

## 1 Title

2 Full: Prevalence, risks and outcomes of frailty in People Experiencing Homelessness: a protocol for  
3 secondary analysis of Health Needs Audit data

4 Short: A protocol for quantifying prevalence and associated factors of frailty in people experiencing  
5 homelessness

6

## 7 Authors

8 Jo Dawes<sup>1\*</sup>, Prof Kate Walters<sup>2</sup>, Dr Rachael Frost<sup>2</sup>, Dr Emmanouil Bagkeris<sup>3</sup>, Dr Alexandra Burton<sup>4</sup>,  
9 Debra Herzberg<sup>5</sup>, Prof Andrew Hayward<sup>1,6</sup>

10 1. Collaborative Centre for Inclusion Health, University College London, UK

11 2. Department of Primary Care and Population Health, University College London, UK

12 3. Great Ormond Street Institute of Child Health, University College London, UK

13 4. Department of Behavioural Science and Health, University College London, UK

14 5. Homeless Link, UK

15 6. United Kingdom Health Security Agency, UK

16 \* Corresponding author

17 Email: [Joanna.dawes@ucl.ac.uk](mailto:Joanna.dawes@ucl.ac.uk) (JD)

18

## 1 Metadata

## 2 Funding

3 JD is funded by a National Institute of Healthcare Research (NIHR) Doctoral Research Fellowship  
4 [NIHR302285]. The funder's role in this research is to provide full fellowship funding for the lead  
5 researcher (JD), as part of a wider research training programme and completion of PhD.

## 6 Competing interests

7 The authors declare that they have no conflict of interest.

## 8 Data availability

9 This study used third party data made available under licence that the author does not have  
10 permission to share. Requests to access the data should be directed to Homeless Link at:  
11 [www.homeless.org.uk](http://www.homeless.org.uk) or email: [info@homelesslink.org.uk](mailto:info@homelesslink.org.uk).

12

## 13 Associated content

14 STROBE checklist

## 1 Abstract

2 Frailty describes a health state related to ageing where people become less resilient to health  
3 challenges and more likely to have adverse outcomes if they become unwell. People experiencing  
4 homelessness (PEH) are known to have poor health, with research suggesting that many become  
5 frail at a younger age than the general population. Previous research using small-scale primary data  
6 collection suggests that the prevalence of frailty in homeless populations varies widely (16-55%),  
7 with variations in sample sizes and settings partially accounting for differences in current estimates.  
8 The prevalence, risks, and outcomes of frailty in PEH are poorly understood. We propose to carry out  
9 a secondary analysis of existing health survey data collected from 2,792 PEH. This will involve  
10 creating a Frailty Index (FI) to identify frail people within the dataset. Regression analyses will be  
11 used to identify associations between potential risk factors and outcomes of frailty in this  
12 population. This protocol will: 1) Outline the creation of a FI to assess the frailty prevalence within a  
13 dataset of health information collected from a cohort of PEH and 2) Describe proposed methods of  
14 regression analysis for identification of associations between frailty and risks factors/outcomes of  
15 frailty in the cohort of PEH within the dataset. The processes described in this paper can inform  
16 future development of FIs in other datasets. It is expected that the FI created will be an appropriate  
17 and robust method for identifying frailty in a cohort of PEH and results of the secondary data  
18 analysis will provide a more robust estimate of the associations between frailty and risk  
19 factors/outcomes.

20

21

## 1 Introduction

## 2 Background

3 Frailty is a health state related to aging where multiple body systems gradually lose their reserves  
4 and resilience, which increase risks of adverse outcomes (e.g. prolonged recovery time, disability or  
5 death) and health service usage (e.g. hospitalisation, attendance to accident and emergency or  
6 greater primary care usage).(1) There are two theoretical approaches to assessing frailty. One is the  
7 phenotypic model(2) where frailty is thought to be represented by the presence of up to five  
8 characteristics of frailty (unintended weight loss, reduced muscle strength, reduced gait speed, self-  
9 reported exhaustion and low energy expenditure). The alternative model is that of cumulative  
10 deficit(3) whereby as a person accumulates health deficits (e.g. loss of hearing, low mood, falls,  
11 various diseases) which can happen as they age, the deficits combined then form a “frailty index”. It  
12 is the model of cumulative deficit that is applied in this research protocol, with the FI recognised as  
13 strongly associated with adverse outcomes including mortality and disability.(4)

14

15 Homelessness can be considered as a continuum, at one extreme defining people experiencing  
16 homelessness (PEH) solely by their absence of shelter, at the other extreme, is the inclusive  
17 definition that a PEH can be someone without access to shelter that meets the basic criteria  
18 considered essential for health and social development. For the purpose of this study, PEH are  
19 defined as fitting the FEANTSA definitions of “roofless” (e.g. no fixed abode, living in a public space)  
20 and “houseless” (e.g. living in hostel, refuge, shelter, temporary accommodation).(5)

