence, will help Wakefield EHCH vanguard understand what is happening on the ground, assess what is working and identify potential areas for further investigation or improvement.

Background

The Wakefield context

Wakefield is located in West Yorkshire and Wakefield Clinical Commissioning Group (CCG) plans and funds health care for 360,000 patients. It is the 65th most deprived district in England (out of 326 districts) with over 40,000 people living in neighbourhoods that are in the top 10% of most deprived in England while other areas are considerably more affluent. Overall, the health of people in Wakefield is worse than the average in England with an ageing population and increasing rates of long-term conditions such as diabetes and coronary heart disease.[1] This is likely to be reflected in Wakefield's care home population.

Wakefield was selected by NHS England to be one of six enhanced health in care homes (EHCH) vanguards in March 2015. The focus of the vanguard is to provide more joined-up care that could reduce unnecessarily long hospital stays and reduce loneliness in care homes and supported living schemes (a supported living scheme is a set-up that offers personal care in a person's home). Phase 1 of the Wakefield EHCH vanguard (referred to as WCH in this document) started in November 2015 when an enhanced care package was introduced in 15 care homes. Elements of the enhanced care package were also implemented in two supported living schemes.[2] In April 2017, phase 2 of WCH started with residents in 12 more care homes getting access to the enhanced care.

Although this study focuses on the effect of the enhanced care package in the first 15 Wakefield care homes introduced in phase 1, it is important to note that there had been ongoing work to transform and integrate health and social care in Wakefield before the start of WCH. Health and social care hubs covering the whole Wakefield district opened in 2014 with the aim of preventing avoidable hospital admissions and supporting services to enable people to be discharged from hospital earlier.

In addition to the WCH, a federation of GP practices in Wakefield called West Wakefield Health and Wellbeing was selected to be a multispecialty community provider (MCP) vanguard in March 2015. The MCP, which aimed to move specialist care out of hospital and into the community, started operating at the end of 2015. This included a system that allowed patients to be seen by a physiotherapist without requiring a GP referral, having pharmacists in GP surgeries, better online access to health care and extended hours for GP services.[3] During the time period that the Improvement Analytics Unit (IAU) was evaluating, the MCP was completely separate from the care home vanguard and unlike the care home vanguard there was no specific focus on older people. In April 2017, the care home vanguard and MCP vanguard in Wakefield came together under the Connecting Care programme.

Wakefield is also part of the West Yorkshire urgent and emergency care vanguard led by the Healthy Futures Programme.[1] Mid Yorkshire Hospitals NHS Trust, the main trust in the area, has been a part of the Future Hospital Programme since October 2015. The purpose of the programme is to improve the care of frail and older people. As a part of the programme, a team called Rapid Elderly Assessment Care Team or REACT was set up. REACT assesses older patients when they arrive at hospital and makes sure their medical needs are identified. This is believed to reduce the length of stay in hospital.[4]

Between December 2016 and April 2017 Yorkshire Ambulance Service ran a project to ensure that falls patients received an appropriate and timely response. One of the goals was to reduce the number of falls patients attending A&E and being admitted to hospital. The project covered all of Wakefield and there was no particular focus on care home residents.[5]

Wakefield care home vanguard

WCH had several aims, some focused on hospital activity and some on the more social aspects of life in the care homes. One of the main aims of the WCH was to make care more coordinated and seamless. It was expected that this would help make sure that urgent care was only provided to those who really needed it as well as making the discharge process smoother and thereby avoiding delays. Another aim was to ensure that each care home resident had an end-of-life plan that could increase the number of people dying in their place of choice. Finally there was also the aim to reduce loneliness and isolation in care homes.[1] The 15 care homes – four residential homes and 11 nursing homes – were selected by Wakefield CCG based on differing criteria. Three care homes had been part of a pilot scheme, some were recommended by GPs and some were picked based on the type of specialties they provided (eg over/under 65s/specialist care) or because of high hospital-level activity. In 11 of the care homes, the changes were implemented in February 2016 while in the remaining four the changes were implemented in September 2016. There were around 1,000 beds in the intervention care homes with the median number of beds per care home being 64. The largest care home had 180 beds, which is almost twice as many beds as the second largest care home. Based on the number of care home beds, the IAU estimates that 1,200 residents will have been in the intervention group during phase 1 and 2,000–2,500 residents will have been in similar care homes in Wakefield.

All care homes in the intervention catered to adults 65 and over and some also catered to other age groups. The focus on WCH was on older people and it is expected that only a few people living in the intervention care homes will have been under 65.

The care home intervention in Wakefield had several different components, broadly composed of services delivered by the voluntary sector, a MDT and new key performance indicators (KPIs) for GPs. These are all described in more detail below.

Voluntary organisations have done a lot of work to get residents to interact with each other and to be more connected to the local community. Health walks, church visits and tea dances have been organised outside of the care home and other activities have included bringing a dog or a choir into the care home.[2] Two programmes in particular worked with care home residents on a one-to-one basis. 'Pull Up A Chair' is a programme run by Age UK and delivered in care homes and supported living schemes. It is a filmed interview and personal diary where residents have a chance to talk about their living situation. If any extra needs are identified these are shared with people who might be able to help.[6] 'Portrait of a life' is a toolkit developed by South West Yorkshire Partnership NHS Foundation Trust to help people working in health and social care settings to do life story work in a dementia and older people's service environment. By getting a fuller picture of a resident's life journey, care can be better tailored to their needs.[7] Through Carers Wakefield, a local charitable voluntary organisation, carers with family or friends in a care home can get information, advice and support when attending meetings and help with finding other services.[8]

Care home staff in care homes supported by WCH can refer residents to a MDT. The MDT was set up as a part of WCH; only residents in the care homes supported by WCH could access the MDT. The MDT was made up of health care professionals from a number of specialisms including mental health, physiotherapy and nursing. Originally the plan was to include a pharmacist but due to funding issues this did not happen. Nevertheless, members of the MDT did have the option of referring residents to a community pharmacist or to the GPs visiting the care home for medicine reviews. Members of the MDT were based in the same location and they meet once a week to plan the care that will be offered to the care home residents in the intervention care homes. The MDT also performed falls risk assessment and staff training. Initially the goal was for the MDT to screen all residents for unmet needs but this was changed to only high-risk residents from April 2016. Care home staff determine who is referred to the MDTs but there are no set criteria to determine who is at 'high risk' and, anecdotally, some care homes referred most of their residents to the MDT. It is estimated that roughly 400 MDT referrals had been made since November 2015.

All residents benefited from staff training indirectly. The staff training was delivered by the MDT and included falls prevention and what to do after a fall, how to screen for malnutrition and swallowing problems, dementia awareness and pressure sore prevention. Not all members of staff attended all training sessions but in total 286 staff members received some training.

Out of roughly 40 GP practices, 26 participated in the WCH. Any GP practice with at least one patient in a care home supported by WCH was selected to participate. All GPs in Wakefield already visited their care home patients when required before the start of the vanguard. But starting from November 2015 a new set of KPIs was set for all GPs in Wakefield CCG as part of the EHCH. With reference to care home residents, these include a face-to-face consultation and provisional health care assessment within 14 days of registering for new residents, a full health care plan within eight weeks of registering and emergency admission reviews within seven days of discharge letter being received. One KPI applied only to participating GP practices – this was linked to administrative support for care homes to access SystmOne – but the other KPIs applied to all GP practices in Wakefield. Some GP surgeries had weekly ward rounds and others visited only when called. The frequency of visits differed between practices and was largely based on the practice's relationship with the care home. Sometimes an advanced nurse practitioner (ANP) would visit instead of a GP, based on the needs of the residents. All care home residents in Wakefield were seen by a GP in their care home. This was not linked to living in an intervention care home. GPs visiting care homes is not linked to the MCP described in the preceding section, even though nine of the GP practices participated in both vanguards.