21

22 PEH have substantially poorer health than the general population, with cohort studies showing 3-6  
23 fold increased mortality risks, high levels of chronic illness and mental health problems.(6, 7) In the  
24 general population of England, 8.1 [95% CI 7.3-8.3]% of people aged >50 years are estimated to be

1 frail.(8) Research suggests prevalence of frailty amongst PEH ranges between 16% and 55%,(9-11)  
2 often presenting at a younger age. However, frailty in PEH is poorly understood, with studies to date  
3 exploring it using primary data collection in specific settings, such as a single homeless hostel, and  
4 with relatively small sample populations (range n=33-247).(9, 10) PEH are known to have a high  
5 need for and usage of healthcare, which frailty may contribute to, so it is important to identify frailty  
6 in this population. To build on the existing literature, this research proposes to explore frailty using  
7 secondary analysis of a dataset of health needs information collected from PEH. This will allow for  
8 inferences to be made from a far larger sample than previous research has achieved. Moreover,  
9 within this dataset, a broad spectrum of types of homelessness is represented, thus allowing for  
10 more accurate assessment of prevalence of frailty and associated risk or outcomes in PEH and  
11 therefore more generalisable results.

12

### 13 Research questions

14 This protocol describes the plan for a cross-sectional analysis of survey data that will address the  
15 following research questions:

- 16 - What is the prevalence of frailty amongst PEH in England?
- 17 - What risk factors and/ or outcomes are associated with frailty in this this population?

18

### 19 Hypotheses being tested

20 The Homeless Link dataset contains a broad spectrum of data variables which could be explored for  
21 associations with frailty. To avoid data dredging(12) we have conducted preliminary exploration of  
22 potential variables with a panel of experts (including researchers and academics with experience of  
23 frailty and FI development, medical doctors working in general practice, elderly care, drug and  
24 alcohol services, homelessness or frailty) familiar with frailty and/ or homelessness and patient and  
25 public involvement and engagement activities (PPIE) with PEH and the staff who care for them to

1 explore, discuss and further refine our hypotheses. Table 1 outlines the hypotheses that will be  
 2 tested:

3

4 Table 1: Hypotheses and null hypotheses to be tested by the proposed research

	<b>Hypothesis</b>	<b>Null hypothesis</b>
<b>Prevalence of frailty in PEH</b>	Prevalence of frailty in PEH in England is greater than prevalence reported in the general population (based on comparative literature)	Prevalence of frailty in PEH in England is the same as the prevalence reported in the general population (based on comparative literature)
	The mean age of PEH identified as frail in England is younger than the mean age of people identified as frail in the general population (based on comparative literature)	The mean age of PEH identified as frail in England is the same as the mean age of people identified as frail in the general population (based on comparative literature)
<b>Risks factors of frailty in PEH</b>	Frailty in PEH is positively associated with addiction (e.g. problematic drug or alcohol use)	There is no association between frailty and addiction (drugs or alcohol) in PEH
	Frailty in PEH is positively associated with impaired cognitive development (e.g. learning disability or autistic spectrum disorders)	There is no association between frailty and impaired cognitive development (e.g. learning disability or autistic spectrum disorders)
	Frailty in PEH is positively associated with sociodemographic factors (e.g. prison, sex work, childhood experience of the care system)	There is no association between frailty and sociodemographic factors (e.g. prison, sex work, childhood experience of the care system)
<b>Outcomes of Frailty in PEH</b>	Frailty in PEH is positively associated with health care usage (e.g. hospital admission, ambulance, accident and emergency use and General practice)	There is no association between frailty and health care usage (e.g. hospital admission, ambulance, accident and emergency use and General practice)
	Frailty in PEH is negatively associated with preventative health care (e.g. vaccination)	There is no association between frailty and preventative health care (e.g. vaccination)

5

## 6 **Aims and contribution**

7 The aim of this research is to use an existing dataset of health needs audit data, collected from PEH  
 8 using a one-to-one, interviewer-led survey, to calculate frailty prevalence and identify frailty risk  
 9 factors and outcomes of frailty in that population. This study will contribute in two ways. Firstly, it  
 10 will provide a calculation of prevalence of frailty in this population from a far larger dataset than any  
 11 previous research has explored- thus improving accuracy and precision of this calculation with the  
 12 potential for supporting investment in services to address frailty in PEH. Secondly, it will explore risk

1 factors for and outcomes of frailty, potentially supporting targeting of interventions to those most at  
2 risk of frailty or of poor outcomes of frailty. The intention is to provide information that can help  
3 direct resources to prevent and address frailty in PEH rather than to determine causality.

4

## 5 Materials and Methods

### 6 Design and setting

7 The design of this study is an analytical, cross-sectional study, using secondary analysis of health  
8 needs audit survey data collected from PEH. The primary dataset was collected from PEH in England  
9 between 2015 and 2022. Settings for data collection included homeless accommodation services,  
10 outreach, day centres, night shelters, and specialist support services.(13)

11

### 12 Dataset

13 The data comprises health information (demographics, physical health, mental health, drug and  
14 alcohol use, health service usage, wellbeing and preventative healthcare)(13) collected by a  
15 partnership of service providers, recruited by the audit commissioner, homelessness charity  
16 Homeless Link. It is important to note that although this survey tool contains substantial  
17 health information relevant to frailty, it was not designed with frailty assessments in mind.  
18 The data were collected between 2015 and 2022 from people across England who were either  
19 homeless at the time of the survey or had experienced homelessness in the preceding 12 months.  
20 Data were collected in accordance with GDPR (2018) by trained individuals using the Homeless  
21 Health Needs Audit (HHNA), an interviewer administered health survey tool. The advantage of using  
22 this dataset is that all participants were either currently homeless or had recently experienced  
23 homelessness leading up to the survey. This is important because there are substantial problems

1 with identifying PEH from other, more commonly used national datasets of health information. For  
2 example, PEH may be difficult to identify due to non-disclosure of homelessness status, providing an  
3 address of a relative/ friend, or stating a hostel or temporary accommodation as their address. To  
4 ensure the research team do not have any prior knowledge of any patterns or summary statistics of  
5 the data at this stage of planning, no reports of information relating to frailty have been shared with  
6 the research team prior to study design and protocol registration. Data was accessed by the UCL  
7 research team on 25<sup>th</sup> July 2023. Data was sufficiently anonymised that no information which made  
8 individual participants identifiable could be accessed.

9

## 10 Sample size, inclusion, and exclusion criteria of the HHNA dataset

11 The sample contains data from 2,792 participants who have consented to their anonymous data  
12 being collected, analysed, and reported. Inclusion criteria are:

- 13 • Participant living in England
- 14 • Participant homeless or experienced homelessness within preceding 12 months of survey
- 15 • Participant provided survey responses to at least 20/ 43 (46.5%) survey questions

16

17 Sample size is determined by the number of people surveyed using the HHNA tool and who have  
18 provided their consent for their anonymous data to be used for research purposes. The data were  
19 collected before the current protocol was conceived, so the sample size can only be based on the  
20 available data.

21

## 22 Participants and recruitment

23 Organisations supporting PEH in England partnered with Homeless Link to collect HHNA data. This  
24 involved training partner organisation staff in how to administer the HHNA survey tool in their local



1 area. The recruiting organisations had detailed understanding of the services and landscape of  
2 homelessness in their local area, so could target and reach areas when PEH would be, thus  
3 optimising opportunities for anyone eligible to participate to be surveyed. Anyone experiencing  
4 homelessness within 12 months of the time of interview was invited to participate in the HHNA  
5 survey. Informed consent was obtained by providing potential participants with an information  
6 sheet outlining the study and how their data would be used. All participants agreeing to be surveyed  
7 provided written consent to participate and for their data to be analysed and reported on. Surveys  
8 took approximately 30 minutes to complete and were usually undertaken in 1:1 case work  
9 sessions.(13)

10

## 11 How materials will be selected and used

12 For the purposes of research planning, Homeless Link shared the HHNA survey tool with our  
13 research team, allowing for creation of a variables and response list (supplementary file 1). Although  
14 Homeless Link descriptive reports of HHNA data are published in the public domain, no secondary  
15 analysis has previously been carried out in this cohort relating to frailty.

16

17 A Data Sharing Agreement was created between Homeless Link (Data Providers) and UCL (Research  
18 Project lead) prior to sharing of data. Also, to comply with the consent provided by participants, all  
19 researchers who will access the data were awarded honorary contracts of employment by Homeless  
20 Link. The dataset was anonymised by Homeless Link before it was shared with UCL, so no names,  
21 unique identification numbers, participants' dates of birth or any other personally identifying  
22 information was shared. Data will be used to summarise the characteristics of the survey population,  
23 create a frailty index, calculate the prevalence of frailty, and identify associations between risk and  
24 outcome variables and frailty. The HHNA dataset has been shared with the research team in its raw  
25 data form, with cleaning currently being undertaken.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

## Research processes

The dataset will be cleaned to remove variables irrelevant to the scope of this study. Anomalies in categorisation will be identified and incorrect or incorrectly formatted data will be removed or corrected.

Descriptive statistics will be used to summarise the characteristics of the sample population, including: Age, gender, sexual orientation, ethnicity, immigration status, employment, history of prison, domestic violence, being in the armed forces, local authority care, disability, housing status, mental health conditions, cognitive condition, alcohol consumption, drug use, smoking, GP registration, average meals per day consumed, perceived health, physical health conditions and healthcare utilisation.