One of the core elements of the intervention aimed to make sure that care home staff had access to the appropriate clinical information. This was considered to be important to support clinically safe early discharge from hospital and the continuing care of the resident in the care homes. The plan was to give care homes access to SystmOne and NHSmail for secure transfer of patient information. This was difficult to implement: only one care home had access to SystmOne by the end of phase 1 and no care home had access to NHSmail.

The start dates and who had access to the different components are outlined in Table 1. Most of the components of the enhanced care package were implemented in February 2016.

However, the GP/ANP component started in November 2015 and the MDT started being set up in November 2015 but was not fully staffed until February 2016, and not all residents had access to all components.

Table 1. Timelines of WCH intervention components

Intervention	Start date	Available to everyone in intervention care home	Available to everyone in non-intervention care home
GP/ANP	November 2015	Yes	Yes
MDT	February 2016 (set-up started in Nov 2015)	By referral, high risk only from April 2016	No
Staff training	February 2016	Yes indirectly	No
Voluntary (eg Portrait of a life)	February 2016	Yes	No

Intended impact of enhanced care in care homes on outcomes

The enhanced care package is expected to make care more coordinated and seamless and lead to urgent care only being provided to those who need it. A strong emphasis is put on reducing demand on secondary care from care home residents and some of the stated intended benefits for residents are to reduce accidents and health deterioration resulting in hospital attendance or admissions. The vanguard aims to ensure that each resident has an end-of-life plan to allow people to die in their place of choice.[1] The enhanced care in the care homes is expected to provide a smoother discharge process, thereby reducing the number of bed days. There is an aim to improve management of long-term conditions, falls and end-of-life care as well as increasing proactive case management and personalised care planning. There is also a strong focus on reducing loneliness and isolation.[9]

The MDT is expected to improve case management and personalised care planning. Staff training should improve management of falls and end-of-life care. 'Pull Up A Chair' and 'Portrait of a life' are examples of initiatives that are believed to reduce loneliness. Closer collaboration between GPs, MDT members and care home staff is believed to make care more coordinated.

Objectives of the analysis

The IAU evaluation will evaluate the effect of receiving enhanced support in care homes for residents aged 65 and over.

The IAU will compare the outcomes of patients who lived in one of the care homes supported by WCH to local matched controls living in other similar care homes in Wakefield CCG. The local matched control group within Wakefield CCG will be similar to the intervention group in terms of a range of observable resident characteristics, such as age, gender, long-term conditions and prior hospital activity at resident level. They will also live in similar care homes,

as determined by variables such as care home type (nursing or residential) and the number of beds in the home, and will have access to similar services, eg Mid Yorkshire Hospitals NHS Trust's Rapid Elderly Assessment Care Team, RAPID.

After matching, appropriate regression models will be used to compare the two groups, thereby allowing for further adjustment for any remaining observable differences between them.

This evaluation aims to assess the effect of the enhanced care package on all residents that lived in a care home in Wakefield CCG at any point during the period mid-February 2016 to mid-March 2017, ie phase 1 of the EHCH vanguard.

Phase 2 started in April 2017, when 12 more care homes in Wakefield CCG joined the intervention.* Although the phase 1 care homes continued to receive the enhanced care package after April 2017, this study will only evaluate the period until March 2017, as the intervention received is likely to have changed as more care homes joined WCH and resources (in particular MDT staffing) did not increase proportionately.

The IAU will not evaluate other potential impacts of the interventions, such as quality of life or staff satisfaction, due to the limitations of the data available to the IAU. Costs will not be evaluated in this study.

Methods

Study design

The IAU will compare the outcomes of patients living in care homes supported by WCH with those of a retrospectively matched local control group of residents living in other care homes in Wakefield CCG. The IAU will match patients based on their baseline patient characteristics such as age, co-morbidities and prior hospital activity, and their care home's characteristics such as the number of beds and type of care home (nursing or residential). In other words, routine resident-level information will be used to characterise the intervention residents at baseline and then to select matched controls that had similar observed characteristics and lived in similar care homes. These matched residents will form the control group for the evaluation of WCH.

Once matched controls have been selected, the effect of living in an intervention care home compared with a non-intervention care home will be estimated by fitting multivariable regression models to a selection of study endpoints (see page 16).

Study cohorts

Definition of intervention population

The intervention population will be defined as those residents aged 65 or older living in any of the 15 intervention care homes in Wakefield CCG at any point between mid-February 2016 and mid-March 2017. Study dates are mid-month due to limitations of the data used (see the

* We did not include phase 2 care homes in the analysis as there would only have been data for approximately the first six months, which was deemed to be too short to detect any impact.

Sources of data section, from page 14, for more information). The exact dates used depend on the date of the NHAIS monthly extraction. Both new residents and residents already living in the care home at the start of the intervention will be included.

Definition of study cohort

The study cohort, consisting of intervention residents and all potential controls aged 65 and over, is determined using both care home and resident characteristics.

The following residents will be excluded:

- residents without full address recorded in the NHAIS data (see below)
- residents younger than 65 when they joined the study
- residents without a recorded month and year of birth
- residents without a record of prior emergency or elective admissions in the three-year pre-period (defined in the Baseline variables section, from page 20). These residents will be excluded because prior hospital data are required to define baseline resident characteristics
- intervention residents who were not matched to a control resident
- residents with a follow-up period of less than a month.

It is unlikely that any intervention residents will be excluded because they could not be matched to a control resident. Characteristics for any excluded intervention residents will be reported on.

For residents that moved from one care home to another during the study period, only the stay in the first care home will be included.

Although the study period extends to mid-March 2017, only residents who moved in by mid-February 2017 will be included to allow for a follow-up period of at least one month.

The study cohort will consist of all residents who lived in a care home in Wakefield CCG at any time during the study period (as defined in the Baseline variables section, from page 20). As residents will be matched on hospital activity before the study period, the IAU needs to differentiate between residents who moved in after the intervention started (so-called new residents) and those that were already care home residents when the intervention started (so-called existing residents). The groups will be analysed together but there will be a slight difference in the matching. In all three groups, only residents aged 65 and over will be included.

At care home level, the study will include care homes located in Wakefield CCG caring for older people, but will exclude care homes that are likely to be specialist care homes for groups other than the frail older population as these would have very different primary care needs and usage of acute services. Data from the Care Quality Commission (CQC), the independent regulator of all health and social care services in England, is used for the inclusion criteria for care homes. The CQC data hold information on care home specialties and other characteristics such as the number of beds. Anecdotal evidence suggests that, especially when the CQC was first established in 2010, registering care homes would often add more categories than the categories were designed to capture. The IAU will therefore endeavour to exclude care

homes likely to be genuine specialist care homes but include care homes for older people that are able to accommodate residents with more complex needs or who are younger than 65. Excluding care homes that are likely to be specialist care homes will therefore be done by excluding care homes that meet both of the following criteria:

- has at least one of the following recorded specialties: learning disabilities or autistic spectrum disorder; people who misuse drugs and alcohol; people with eating disorders; people detained under the Mental Health Act; or people with sensory impairment
- is recorded as catering to additional age groups, other than just those aged 65 and over.

The study will not exclude care homes with specialisms in dementia care, mental health care or physical disability, as these categories are not inconsistent with the needs of frail older patients.