## Creation of Frailty Index

Procedures for creating a FI for datasets is outlined by Searle et al(14) and Theou et al(15) who describe step-by-step procedures for FI creation. It is recommended that inclusion of variables in an FI, is based on the following criteria:(14, 15)

- 1) Variable must be health-related(14, 15)
- 2) Variable must increase with age(14, 15)
- 3) Variable must not saturate too early (e.g. not be universal in the adult population by midlife)(14)
- 4) The total included variables must cover a range of physiological systems and processes.(14, 15)
- 5) Variable must be present in at least 1% of study population(14, 15)

- 1                   6) Variable must have no more than 5% missing data(14, 15)
- 2                   7) Highly correlated screen variables ( $r>0.95$ ) will be excluded by removing the variable
- 3                   with the lowest response rate (15)

## 5   Modified Delphi for FI creation

6   In recognition that there is a degree of subjectivity regarding which variables from the HHNA dataset  
7   should/ should not be included in the FI, a modified Delphi process was adopted for creation of the  
8   FI. The Delphi method is a structured method of developing consensus amongst a panel of experts  
9   and is an accepted approach for this purpose.(16) A panel of ten experts working in the field of  
10  frailty and/or inclusion health in the UK or Canada were invited by the research team to contribute  
11  (six agreed to participate). All who agreed to participate were then invited via email, then met online  
12  for a preliminary meeting to determine the nature of their contribution, including agreeing how they  
13  would inform the selection of variables. The adapted Delphi method is outlined in Fig 1.

14

15  Fig 1: Outline of modified Delphi Method for identification of variables

16

17  A spread sheet of all 117 potential variables was circulated amongst the expert panel, with each  
18  member independently reviewing the variables and stating on the spreadsheet whether each  
19  variable should be included in either the FI, based on the stated inclusion criteria (e.g. whether they  
20  are health related; increase in prevalence with age; and whether they did not saturate early) or  
21  considered a risk factor/ outcome for the secondary data analysis. A deadline was set for return of  
22  completed spreadsheets to the lead author (JD), whereupon she collated responses, identifying  
23  areas of consensus and disagreement. A second meeting was convened to share the findings of the  
24  first round and to discuss/ agree any differences of opinion with the panel. After two Delphi rounds

1 with the expert panel, of the potential 117 variables in the HHNA, 41 were viewed as potentially  
2 viable for inclusion in a FI for this dataset. Searle et al suggest that 30-40 variables are optimal.

3

4 After the dataset is cleaned and reviewed, missing data will be identified, the reason for missingness  
5 considered and how missing data may impact inclusion of each variable in the FI and data analysis.

6 Following completion of data scoping, the lead author will determine if the proposed variables cover  
7 a range of physiological systems and processes. The final variable list will be circulated with the  
8 expert panel, providing an opportunity to review the FI as a whole and comment on its

9 appropriateness to identify frailty in a homeless cohort, including any concerns they have about its  
10 likely accuracy.

11

## 12 **Determining frailty using FI**

13 To calculate each participant's FI score, the sum of the deficit scores will be divided by the total  
14 number of deficits measured:

$$15 \quad FI = \frac{N \text{ of deficits present within the individual}}{N \text{ of deficits measured for that individual}}$$

16 Although any FI score ranges from 0 to a theoretical maximum of 1, it is noted that they consistently  
17 show a sub-maximal limit at around 2/3 of the deficits considered. For example, if a frailty index is  
18 composed of 60 items, the most items that any participant will likely indicate as being relevant to  
19 them is up to 40.(14, 17) For the purposes of analysis, there is debate regarding the determination  
20 of a cut off value of a dichotomous "frail" or "not frail",(18) whereas keeping the FI scores as a  
21 continuous variable, allows for analysis to determine how frailty scores change with associated risk  
22 factors or outcomes. However, for identifying prevalence of frailty, a cut off figure is required, with  
23 different studies citing different values. A systematic review of prevalence of frailty in community-  
24 dwelling older people described a FI with a cut off of 0.3.(19) Whereas Rockwood et al suggest 0.25

1 is generally considered of value.(20) The Fried Phenotypic model classifies people as “not frail”,  
2 “pre-frail” and “frail”.(2) To reflect these subgroups, Song et al(21) proposed scores of  $FI \leq 0.08$  as  
3 ‘non-frail’,  $FI \geq 0.25$  as ‘frail’, and the rest as ‘pre-frail’.(21) For the purposes of this research we will  
4 follow the work of Song et al(21) to determine non-frail, pre-frail and frail people in the dataset,  
5 since we believe the inclusion of a pre-frail category is likely to be useful in a cohort of PEH, whose  
6 age range is younger than many groups where frailty has been researched (the predominant age  
7 range is 25-54 years in the HHNA dataset). This decision is supported by the work of Gordon et al(18)  
8 whose systematic scoping review indicated 0.25 as the most commonly used score to determine  
9 frailty.

10

#### 11 **Coding individual variables in FI**

12 All binary variables will be coded using ‘0’ for no deficit and ‘1’ for presence of a deficit. Ordinal and  
13 continuous variables will be accommodated by grading the continuum into a rank or score between  
14 0 (where there is no deficit) and 1 (where there is a deficit). For example, a continuous score (such  
15 as, rate your health out of 100, where 100 is best health and 0 is worst health) a score of 100 will be  
16 rated 0, and a score of 50 would be rated 0.5 etc.

17

#### 18 **Analysis Plan for identification of associated risks and outcomes of frailty**

##### 19 **Public and Patient Involvement and Engagement (PPIE)**

20 In recognition that people who are affected by research should have a say in how it is designed,  
21 undertaken and disseminated,(22) people with lived experience of homelessness and frailty and  
22 those involved in their care, have contributed to a series of PPIE activities informing this research to  
23 date. Firstly, in June 2021, an online session was run in partnership with “Expert Focus”, an  
24 organisation that supports PEH to engage with research activities, by minimising the impact of digital  
25 exclusion. During the session we discuss conceptual ideas for the research proposal and funding

1 application and the attendees provided feedback on our Plain English Summary of the research,  
2 which had been circulated in advance.

3

4 After successfully securing of research funding, two further workshops were carried out in  
5 November 2022 (one with PEH and one with clinicians working with PEH) which allowed for more  
6 detailed consideration of this research. A recommendation from PEH that emerged addressed the  
7 importance of language of frailty, with attendees suggesting that using words such as “wellbeing” or  
8 “general health” might be more effective for communicating research with PEH, stating that the  
9 term “frailty” carried unappealing connotations of weakness and illness that PEH may not wish to  
10 identify with.

11

12 Finally, review of the survey tool and dataset highlighted many variables which could be considered  
13 appropriate for exploration of associations with frailty. It is considered poor research practice to  
14 simply explore all potential variables for associations, as this risks generating misleading results.(12)

15 So, in October and November 2023, to understand which variables should be prioritised for  
16 statistical exploration of association with frailty, two PPIE workshops (one in person and one via  
17 video call) were held with health and social care outreach clinicians working with frail PEH.

18 Additionally, a day of outreach PPIE was held with PEH who themselves were frail to discuss these.

19 Through informal and accessible discussions about the key areas of health in the HHNA, the research  
20 team were able to refine their hypotheses and variables of interest.

21

## 22 [Statistical models](#)

23 Descriptive statistics (proportions, means, standard deviations, medians and interquartile ranges  
24 depending on whether a variable is categorical or continuous numeric—normally distributed or

1 skewed) will be used to summarise the population characteristics and state the prevalence  
2 (proportion of sample population) of frailty and pre-frailty in the cohort. The creation of an FI for this  
3 data allows for it to be expressed both continuously and categorically (not frail, pre-frail, frail).  
4 Univariable and multivariable logistic regression analysis will be used to assess the factors associated  
5 with frailty when describing associations with frailty expressed categorically and linear regression  
6 will be used when expressing frailty on a continuous scale. Regression coefficients and 95%  
7 confidence intervals will be reported. In the cases of some variables, it might be debatable whether  
8 they are a risk factor, an outcome or risk factor and/or an outcome of frailty. For example, if a  
9 relationship between alcohol or drug use and frailty is identified, the cross-sectional nature of this  
10 research means it will not be possible to determine the direction of that relationship, i.e., does  
11 increased alcohol or drug use issues result in increased frailty? Or does the presence of frailty cause  
12 someone's alcohol or drug use to be altered? Directional acyclic graphs (DAGs) will be used to  
13 identify potential confounding factors that should be accounted for in the multivariable models  
14 designed to test associations between different relevant variables and frailty amongst PEH.

15

## 16 Transformations

17 Data transformation is a technical statistical process that can be used to make data better meet  
18 distributional assumptions for parametric tests. It is expected that not all variable data will be  
19 normally distributed, so transformations may be required. Transformations will be guided using the  
20 Tukey's Transformation Ladder.(23)

21

## 22 Inference criteria

23 Linear and logistic regression analysis will be used to identify associations between dependent and  
24 independent variables. In the case of risk factor analysis (for example past adverse experiences) the  
25 exposure would be the independent variable and frailty would be the dependent variable.

1

## 2 Data exclusion

3 Assessment for exclusion of variables was carried out based on the most recent iteration of the  
4 survey tool. Prior to receiving the dataset, it was possible to identify 14 variables that would not be  
5 useful to this research, such as variables with limited responses that only applied to participants  
6 limited by sex and age range (e.g. uptake of cervical smear, uptake of mammogram, access to free  
7 sanitary products) and therefore were not shared by Homeless Link (supplementary file 2). Following  
8 creation of the FI a further 22 variables were identified through the PPIE process with clinicians that  
9 were not relevant to risk factor or outcome analysis, and so could be excluded. Once the dataset is  
10 cleaned, we will identify if any further data needs to be excluded. For example, some variables  
11 within the HHNA were only collected in more recent iterations of the survey tool. In these cases,  
12 where there are substantial levels of missing data, the variables may need to be excluded from the  
13 frailty index and/or the regression analysis.