Although care homes that cater to younger age groups as well as older are not excluded, the resident-level inclusion criterion on age ensures that only residents aged 65 or older will form part of the evaluation. In addition, what age groups a care home caters to will be included in the matching.

Cohort group 1 will be based on residents that moved into a care home between mid-August 2014 (first available NHAIS extraction, see the Sources of data section, below, for more information) and the date the care home joined the intervention (either February 2016 or September 2016, depending on the care home). Only residents that stayed in the care home at least until one month after the intervention started will be included, to allow for a follow-up period of at least one month after the start of the intervention.

Cohort group 2 will be based on residents who moved into a care home between the start of the intervention and the end of the study period, ie between either mid-February 2016 or mid-September 2016 and mid-February 2017.

The IAU expects there to be 15 intervention care homes (11 nursing homes with 847 beds and four residential homes with 152 beds) and around 37 control care homes (eight nursing homes with 468 beds and 29 residential homes with 807 beds). Based on the number of care home beds, it is estimated that there will be 1,200 intervention residents and 2,000–2,500 potential control residents.

Sources of data

Data from four different sources will be linked together using pseudonymised patient identifiers (IDs). Each data source is described in more detail below.

SUS national administrative data

The IAU will have access to pseudonymised (ie anonymised in line with the Information Commissioner's Office Code of Practice on Anonymisation) SUS national administrative data, provided by the National Commissioning Data Repository (NCDR). SUS is a comprehensive repository for secondary health care data in England that is paid for by the NHS. It is used to support the NHS in the delivery of health care services and to trigger reimbursement for secondary care activity.

The IAU will create the analysis data set using SUS data for the period from mid-August 2011 to mid-March 2017. This will cover the study period, as well as three years preceding the intervention, which is data needed for measuring study covariates such as co-morbidities.

Data derived from the CQC

Since 2010 the independent regulator of health and adult social care in England, the CQC, regularly monitors, inspects and regulates health care services (including care homes) to ensure they meet fundamental standards of quality and safety. The IAU will obtain the full address of care homes from the CQC as well as data on care home type, capacity and specialties required to meet the inclusion/exclusion criteria at the care home analysis level as outlined in the section on Study cohorts. The CQC regularly carries out inspections of all care homes in England and the IAU considered including performance rating of care homes in the study, but as less than half of the care homes had been inspected under the new system when the interventions started, this is not possible. It is worth noting that the CQC registry is not designed for research purposes, nor is it properly validated; as such, there is the possibility that the lists of care homes on CQC record and those supplied by the vanguard team indicate inconsistent specialties. Therefore, care will be taken to ensure that genuine specialist care homes not fitting the inclusion criteria outlined previously are excluded from the analysis pool, as detailed in the Study cohorts section (see page 12).

Data derived from National Health Applications and Infrastructure Services

The NCDR holds monthly extracts from NHAIS from August 2014 up to a month before the date of the data transfer to the IAU. These monthly extracts, created on the first Sunday after the 13th day of the month, contain demographic information about all registrations with general practices in England, including date of birth, full residential address and the general practice at which the patient is registered. The date of death is also recorded in NHAIS for patients who died in the last five years.

The NCDR will identify from NHAIS extracts the month in which a resident has entered or left (through death or relocation, as applicable) a care home in Wakefield CCG during the study through an examination of changes in each resident's address history on record. The NCDR will then pseudonymise NHAIS information on each care home resident's month and year of birth, and estimate the dates of death, relocation into/out of a care home in Wakefield CCG and registering with a practice by using the date of the monthly NHAIS extract and make these available to the IAU. Via a pseudonymised patient identifier this data will be linked to the SUS database.

In summary, the NCDR will derive and provide the following limited data from the NHAIS database, for the period mid-August 2014 to mid-March 2017, for residents living in a care home in Wakefield CCG during the study period:

- residents' month and year of birth
- residents' estimated date of death
- residents' estimated date of moving to/from a Wakefield care home
- residents' general practice code and the estimated date of registering with the general practice.

The data derived from NHAIS will be linked to the SUS data via a pseudonymised patient ID.

The above outlined data extraction and linkage process will enable identification of the whole care home resident population in Wakefield CCG between mid-February 2016 and mid-March 2017 as well as when they moved to a care home – even those with no hospital admission record during the follow-up period – required for the evaluation. All linked secondary health care and care home data informing the present evaluation will be stored, processed and analysed by the IAU within an accredited secure data environment located in the Health Foundation. The overall approach to information governance used on this evaluation has been scrutinised by the IAU's Programme Oversight Group and information governance experts at NHS England and NHS Digital. The IAU at no point will have access to patient identifiable information related to this evaluation, nor plans on utilising more than the strictly necessary amount of data.

It is plausible that some GP practices are better at updating addresses than others so there could be systematic differences based on which GP a resident is registered with. In a previous study by the IAU where GP alignment was part of the intervention and therefore addresses were likely to be updated promptly when moving into a care home, an increase in hospital activity in the three months before moving into a care home and a drop in hospital activity in the following period were observed.[10] The IAU will investigate whether addresses seem to be updated in a timely manner by plotting hospital-level activity before and after moving into a care home to see if the same pattern as described above can be seen in Wakefield.

Data from Mid Yorkshire Hospitals NHS Trust

The IAU will request the following information from Mid Yorkshire Hospitals NHS Trust:

- pseudonymised patient IDs for patients who were referred to MDTs from November 2015 to March 2017
- date of referral.

Information on patient IDs and date of referral will allow the IAU to identify who was referred to a MDT and when they were referred. This will be used for the exploratory analysis detailed in Chapter 3 (see page 25).

The pseudonymised data from Mid Yorkshire Hospitals NHS Trust will be transferred to the Data Services for Commissioners' regional office (DSCRO), where it will be re-pseudonymised using a key common to the SUS data. It will then be transferred to the secure environment within the IAU and finally linked to the SUS data via the pseudonymised patient IDs.

Study endpoints

Primary outcomes

The primary outcomes are:

- the number of emergency admissions per patient over the follow-up period
- the proportion of total emergency bed days per patient, relative to the total number of days at risk (ie days in the follow-up period).

The primary endpoints will be modelled allowing for varying length of the follow-up periods, by including an offset for amount of time at risk in the statistical analysis. The length of the follow-up period will differ between patients, depending on the date on which they entered the study and whether they died or left the care home during the study period (see the Variable definitions section, page 18).

Emergency admissions are defined as separate hospital spells that either occur through the emergency room or as a result of direct, urgent referrals from a GP or other professional.

A bed day is defined as a night in hospital following an elective or emergency admission but excluding 'regular day/night attendances' (determined by a specific code in the raw SUS data). An admission and discharge within the same day will not count towards the total number of bed days. This is consistent with the definitions of bed day used within NHS England[11] and the NHS England New Models of Care dashboard, which displays outcome data for all vanguard sites.[12] This endpoint reflects changes to the length of stay in hospital as well as the number of admissions.

Secondary outcomes

Secondary outcomes, all calculated over the follow-up period, will be:

- the number of potentially avoidable emergency admissions per resident (defined below)
- the number of A&E attendances per resident
- the number of elective admissions per resident
- the number of outpatient attendances (excluding 'did not attends', and defined using code Attended=5 or 6 in SUS) per resident
- the proportion of total elective bed days per patient, relative to the total number of days at risk (ie days in the follow-up period)
- the proportion of deaths in hospital (used as a proxy for end-of-life care not occurring in the patient's place of choice)
- elective admissions are defined as non-emergency admissions, excluding maternity cases and 'regular day/night attendances' (determined by a specific code in the raw SUS data).

A list of potentially avoidable admission conditions in frail older people, also used by the CQC, will be used for this study. This list of conditions is not specific to care home residents but focuses on older people experiencing health and social care and includes: acute lower respiratory tract infections (such as acute bronchitis); chronic lower respiratory tract infections (such as emphysema and other chronic lung diseases); pressure sores; diabetes; food and drink issues (such as abnormal weight loss and poor intake of food and water due to self-neglect); food and liquid pneumonitis (inhaling food or drink); fracture and sprains; intestinal infections; pneumonia; and UTIs.[13,14]

The IAU will evaluate the number of outpatient attendances and elective admissions per resident. Both these outcomes could either decrease or increase with good care. The intervention could mean that some of this care could be avoided or even given in the care home leading to a decrease in activity. On the other hand, it could also mean that extra care

needs are identified which could lead to an increase in activity. Although interpretation may be difficult, these outcomes help create a fuller picture of how care home residents use secondary care, and they are therefore relevant endpoints.

The proportion of deaths in hospital will be used as a proxy for patients not dying in their preferred place of death. While patients may not die in their preferred place outside of hospital, it is assumed that locations outside of hospital (eg home, care home, hospice) would always be preferred over dying in hospital. Proportion of deaths is calculated by combining information on hospital deaths from the SUS data with information on all deaths from the NHAIS data. Only residents that died are included in the analysis of this endpoint.

Variable definitions

Bedding-in period and study period

As described in the Summary section from page 5, the intervention is composed of several components. In November 2015, the new KPIs were set for GPs and the MDT set-up started. Initially the MDT was understaffed, and it could not be considered to be fully up and running until February 2016. The initiatives delivered by Age UK and the training for care home staff started in February 2016. The three-month period between mid-November 2015 and mid-February 2016 will therefore be treated as a bedding-in period (ie a period omitted from the analysis to allow time for the interventions to become established). No outcome data from the bedding-in period will be used in the analysis.

The study period, ie the period that this evaluation covers, starts mid-February 2016 and ends in mid-March 2017 (the exact date being the date of the NHAIS monthly extraction).

Four care homes joined the intervention in September 2016. For residents in those care homes the study period will start mid-September 2016 and end mid-March 2017.

Exposure variable

A person is considered to be in the intervention group (ie exposed to the intervention) if they lived in a care home supported by WCH at some point between mid-February 2016 and mid-March 2017. Only residents who moved in by mid-February 2017 will be included to make sure that everyone has at least one month of exposure to the intervention. Persons in the intervention group will be identified using data from NHAIS as described in the Sources of data section, page 14.

Move-in date, index dates and follow-up period

A resident's move-in date will be the date of the monthly NHAIS extract in which they first appear with a care home address.

Outcome data on each resident will be collected for the specific period that they were living in a care home during the study period. This is known as the follow-up period. The start of the follow-up period, the so-called index date, depends on the resident's move-in date: for

residents who were already living in a care home at the start of the intervention, the index date is the start of the intervention (mid-February or mid-September 2016) and for residents who moved in after the intervention start, the index date is the date that they moved in.

The follow-up period ends when the person either dies, leaves the care home or the study period ends. The end date is defined as the earliest of the following dates:

- the end of the study period, ie mid-March 2017
- the date of death, set to the date of the monthly NHS extraction in which death is first recorded
- the date of leaving the care home, set to the day before the date of the monthly NHAIS extraction in which they were no longer registered as living at that care home.

A person's follow-up period may therefore be between approximately one and 13 months.

The date of death will be estimated as the extraction date of the month in which they died. However, as the true death date could be as early as the day after the previous extraction date, there is a risk of overestimating the follow-up period by up to a month, during which period the patient would have no recorded hospital activity. However, as it is unlikely that the day of the month that a person dies is other than random, this is unlikely to introduce bias between the intervention and control groups.

The date of moving out of the care home may also have occurred up to a month earlier. However, hospital activity will be recorded throughout, and it is unlikely that the dates of moving in/out are other than random and could therefore introduce bias between the intervention and control groups.

Pre-period and pre-move-in period

Previously diagnosed health conditions (co-morbidities) will be identified using data recorded during the 'pre-period', which is the period before a resident's index date.

For co-morbidities, patient-level data for three years prior to the index date will be used, consistent with for example the 'patients at risk of readmission' (PARR) predictive model.[15] A longer look-back period means it is more likely that someone has been admitted to hospital in that period and can therefore be included in the study.[16] The number of new patients identified through prior hospital admissions progressively diminishes over increasing look-back periods.[15] A longer look-back period also allows for more co-morbidities to be identified.[15,17] Most studies that explore varying look-back periods do not go beyond one year;[16] however, one study investigated up to five years' look-back.[16] A comparison of model fit and predictive ability for both modelling deaths and readmissions within 30 days found that both improved over the length of the look-back period, with a five-year look-back period being better than a three-year look-back period.[16] However, the magnitude of improvement progressively diminished over increasing look-back periods,[16,18] with an increase in co-morbidity prevalence of less than 1% when including admissions up to three years compared with two.[17] As some co-morbidities may resolve over time, a three-year look-back should adequately balance the need to identify patient characteristics while not unduly identifying historic co-morbidities that have since been resolved.

Residents who moved into a care home before the intervention started will have spent part of their pre-period in a care home. It is assumed that living in a care home does not affect the detection or documentation of co-morbidities during a hospital admission.

Prior hospital activity will be assembled using data recorded during the 'pre-move-in period', which is the period before a resident's move-in date. Data for one or two years before the individual moved into the care home will be used (the length of time depending on the activity, see Table 2).

Baseline variables

Baseline variables will be included in both the matching and the regression models. All baseline variables are calculated on either pre-period or pre-move-in data. Potential baseline variables to include in the matching or as covariates in the modelling are listed in Table 2.

Table 2. Resident-level baseline variables

Category	Variables at resident level
Demographics and socio-demographics	• Approximate age at index date • Gender • Ethnicity • Average socio-economic deprivation deciles, based on the Index of Multiple Deprivation (IMD) 2015, available at the lower layer super output area (LSOA) level. This will be based on where the person lived before moving into a care home
Prior hospital use	• Number of emergency admissions in year -1, the last year of the pre-move-in period (ie 365 days to 1 day before move-in date) • Number of emergency admissions in year -2 of the pre-move-in period (ie 730 to 366 days before index date) • Number of potentially avoidable emergency admissions in year -1 of the pre-move-in period • Number of elective admissions in year -1 of the pre-move-in period • Number of A&E attendances in year -1 of the pre-move-in period • Number of outpatient attendances in year -1 of the pre-move-in period • Number of emergency bed days in year -1 and -2 of the pre-move-in period • Number of elective bed days in year -1 of the pre-move-in period
Health variables	• Specific co-morbidities linked to frailty, identified in the pre-period[19] • Co-morbidities associated with emergency admissions, as identified by the PARR-30 model • The Charlson Co-morbidity Index (over a three-year look-back period)[20]
Seasonality	• Index date quarter
Time spent in care home before start of intervention	• Time in months between move-in date and index date

Average socio-economic deprivation quintiles will be defined based on the resident's most recent address before moving into a care home. This will better reflect the characteristics of the resident than the address of the care home.[15]

All co-morbidities (health variables in Table 2) will be defined using data from any diagnosis field in any hospital admission in the three-year pre-period.