14

## 15 Missing data

16 Once the dataset is cleaned and basic summaries of variables and observations made, we will gauge  
17 the volume and patterns of missing data. Where possible, we will determine the likely reason for  
18 missing data (completely missing at random, missing at random or missing not at random) by  
19 observing patterns in the missingness and communicating with our colleagues at Homeless Link who  
20 have detailed insights into the data collection processes. Data might be missing at random if some  
21 data collectors failed to ask all questions on the survey tool, or missing not at random if some  
22 variable questions do not exist in survey waves or responses to contentious topics, such as sexuality  
23 or connection to criminal activity. There are established strategies to deal with missing data. For  
24 example, with variables selected to be in the FI, it may be possible to address missing data by coding  
25 the variable to binary, with 1= known to have condition x and 0= not known to have condition x,



1 where 0 will also be the code for missing data. However, a limitation of this could be an  
2 underestimation of prevalence of conditions contributing to the FI, thus potentially underestimating  
3 levels of frailty in the cohort. After identifying all complete cases from the dataset, we will determine  
4 if multiple imputation, complete case analysis or coding “missing” as a variable category, may be an  
5 appropriate strategy depending on the extent, patterns, and causes of missingness.

6

## 7 Exploratory analysis

8 Exploratory descriptive statistics (frequency, calculation of means and Standard Deviations) will be  
9 reported to show the characteristics of the participant sample and identify if any of the variables are  
10 sufficiently uncommon within the dataset (>1%) or have >5% missing data, leading to removal from  
11 the FI. When the FI is finalised, prevalence of frailty in the population will be calculated by selecting  
12 all participants with a minimum of 80% complete observation data (Theou et al(15) state that a  
13 frailty index score should not be calculated for individuals missing more than 20% of the frailty index  
14 items) and determining their individual frailty score. Once calculated, the frailty scores for each  
15 individual participant will become an additional continuous variable that can be used, to carry out  
16 linear regression to explore the associations between frailty in PEH and risks and outcomes.

17

## 18 Data management plans

19 The data comprises anonymised human health information, gathered from people who were  
20 homeless at or within 12 months of the data being collected. Data is currently securely stored and  
21 processed by Homeless Link (data owner) and UCL (research group, data controller and processor). A  
22 Data Sharing Agreement (DSA) between Homeless Link and UCL is place. The data is stored in the  
23 UCL Data Safe Haven (DSH), a secure server suitable for the safe storage and analysis of healthcare  
24 data. The dataset contains sensitive data (special category data, e.g. ethnicity, health and sex life or  
25 sexual orientation)(24) however, the data were anonymised by Homeless Link before transfer to

1 UCL, so the risk of deductive identification of any individual is extremely low. The UCL policy on  
2 retention and deletion of data states that data must be stored for 10 years following completion of  
3 the project (e.g., publication of final report). After this point the data will be deleted from the UCL  
4 DSH.(25)

5

## 6 Ethical considerations and declarations

7 Ethical approval for this study was provided by the Research Ethics Committee, University College,  
8 London [Project ID: 25071/001] on 7 June 2023. This secondary analysis of existing data, originally  
9 collected by a partnership of homelessness service providers recruited by the audit commissioner,  
10 Homeless Link, then anonymised by Homeless Link prior to sharing with UCL carries very low risk  
11 ethically. Because the data has already been collected, it carries no additional time burden or cost to  
12 participants. There is no conflict of interest between the research team, Homeless Link or the  
13 participants who have provided data for the research. The presence of a robust DSA agreement  
14 between Homeless Link and UCL minimises the risk of data breaches. Additionally, in the reporting  
15 of research results, all findings will be reported using a minimum cell size of 10, with specific  
16 identifiers, such as age bands location kept broad in reporting.

17

18 When balancing the risks and benefits of this research the benefits far outweigh the risks. Although  
19 there is unlikely to be direct benefit of this research to any individual who provided their data, the  
20 benefits in the long term could highlight a health need in this otherwise vulnerable population and  
21 inform policy, practice, investment in services and targeting of services to improve the health  
22 landscape for PEH.

23

## 1 Status and timeline of study

2 At the point of protocol submission, all ethical approvals, data sharing agreements and data transfer  
3 are complete. Data review and cleaning are currently underway. For frailty index creation, the  
4 modified Delphi process with the expert panel is complete and a preliminary proposed list of  
5 variables for inclusion exists. The next phase for the FI will involve screening of these variables in line  
6 with recommendations of Searle et al(14) and Theou et al(15) to exclude any proposed variables  
7 which do not meet the inclusion criteria. PPIE activities are underway to prioritise variables for  
8 exploration of risk and outcomes of frailty. Data analysis to report prevalence of frailty and risk  
9 factor/ outcome associations is expected to be complete by May 2024.