Co-morbidities linked to frailty are: anxiety or depression; functional dependence; falls and significant fracture; incontinence; mobility problems; pressure ulcers; and cognitive impairment (composite of delirium, dementia and senility).[18]

Co-morbidities included in the PARR algorithm are: alcohol related diagnoses; cerebrovascular disease; chronic obstructive pulmonary disease; connective tissue disease/rheumatoid arthritis; developmental disability; diabetes; ischaemic heart disease; peripheral vascular disease; renal failure; and sickle cell disease.

Although some of the co-morbidities can be cured, the assumption will be that any issues reported in the pre-period will be relevant to the overall health or frailty of the person and therefore can be included as a potential covariate.

At care home level, potential covariates are shown in Table 3.

Table 3. Care home-level baseline variables

Category	Variable at care home level
Care home characteristics	• Care home type: nursing or residential (CQC data) • Number of beds available (CQC data) • Whether the care home caters exclusively to adults aged 65 and over, or also to another age group (adults under 65 or whole population) (CQC data)
Demographics	• Urban/rural classification at lower layer super output area (LSOA) level, based on the 2011 census
Hospital activity	Crude rates of: • emergency admissions • potentially avoidable emergency admissions • elective admissions • A&E attendances • outpatient attendances • emergency bed days • elective bed days at care home level in the year prior to the start of the intervention

Statistical methods

Identifying control group

The control group will be determined using matching methods to optimise similarity with the intervention group with respect to variables that are likely to be predictive of any of the outcomes.

Matched control observations will be selected using the genetic matching algorithm, which is a computer-intensive search procedure that produces more closely balanced groups than traditional approaches such as nearest neighbour matching or the propensity score.[21] The algorithm measures the similarity of pairs of patients using distance metrics that are generalised versions of Mahalanobis distance.[20,22] The distance metric contains weight parameters, which are optimised to produce a matched group that is as similar as possible to the intervention group.

The IAU will match 1:1, ie one control resident to one intervention resident. Matching will be done with replacement, ie the same control resident can be matched to several intervention residents. When matching with replacement, the matched controls are no longer independent and to minimise this problem the number of times the same control is used will be monitored.

The genetic matching algorithm will try various distance functions to determine the 'closeness' of the match. However, for some variables a match may be required to be exact or to fall within a pre-fixed 'caliper', whereby the variables are required to be within a fixed distance of one another. Table 4 shows those variables where special matching methods are to be applied.

Table 4. Matching variables and method of matching

Variable	Method of matching
Nursing or residential care home	Exact
Cohort group	Exact

Care home type (ie whether nursing or residential) will be matched on exactly, as the level of care provided in the two types of care home is different. Cohort group will be matched on exactly because of the different time periods used for prior hospital activity. The prior hospital activity before moving into a care home will be further back in time for the groups that were already in the care home at the start of the intervention and therefore these variables could have different abilities to predict outcomes. Another reason to match exactly on cohort is that residents will have had access to the interventions for different proportions of their care home stay and at different stages of their care home stay. It is not unlikely that having access to the enhanced care package from the point of moving in will be different to only having access towards the end of the person's stay. Exact matching also facilitates a subgroup analysis.

It is expected that all intervention patients will be matched to controls. Characteristics for any excluded intervention residents will be reported on.

As the IAU is not matching on length of time in the study (as this could be correlated with quality of care), the intervention and their control residents may differ in this respect. The IAU will conduct descriptive analyses regarding length of time in the study to check for differences between the control and intervention groups.

Choice of matching variables

The IAU will match on resident and care home characteristics to ensure similar patient case-mix between the groups. Although the IAU will ultimately aim to get balance across all variables as detailed in the Baseline variables section (page 20), this may not be possible. If this is the case, the IAU will define a subset of variables to be included in the matching algorithm. The IAU will empirically explore which covariates are most predictive of emergency admissions among the potential control group and whether the conditions are prevalent in the intervention and control groups. The subset of variables included in the matching algorithm will be adapted to optimise balance between the two groups on those variables considered most strongly predictive of the primary outcome, eg the prior numbers of emergency admissions, but also aiming to optimise balance across the wider set of variables.

Matching parameters

Table 5 gives the matching algorithm parameters to be used for this analysis.

Table 5. Matching parameters

Parameter	Value
Estimand	ATT
Number of controls per intervention patient	1
Control sampling method	With replacement
Population size	2,000
Maximum generation	1,000
Number of generations to wait	100

ATT = average treatment effect for the treated

Diagnostics

Balance will be assessed across all baseline variables listed in Tables 2 and 3 (pages 20–21) even if not all variables will be included in the matching algorithm.

Balance will be assessed using the standardised difference, which is defined as the difference in means as a proportion of the treatment group standard deviation.[23] Although the standardised difference should ideally be minimised without limit, a standardised difference below 10% has been used to describe negligible imbalance.[24] The standardised difference is a better measure of balance than formal statistical tests, as the latter depend on the size of the groups, as well as on the level of similarity.[25]

Statistical analysis

The IAU aims to estimate the 'average treatment effect for the treated' (ATT). This is the effect of the intervention for those who received it, ie those living in the care homes supported by WCH. This does not necessarily reflect the impact that implementing the intervention in all care homes in Wakefield CCG would have as the care homes were not selected at random. Once matched controls have been selected, the IAU will estimate the effect of living in an intervention care home compared with the control group by fitting multivariable regression models, both unadjusted and adjusted for covariates. The adjusted model will contain all variables that were used in the matching process to adjust for any remaining observed imbalance, as well as any other covariates predictive of outcome. Index date will be included as a quarter categorical variable to account for seasonality. Modelling checks for collinearity will be carried out and, if appropriate, the list of covariates will be changed accordingly.

Each outcome will be analysed by fitting a regression model that is appropriate to the type of outcome and the distributional properties of the data.

The intervention was provided at care home level and therefore the data are likely to be clustered at care home level. Explored modelling options will include multilevel models with care home as a random effect to capture the clustered nature of the data. A random effect modelling approach would provide more precise local estimates in the event of issues with data sparsity from particular care homes.

All count outcomes will be modelled allowing for varying study lengths by including an offset for amount of time at risk in the statistical analysis.

The count variable endpoints, eg number of emergency admissions per patient, will be analysed using a generalised linear model suitable for count data, such as a Poisson, to estimate the rate ratio between intervention and matched control patients. The effect of the intervention on the absolute (as opposed to relative) number of events (eg admissions) will also be estimated.

Model fit will be assessed by examining diagnostic statistics and over-dispersion parameters (eg the ratio of the residual deviance to the residual degrees of freedom), and excess zeros by comparing predicted and observed proportion of zero counts. If over-dispersion is detected then an alternative model, such as a negative binomial or a hurdle model for count data, will be fitted. Where a similarly good fit is obtained using multiple models a choice will be made by comparing the log-likelihood ratio and the Akaike Information Criterion (AIC), as appropriate.

As a guide, Table 6 details the typical regression models and alternatives for each outcome.

Table 6. Regression models for each outcome

Type of endpoint	Outcome	Initial model	Alternative model	Diagnostics
Count data	Emergency admissions/ Number of potentially avoidable hospital admissions/ elective admissions/ A&E attendances	Poisson	Negative binomial/ hurdle/ Zero-inflated Poisson	Over-dispersion Model fit Excess zeros
Proportions	Proportion of emergency/elective bed days (of whole follow-up period) Proportion of deaths in hospital/ patients who die (sensitivity analysis)	Binomial	Quasi-binomial/ negative binomial	Model fit Distribution of model residuals Over-dispersion Heteroscedasticity diagnostics

To account for differing attrition arising from death, moving away from the area or different entry dates into the study, an offset of the number of days in the study will be added to the model. However, the offset assumes that the number of days that are 'missing' is random and that the rate of outcomes, eg emergency admissions, is constant, when in fact this is unlikely to be the case. For example, a resident may use more hospital services in the final months of life. The IAU will examine the length of time people were followed up in the study and reasons for leaving the study between the groups. As detailed in the Variable definitions section (page 18), a resident's follow-up period ended (was censored) when the resident either died or moved out of the care home, or the study period ended. If the length of time residents were followed up in the study or the reasons for leaving the study differed between the groups, the IAU will consider doing an alternative analysis as a sensitivity analysis.

Exploratory analysis

In addition to estimating the overall effect of the WCH, the IAU will, if possible, do some exploratory analysis to get a better understanding of the effect of the MDT. This would require access to a pseudonymised list of patients who were referred to the MDTs from Wakefield CCG to identify those residents who also received care from the MDTs (see the Sources of data section, from page 14).

All WCH residents received the 'basic vanguard package', ie staff training, Carers Wakefield, Portrait of a life and Pull Up A Chair initiatives. By identifying WCH residents who did not also receive MDT care, the IAU can identify residents who only received the basic package.

Residents can be referred to the MDT more than once and some descriptive statistics on the frequency and timings of the MDT will be created to get a better understanding of MDT referral patterns. A MDT referral can have been triggered by hospital activity, eg after a fall, so if the outcome is measured from the index date it could look like residents that are referred

to a MDT actually have higher hospital activity. For those residents, the index would be reset to be the date for referral. If most MDT referrals are towards the end of residents' lives, then maybe only a subset of outcomes are appropriate to look at.

The exploratory analysis will be performed on the same matched control group as the main analysis. The IAU could estimate the effect of getting MDT care in addition to the basic vanguard package or the effect of getting the basic vanguard care and MDT care compared with standard care.

The main limitation of this analysis is the risk of low numbers. It is estimated that there were 400 referrals made to MDT but at this stage it is unclear how many individuals were referred to MDT. If for example the 400 referrals were split over 50 residents, then the numbers would be too low to detect any effect. However, the limitations will be better understood once the IAU has access to the data. A decision about whether to include the results of this exploratory analysis in the final report will depend on how reliable the IAU considers the results.

Subgroup analysis

No results will be presented at care home level, as this could potentially jeopardise patient confidentiality. As some of the care homes are small, patients may have been identifiable based on their characteristics. However, the population and the care delivered (not related to the vanguard) are likely to be different in nursing homes compared with residential homes so the intervention might have a different effect. Therefore, the IAU will conduct a subgroup analysis according to whether a care home is nursing or residential if there are sufficient numbers of residents in each group. There could also be a difference in the effect of the intervention between residents who moved in before or after the intervention started (cohort 1 versus cohort 2). Residents started receiving the intervention at different stages of their care home stay and residents in cohort 1 are likely to be frailer when the intervention started. If there are sufficient numbers, the IAU will do a subgroup analysis based on cohort. This also has the advantage of allowing for comparison with other IAU care home studies that only look at new residents.

No other subgroup analysis is planned.

Sensitivity analyses

One of the main threats to the validity of this study is unobserved confounding. That is, although the intervention and matched control groups are expected to be similar in terms of observed variables (such as age and prior number of hospital admissions), there may be systematic differences between these groups that are not observed, for example, the care home may have been selected because staff in the vanguard care homes might have been more open to change.

Although there is no definitive way to assess the effect of unobserved confounding, rates of the intervention and matched control groups on an endpoint unrelated to the intervention can be compared.[26] On the assumption that the intervention is unlikely to have had a large positive or negative impact on overall mortality within the follow-up period, then differences

in mortality rates would make us doubt the performance of the matching. For example, if enrolled patients died at a higher rate than matched control patients, this might suggest that they were in worse health than controls at the point of enrolment.[27] However, there is also a possibility that good care may result in prolonged life and therefore fewer deaths during our follow-up period. The IAU will therefore compare the rates of all-cause mortality over the study period as well as perform a survival analysis using Kaplan-Meier estimates, censoring patients who moved or reached the end of the study period. Similar mortality and survival rates will be indicative of balanced groups, while differences would need to be interpreted with caution.

If there is found to be a difference in mortality between the intervention and control groups, possible reasons for the imbalance will be explored and, if possible, the cohort will be subset in such a way as to remove any bias.

The IAU will also consider doing a difference-in-difference style analysis after finalising the final report.

Sample size calculation

No sample size was calculated for this study. The analysis is considered informative and will be carried out regardless of whether our study population is of a sufficient size to detect a statistically significant difference.

Limitations and sources of bias

Threats to validity

Internal validity

One of the main threats to the validity of this study is unobserved confounding. Unobserved confounding can occur at hospital, care home or resident level. The uncertainty of the results due to potential unobserved confounding will not be reflected in the confidence intervals or p-values, as these capture other kinds of uncertainties. The risk of unobserved confounding is mitigated by using local controls within Wakefield CCG.[28] Here, the local control group will have access to the same services, eg hospital services, thereby minimising the risk of bias. For example, differences in coding practices between hospitals could bias the detection of co-morbidities and therefore risk adjustment,[29,30] or hospital interventions could impact on outcomes. The IAU will perform some sensitivity analysis (see page 26) to try to assess the presence of unobserved confounding.

Furthermore, although matching will not ensure balanced groups on unobservable characteristics, it will ensure that the control and intervention groups are similar to a reasonable degree in observable variables, which is expected to produce a more similar control group than a non-matched comparison group. Using a local control group will also make it more likely that unobserved variables are similar between the two groups.

The care homes supported by WCH were not selected at random and may therefore be different from the care homes in the control group. Although observed care home characteristics will be matched on and adjusted for in the regressions, it is likely that there

are other care home factors that cannot be controlled for. Such unobserved differences could explain some or all of any difference in outcomes observed between the intervention and control group.

There is risk of self-selection into the care homes for cohort 2, meaning that people choosing a care home could base that decision on whether or not a care home was part of the intervention. If this were the case, residents in the intervention and control care homes could be systematically different – potentially in ways that cannot be observed and therefore not controlled for. The IAU does not think that whether a care home was in the intervention or not would have played a strong role in the choice of a care home as other factors such as care home location are more likely to be of greater importance.

Data on previous hospital activity is a strong predictor of future hospital use and is therefore used in the matching and regressions. Ideally hospital activity right before the start of the intervention would be used but for cohort 1 this period was when people lived in a care home whereas for cohort 2 it was right before they entered the care home. So using hospital activity right before the start of the intervention would not be comparing 'like for like' as hospital activity can differ between care home residents and people living at home.[10] For cohort 2 (ie residents who moved into a care home after the intervention started), hospital activity is captured right up until the resident started receiving the intervention. However, in cohort 1 (ie residents who already lived in the care home at the start of the intervention), any hospital activity while in the care home before the intervention started will be discarded and instead hospital activity before entering the care home is used in the matching algorithm and regression modelling. For cohort 1, patients will therefore be matched on hospital activity for up to 24 months before the follow-up period, thereby losing potentially valuable recent data on prior hospital use. Time between move-in date and index date will be one of the matching variables, to ensure that the cohort 1 residents are matched 'like for like'.

If important predictive covariates are omitted from the matching and regression models, then unobserved differences between the groups may contribute to any or part of any difference in outcomes. These differences would in turn then erroneously be attributed to the intervention.