10

## 11 Discussion

12 A strength of this proposed research is that it will provide a new level of understanding of frailty  
13 amongst PEH in England. By using a large dataset of healthcare data collected entirely from people  
14 who were homeless within a 12-month period prior to undertaking the survey, this study enables  
15 accurate representation of the population on a far larger scale than has been possible in previous  
16 small-scale studies exploring frailty. Another strength is the robust approach to development of the  
17 Frailty Index. By using an established framework for considering which variables should contribute to  
18 frailty indexes and an adapted Delphi method, undertaken in a systematic way and by inviting input  
19 from clinical and academic experts in the fields of frailty and homelessness, the development of the  
20 FI is well informed. Finally, involving, and engaging people with lived experience of homelessness  
21 and being frail and the staff who care for them in prioritising which variables should be considered  
22 for exploration of associations with frailty and stating hypotheses in advance of data analysis means  
23 that the risk of type one errors and poor statistical practices, such as data dredging are removed.

24

1 Limitations of this research must also be acknowledged. Firstly, the development of the FI for this  
2 research can only be applied to the Homeless Link dataset. So, although the principles of how the FI  
3 is developed can be mirrored by other researchers, the FI itself cannot be transferred to a different  
4 database which has not collected the variables included in the index. However, it will be made public  
5 and shared, so any future research related to frailty and using the Homeless Link database or other  
6 databases with relevant data, can harness the knowledge gained from this proposal and benefit  
7 future research. Also, this cross-sectional observational study does not have a comparison group for  
8 the assessment of prevalence of frailty, so we will make comparisons with findings of prevalence in  
9 the literature- recognising that this will likely have differences in population characteristics and how  
10 frailty was determined. These limitations will be noted in reporting. Additionally, we cannot identify  
11 temporal associations of risk factors and outcomes with frailty status, thus it is not possible to make  
12 any interpretations of the findings related to causality. It must be noted that due to data collection  
13 and anonymisation processes, it is not possible to identify repeat observations from the dataset, for  
14 example, if participants have completed the HHNA survey on more than one occasion, or how that  
15 may impact analysis. Amendments to the proposed study will be dealt with appropriately depending  
16 on the nature of the amendment. For example, amendments related to ethics will be submitted to  
17 the UCL Research Ethics Committee.

18

19 This research will have a broad dissemination plan, which will be widely accessible to a variety of  
20 audiences. We intend for PEH to learn about this research and/or engage with it in the future, by  
21 publicising it in an accessibly written format in a nationally published magazine widely viewed by  
22 PEH, such as Pavement or the Big Issue. We also intend to reach a wide range of staff working in the  
23 field of inclusion health, including those involved in commissioning of inclusion health services (e.g.  
24 Integrated Care Boards) or policy regarding PEH (e.g. OHID or UKHSA or DHSC or DLUCH) by  
25 communicating through peer reviewed publications, relevant conferences, staff training and

1 webinars. We anticipate that there will be a degree of snowballing with this approach- as staff  
2 working in Inclusion Health learn about our findings, in turn it should reach many PEH through  
3 communication in their places of accommodation, support services or healthcare. We do not  
4 anticipate any ethical issues to arise from this. By sharing our research there are important  
5 implications, including a greater understanding of the issue of frailty in PEH a topic which is widely  
6 acknowledged as important within the community of those with lived experience of homelessness  
7 and people who care for them. By quantifying the prevalence of frailty, we create an objective  
8 advocacy tool for policy change and by identifying associations with frailty, we gain insights into how  
9 interventions and future research should be targeted to address this important health issue.

10

## 11 Authors' contributions

12 JD, KW, RF and AH were responsible for design, methodology and conception of this research  
13 protocol. DH was responsible for data collection and data management. JD and AH were responsible  
14 for liaison with colleagues at Homeless Link to acquire the data for analysis. JD was responsible for  
15 funding acquisition, project administration, drafting and critically revising the article for important  
16 intellectual content. AH, KW, EB, RF and AB are all responsible for supervision of this research. All  
17 authors have approved the final version to be published and agree to be accountable for all aspects  
18 of the work in ensuring that questions related to the accuracy or integrity of any part of the work are  
19 appropriately investigated and resolved.

20

## 21 Acknowledgements

22 The authors would like to thank Sophie Boobis from Homeless Link for her commitment, time and  
23 support to this proposed research. We would also like to thank our PPIE and Delphi expert panel  
24 contributors for sharing their knowledge and expertise in the development of this research.