There may be varying levels of engagement and interest in the MDTs among the health care professionals able to refer patients to the MDTs in different care homes. The decision to refer may therefore also be influenced by other factors such as primary care staff engagement or knowledge of the MDTs. As information on why a referral was made is not available, this cannot be accounted for in the analysis.

In some care homes, GPs visited weekly and in others, GPs visited on request of the care home. The effect of weekly GP visits may be different from that of GP visits on demand but different GP schedules cannot be identified in the data. If one set-up is better and there is a systematic difference between the intervention and control group so that more control residents have access to the better set-up then this will further bias any comparison between the control and intervention group.

The number of care home residents aged 65 and over in the non-intervention care homes in Wakefield is expected to be 2,000–2,500, which is a small pool of potential controls for the 1,200 expected number of residents in the intervention group. This means that it may

be difficult to find adequate balance across covariates. In addition, intervention care homes have on average more care home beds and the ratio of nursing to residential care homes is different in the intervention and potential control group. However, it is expected that the advantage of using a local control group still outweighs the problems linked to a small pool of potential comparators, especially as there were several other interventions within the CCG (see Chapter 1, Background, page 8).[27]

The use of a local control group increases the risk of contamination of the control group (ie when the control group has access to some or all of the interventions). This could for example happen if an intervention care home staff member received training and then started working in a control care home. This can bias the estimate of the effectiveness of the interventions. Using a local control group has advantages, such as unobservable characteristics being more likely to be similar, that outweigh the problems outlined above.

If the regression model is mis-specified, this could lead to biased inferences. However, matching on key variables before running the regression decreases the dependency on the specification.

Since the IAU will use data that are linked based on the full residential address, care home resident identification will be precise in most cases. However, this assumes that the addresses are updated in a timely manner in NHAIS. If addresses are not updated in NHAIS or not updated in a timely manner, then the distinction between the pre-period and the start of the intervention will be off, which will affect both the matching and the analysis and therefore bias the results. This is difficult to check. The IAU will also look at overall care home occupancy rates as very low occupancy rates could mean that residents are not being identified.

There was one intervention care home that shares a postcode with two care homes that were excluded after applying the exclusion criteria. Since the addresses will be very similar it is possible that some people have been wrongly assigned to one of the other care homes.

There is a risk that residents might be wrongly assigned to one of the other two care homes when updating their address with their GP. The IAU cannot identify specific care homes in the analysis data set so cannot check for this but, as mentioned above, occupancy rates for all care homes will be monitored.

If there are staff living on the premises of the care home who are aged 65 and over, then these staff will be wrongly categorised as part of the study cohort. The IAU is not aware that this is the case for any of the care homes.

External validity

This study will evaluate the effect of living in an intervention care home for a variable period of up to approximately 13 months. For many residents, the follow-up period will be much shorter. It is possible that the study period is not long enough to allow for the long-term impact of living in an intervention care home to become apparent.

The study is restricted to care home residents aged 65 and over. Results may not apply to a younger care home population.

New KPIs for GPs in Wakefield were set in November 2015 and all care home residents in Wakefield CCG received this level of care. If these KPIs are different from the KPIs in the rest of the country, then the effect of the basic vanguard package and MDT might be different if implemented somewhere else.

This study estimates the average effect on the treated. Implementing the same changes in all care homes in Wakefield might not have the same effect. Possible reasons for this could be that the care homes that first received the interventions are different or that scaling up the interventions would change the way they are delivered. For example, the MDT would have to be larger, which might make coordinated working more difficult.

Statistical conclusion validity

The IAU is expecting a sample size of approximately 1,200 residents in the intervention group. This may be too small a sample size to identify a significant effect if the effect is small or there is large variability within the groups.

A limitation of this study is that distinguishing between additional hospital activity that is due to good and timely care and other hospital activity is not possible.

Construct validity

SUS is an administrative database and has not been subjected to the cleaning rules that Hospital Episode Statistics are. However, the IAU Data Management Team will perform data checks and cleaning.

As the IAU only has access to the monthly table extracts from NHAIS and there is no date of when between the monthly snapshots a change occurred, index and end dates are sometimes approximated, using the date of the data extraction. As a result, the number of days in the study, which will be used to determine the offset in the models, will also only be approximate in some cases.

One of the aims of the intervention was to improve the quality of life of care home residents. The IAU will not be able to evaluate the effect on quality of life, level of isolation or improvement in working relationships, as the IAU only had access to secondary care data. The IAU evaluation is however supplemented by a local evaluation, which will do more qualitative work.

The aim to reduce secondary care activity is partly driven by quality of life factors but also by the fact that secondary care is expensive. The IAU will evaluate the effect on secondary care use but it will not evaluate the impact on the cost of care.

Reporting

General reporting considerations

Results will be reported as the relevant measure of effect, such as odds or rate ratios, plus 95% confidence intervals and p-values. Absolute numbers may also be presented, where appropriate. Both the post-matching unadjusted and adjusted analysis will be presented, and the variables used in the adjustment noted. Results will be presented to two decimal places for effect size and confidence intervals. P-values will be shown to three significant digits.

Special reporting requirements for this study

At a minimum the IAU will:

- adhere to the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) statement and the RECORD (Reporting of studies Conducted using[31] Observational Routinely-collected Data) guidelines[32]
- adhere to the NHS Digital (previously Health and Social Care Information Centre, HSCIC) small number rules[33]
- comply with the statistical code of practice.

Tables and figures for reporting matching results

Tables

A baseline table showing descriptive statistics for the intervention group and the matched and unmatched control populations, with:

- continuous variables summarised by mean (standard deviation, SD) or median (interquartile range, IQR) depending on the distribution
- categorical variables summarised by number (%)
- standardised differences calculated for the intervention group versus the unmatched and matched control groups, and variance ratio for continuous variables.

Figures

The following figures would be a minimal requirement:

- dot plot showing the standardised differences from both the matched and unmatched sample
- bar chart of co-morbidities from both the matched and unmatched sample.

Histograms illustrating the baseline characteristics of the intervention and matched control groups in more detail will be produced and reviewed by the IAU team. If histograms are to be included in the report, then these will be re-banded if necessary to ensure minimum cell counts of above 10.

Addendum

This section has been added to the original statistical analysis protocol (SAP) document to provide further clarifications and information on modifications to the original SAP. The original SAP was agreed in January 2018 with this section added in April 2019.

Definition of study cohort, page 13

In addition to the exclusion criteria described on page 13 of this document, the following exclusion criterion was applied prior to the analysis:

- Residents with no valid Lower Layer Super Output Area (LSOA) geographic area – this variable was used to determine a patient's level of socio-economic deprivation.

Move-in date, index dates and follow-up period, page 18

Here we provide further clarification on how index dates were assigned to potential control residents. For some potential control residents two records with different potential index dates were created to mimic the start of the intervention.

Identifying control group, page 22

Due to the low number of potential control residents, matching on cohort was not done by exact matching. We limited the number of times a record could be reused in the matching to three.

Statistical analysis, page 24

Modelling options also included adjusting for only a list of 'core' variables and those variables considered most predictive of the outcome, as the low number of events for some of the outcomes may otherwise have led to over-parameterized models. The most predictive variables were identified by running a lasso regression and where possible included all core variables as a minimum.[34] Variables considered 'core' were: age; gender (male or not male); IMD quintile; type of care home; cohort; Charlson Index; number of Elixhauser co-morbidities; number of frailty co-morbidities; and hospital use in year prior to index date. The number of beds in the care home was not included in the regression analysis as this variable had an atypical distribution that would have led to the overestimation of the intervention effect. Crude rates of hospital use at care home level in the year prior to the start of the intervention were not included in the regressions as due to the low number of care home residents identified we did not consider the care home rates to be robust enough. Modelling options did not include multilevel models with care home as a random effect.

Tables and figures for reporting statistical results

Tables

The following tables will be provided by the IAU at a minimum:

- a table showing the unadjusted estimates of treatment effect for the intervention and matched control groups:
  - for binary outcomes, the number and proportion in each group
  - for count data, the number of events and person time of exposure
  - for continuous data, the mean and standard error
  - the size of the measure effect (eg odds ratio, rate ratio, hazard ratio or mean difference) together with a 95% confidence interval
  - for a difference-in-difference type analysis, summary results in each time period, their difference and the difference between groups over time

- a table showing the adjusted results:
  - the size of the adjusted measure together with a 95% confidence interval and p-value
  - all adjustment variables with, in some cases, the relevant effect sizes and 95% confidence intervals.

Figures

The following figures would be a minimal requirement:

- forest plot showing the crude and adjusted results for each outcome measure.

Tables and figures for sensitivity analyses

The following tables may be produced:

- a table for showing study length and reasons for censoring (leaving the study)
- a table of the results of all tested regression models to show if the effect size and significance of the selected best model were sensitive to the choice of model.

Data cleaning and validation checks

Data cleaning and data validation checks will be performed by the IAU Data Management Team, in accordance with the IAU Data Management Quality Assurance Process. Where appropriate, a summary of the results can be produced.

Data flow diagrams

The following diagrams will be produced:

- data linkage process diagram
- data flow diagram of study cohort selection.

Acknowledgements

We are grateful for the support of the Wakefield CCG team. We are also grateful to Noemi Kreif and Jonathan Stokes for their peer review of an earlier version of this statistical analysis plan.

References

1. Connecting Care Wakefield District. Connecting Care+ [webpage]. *Connecting Care Wakefield District*; 2017 (https://connectingcarewakefield.org).
2. Jacques P. *Connecting Care Wakefield District: Wakefield Care Home Vanguard* Wakefield Public Health Intelligence; 2017.
3. NHS England. Wakefield Connecting Care Multispecialty Community Provider [webpage]. 2017 (www.england.nhs.uk/ourwork/new-care-models/vanguards/care-models/community-sites/west-wakefield).
4. Royal College of Physicians. [webpage] (www.rcplondon.ac.uk).
5. Mcdonach E, Mohammed MA. Yorkshire Ambulance Service: Frailty & Falls Pilots Project Findings from an External Developmental Evaluation – August 2017. Vol. 44. 2017.
6. Age UK Wakefield District. Pull Up A Chair [webpage]. Age UK Wakefield District; 2017 (www.ageuk.org.uk/wakefielddistrict/about-us/news/articles/2017/pull-up-a-chair).
7. About Portrait of a Life: South West Yorkshire Partnership NHS Foundation Trust [Internet]. [cited 25 September 2017] (www.southwestyorkshire.nhs.uk/2017/05/19/portrait-of-a-life-e-learning-helping-to-look-at-the-past-the-present-and-the-person).
8. Tai D, Dick P, To T, Wright JG. Development of pediatric comorbidity prediction model. *Arch Pediatr Adolesc Med*. 2006; 160(3): 293–9 (http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=16520449).
9. Jacques P. 07b-appendices-Care-Home-Evaluation-Appendices-2016-2017 [Internet]. 2017 [cited 20 September 2017]. Available on request.
10. Lloyd T, Wolters A, Steventon A. The impact of providing enhanced support for care home residents in Rushcliffe. The Health Foundation; 2017 (www.health.org.uk/sites/health/files/IAURushcliffe.pdf).
11. Department of Health. KH02 Quarterly bed availability and occupancy – Data definitions [webpage]. *Department of Health*; 2010 (www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy).
12. Operational Research and Evaluation Unit NE. *Evaluation strategy for new care model vanguards.* 2016.
13. Care Quality Commission. The state of health care and adult social care in England in 2012/13. CQC; 2013 (www.ncbi.nlm.nih.gov/pubmed/21727839).
14. Care Quality Commission. *The state of health care and adult social care in England. Technical Annex 1: Avoidable admissions.* CQC; 2013.
15. Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ*. 2006; 333(7563): 327.
16. Stukenborg G, Wagner D, Connors A. Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Med Care*. 2001; 39(7): 727–39 (www.ncbi.nlm.nih.gov/pubmed/11458137).
17. Preen DB, Holman CDAJ, Spilsbury K, Semmens JB, Brameld KJ. Length of comorbidity lookback period affected regression model performance of administrative health data. *J Clin Epidemiol.* 2006; 59(9): 940–6.
18. Lee DS, Donovan L, Austin PC, Gong Y, Liu PP, Jean L, et al. Comparison of Coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care.* 2011; 43(2): 182–8.

19. Soong J, Poots AJ, Scott S, Donald K, Woodcock T, Lovett D, et al. Quantifying the prevalence of frailty in English hospitals. *BMJ Open*. 2015; 5(10) (http://bmjopen.bmj.com/content/5/10/e008456).
20. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis*. 1987; 40(5): 373–83.
21. Sekhon JS, Grieve RD. A matching method for improving covariate balance in cost-effectiveness analyses. *Health Econ*. 2012; 21(6): 695–714.
22. Sekhon J. Multivariate and Propensity Score Matching. *J Stat Softw*. 2011; 42(7): 52.
23. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009; 28(25): 3083–107.
24. Normand S-LT, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol*. 2001; 54(4): 387–98.
25. Imai K, King G, Stuart EA. Misunderstandings between experimentalists and observationalists about causal inference. *J R Stat Soc, Ser A (Stat Soc)*. 2008; 171: 481–502.
26. West SG, Thoemmes F. Campbell's and Rubin's perspectives on causal inference. *Psychol Methods*. 2010; 15(1): 18–37.
27. Steventon A, Bardsley M, Billings J, Georghiou T, Lewis G. *An evaluation of the impact of interventions on hospital use*. Nuffield Trust; 2011.
28. Steventon A, Grieve R, Sekhon JS. A comparison of alternative strategies for choosing control populations in observational studies. *Heal Serv Outcomes Res Methodol*. 2015; 15(3–4):157–81.
29. Bottle A, Aylin P. Comorbidity scores for administrative data benefited from adaptation to local coding and diagnostic practices. *J Clin Epidemiol*. 2011; 64(12): 1426–33 (http://dx.doi.org/10.1016/j.jclinepi.2011.04.004).
30. Wennberg JE, Staiger DO, Sharp SM, Gottlieb DJ, Bevan G, McPherson K, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. *BMJ*. 2013; 346: f549 (www.bmj.com/content/346/bmj.f549).
31. STROBE (Strengthening the Reporting of Observational studies in Epidemiology). *STROBE Statement – checklist of items that should be included in reports of observational studies*. 2014 (www.strobe-statement.org/index.php?id=strobe-home).
32. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Med*. 2015; 12(10): e1001885.
33. The Health and Social Care Information Centre. *Hospital Episode Statistics (HES) analysis guide*. HSCIC; 2015 (http://content.digital.nhs.uk/media/1592/HES-analysis-guide/pdf/HES_Analysis_Guide_Jan_2014.pdf).
34. Friedman J, Hastie T, Tibshirani R, et al. Package 'glmnet'. 2018 (https://cran.r-project.org/web/packages/glmnet/glmnet.pdf).