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  
31  
32  
33  
34  
35  
36  
37  
38

## Supporting information

- S1: Supplementary file 1: All variables collected using Homeless Link HHNA
- S2: Fig 1: Outline of modified Delphi Method for identification of variables
- S3: Supplementary file 2: HHNA variables excluded and not shared following survey tool review
- S4: STROBE checklist

## References

1. British Geriatric Society. Fit for Frailty part 1- consensus best practice guidance for the care of older people living in community and outpatient settings - a report from the British Geriatrics Society. 2014.
2. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. *The Journals of Gerontology: Series A*. 2001;56(3):M146-M57.
3. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. *Cmaj*. 2005;173(5):489-95.
4. Kojima G, Iliffe S, Walters K. Frailty index as a predictor of mortality: a systematic review and meta-analysis. *Age Ageing*. 2018;47(2):193-200.
5. Edgar B, Meert H, Doherty J. European Observatory on Homelessness: Third Review of Statistics on Homelessness in Europe. Developing and Operational Definition of homelessness. . European Federation of National Organisations Working with the Homeless (FEANTSA); 2004.
6. Lewer D, Aldridge RW, Menezes D, Sawyer C, Zaninotto P, Dedicoat M, et al. Health-related quality of life and prevalence of six chronic diseases in homeless and housed people: a cross-sectional study in London and Birmingham, England. *BMJ Open*. 2019;9(4):e025192.
7. Aldridge RW, Story A, Hwang SW, Nordentoft M, Luchenski SA, Hartwell G, et al. Morbidity and mortality in homeless individuals, prisoners, sex workers, and individuals with substance use disorders in high-income countries: a systematic review and meta-analysis. *The Lancet*. 2018;391(10117):241-50.
8. Sinclair DR, Maharani A, Chandola T, Bower P, Hanratty B, Nazroo J, et al. Frailty among Older Adults and Its Distribution in England. *The Journal of Frailty & Aging*. 2022;11(2):163-8.
9. Rogans-Watson R, Shulman C, Lewer D, Armstrong M, Hudson B. Premature frailty, geriatric conditions and multimorbidity among people experiencing homelessness: a cross-sectional observational study in a London hostel. *Housing, Care and Support*. 2020;23(3/4):77-91.
10. Brown RT, Kiely DK, Mitchell SL, Bharel M. Geriatric syndromes in older homeless adults. *Journal of General Internal Medicine*. 2012;27(1):16-22.
11. Cheallaigh C, Souza Ed, O'Gara R, Power C, Rajan M, Spigos J. Premature aging in the homeless population. *Depaul Health Initiative*. 2018.
12. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*. 2016;70(2):129-33.

- 1 13. Hertzberg D, Boobis S. The Unhealthy State of Homelessness 2022 Findings from the  
2 Homeless Health Needs Audit. London: Homeless Link; 2022.
- 3 14. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating  
4 a frailty index. BMC Geriatrics. 2008;8(1):24.
- 5 15. Theou O, Haviva C, Wallace L, Rockwood K. How to Construct a Frailty Index From an Existing  
6 Dataset in 10 Steps (Accepted, in press). Age and Ageing. 2023.
- 7 16. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its  
8 appropriateness. World J Methodol. 2021;11(4):116-29.
- 9 17. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol  
10 Sci Med Sci. 2007;62(7):722-7.
- 11 18. Gordon EH, Reid N, Khetani IS, Hubbard RE. How frail is frail? A systematic scoping review  
12 and synthesis of high impact studies. BMC Geriatrics. 2021;21(1):719.
- 13 19. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of Frailty in Community-  
14 Dwelling Older Persons: A Systematic Review. Journal of the American Geriatrics Society.  
15 2012;60(8):1487-92.
- 16 20. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine  
17 defined by frailty. Clin Geriatr Med. 2011;27(1):17-26.
- 18 21. Song X, Mitnitski A, Fau - Rockwood K, Rockwood K. Prevalence and 10-year outcomes of  
19 frailty in older adults in relation to deficit accumulation. 2010(1532-5415 (Electronic)).
- 20 22. Staley K. Exploring impact: public involvement in NHS, public health and social care research:  
21 National Institute for Health Research; 2009.
- 22 23. Tukey JW. Exploratory Data Analysis. Reading, MA: Addison-Wesley; 1977 1977.
- 23 24. Office ICs. Special Category Data 2023 [Available from: [https://ico.org.uk/for-](https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/lawful-basis-for-processing/special-category-data/)  
24 [organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/lawful-basis-for-](https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/lawful-basis-for-processing/special-category-data/)  
25 [processing/special-category-data/](https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/lawful-basis-for-processing/special-category-data/).
- 26 25. UCL. UCL Research Data Policy. In: Services URlA, (RIISG) G, editors. version 6 ed2020.
- 27

## Problem area

Identification that HHNA contained 117 variables relating to health, but determining which were most appropriate to include in an FI required clinical expertise.

## Panel members

8 researchers and clinicians with expertise in frailty/ FI development, general practice, elderly care, drug and alcohol services, or homelessness invited to participate.

6 were available and agreed to contribute.

## Delphi rounds

Round 1 (online meeting): Agreed definition of frailty, FI inclusion criteria, process for individual panel voting for variables.

Round 2 (email): Panel members submitted their votes for proposed variables for inclusion/ exclusion via email. Lead researcher collated and summarised areas of consensus/ disagreement.

Round 3 (online meeting): Discussion to reach consensus where disagreement existed for variables.

## Closing criteria

Final decision regarding variables for inclusion in FI collated by lead researcher and presented to Delphi panel members and research team for final agreement.

Figure