The Dynamics of Frailty Development and Progression in Older Adults in England (2006 – 2017): A Retrospective Cohort Profile

Carole Fogg
University of Southampton
Simon DS Fraser
University of Southampton
Paul Roderick
University of Southampton
Simon Lusignan
University of Oxford
Andrew Clegg
University of Leeds, Bradford Teaching Hospitals NHS Foundation Trust
Sally Brailsford
University of Southampton
Abigail Barkham
Southern Health NHS Foundation Trust
Harnish P Patel
University Hospitals Southampton NHS Foundation Trust, Southampton General Hospital
Vivienne Windle
University of Southampton
Scott Harris
University of Southampton
Shihua Zhu
University of Southampton
Tracey England
University of Southampton
Dave Evenden
University of Southampton
Francesca Lambert
University of Southampton
Bronagh Walsh (b.m.walsh@soton.ac.uk)
University of Southampton

Research Article

Keywords: Frailty, cohort study, adults, trajectories, computer simulation modelling, primary care, service use, electronic health records

Posted Date: June 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-547404/v1

Abstract

Background: Frailty is a common condition in older adults and has a major impact on patient outcomes and service use. Information on the prevalence in middle-aged adults and the patterns of progression of frailty at an individual and population level is scarce. To address this, a cohort was defined from a large primary care database in England to describe the epidemiology of frailty and understand the dynamics of frailty within individuals and across the population. This article describes the structure of the dataset, cohort characteristics and planned analyses.

Methods: Retrospective cohort study using electronic health records. Participants were aged ≥50 years registered in practices contributing to the Oxford Royal College of General Practitioners Research and Surveillance Centre between 2006 to 2017. Data include GP practice details, patient sociodemographic and clinical characteristics, twice-yearly electronic Frailty Index (eFI), deaths, medication use and primary and secondary care health service use. GP and patient characteristics at cohort entry are described. Planned analyses include frailty incidence, prevalence and transitions between frailty categories, predictors of transitions and trajectories, and relationships between frailty, service use and outcomes. Techniques including regression, multistate models and cost analyses will be used.

Results: The cohort includes 2,177,656 patients, contributing 15,552,946 person-years, registered at 419 primary care practices in England. The mean age was 61 years, 52.1% of the cohort was female, and 77.6% lived in urban environments. Frailty increased with age, affecting 10% of adults aged 50-64 and 43.7% of adults aged ≥65. The prevalence of long-term conditions and specific frailty deficits increased with age, as did the eFI and the severity of frailty categories.

Conclusion: A comprehensive understanding of frailty dynamics will inform predictions of current and future care needs to facilitate timely planning of appropriate interventions, service configurations and workforce requirements. Analysis of this large, nationally representative cohort including participants aged ≥50 will capture earlier transitions to frailty and enable a detailed understanding of progression and impact. These results will inform novel simulation models which predict future health and service needs of older people living with frailty.

Study registration: registered on www.clinicaltrials.gov October 25th 2019, NCT04139278.

Introduction

Frailty is a state of vulnerability to minor stressors such as infections or falls, which leads to a disproportionate decline in overall health status and increased risk of disability and death [1, 2]. Frailty can be identified and classified by severity using either a phenotype model following a physical assessment [3], or by using a frailty score or 'index' based on an accumulation of conditions or deficits, which can be derived from routinely collected data [4]. A pooled analysis has estimated the prevalence of frailty in older adults at 18% in all settings, and 12% in community-based studies [5]. Frailty is more common as age increases, but is not synonymous with ageing, with age only partly explaining frailty

trajectories [6]. Little is known about the onset of frailty and its development in middle-aged and younger old adults.

Although frailty research has expanded over the last decade, use of different measures and concepts of frailty, small cohort sizes that are not representative of the wider older population, and variable lengths of follow-up have led to uncertain conclusions on the occurrence and progression of frailty within ageing populations [7]. There is therefore a need for larger, longer cohort representative studies which include standardised methods of establishing frailty, with key covariates related to frailty risk and risk of other adverse outcomes and clear descriptions of populations and methods. An electronic Frailty Index (eFI) has been developed and validated using routine primary care data and is intended for application to routine data collected by GPs [8, 9]. The use of the eFI as a risk stratification tool is expanding and has become standard throughout primary care in England, making the eFI an appropriate measure for epidemiological studies of frailty at population level. As the eFI score can be calculated from routine primary care records, it is possible to apply the tool to the study of large datasets to generate widely understood and transferable findings [10, 11].

The cohort described in this paper is part of a larger project which aims to describe the epidemiology and dynamics of frailty in the adult population and its impact on patients and health and social care services, in order to develop simulation models which will inform the development of guidelines and tools to facilitate commissioning and service development [12]. Specifically, the objectives of the larger project are to establish a cohort with which to estimate the incidence and prevalence of frailty states in an ageing population; to describe frailty trajectories and transitions in severity over time; to explore the drivers of frailty progression; to examine the impact of frailty on service use and costs and explore associations between patient and practice characteristics and these outcomes. Analyses to meet these objectives will inform the final part of the project, in which a simulation model will be created to predict future trends in frailty and need for services, and explore ‘what-if’ scenarios proposed by health and social care professionals, commissioners, patients and the public to enable informed service planning.

This paper gives an overview of the cohort study design, the dataset structure, key variables, future data linkages and planned statistical analyses. Baseline descriptive data both for the general practices from which patient records originated and for the participants is presented.

Methods

Study Design

Retrospective cohort using routinely collected electronic health records.

Study setting

The UK has a registration-based primary healthcare system where patients are registered with a single general practice. Patients are allocated a personal lifetime unique identifier, (a National Health Service

(NHS) number), which reduces the risk of duplicate records and facilitates the linkage of primary care to other healthcare datasets. The Oxford-Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) is a network of around 5% of primary care practices in England which contribute electronic health record data voluntarily. Practices registered with the RCGP RSC have been shown to be nationally representative in terms of the population served and health outcomes [13].

Eligibility criteria

The inclusion criteria were patients (i) aged ≥50 years, (ii) registered at a general practice contributing to the Oxford RCGP RSC network database and (iii) registered at any time between 2006 to 2017. Potential duplicate patients were excluded, i.e. more than one sex present for a patient record, duplicated calendar years of data and differing birthdates in the patient record, and patients with missing or impossible birthdates. Yearly records were excluded as follows: (i) where a patient changed practice within a calendar year and had duplicate yearly records, the yearly record with the longest period was kept if continuous years of data were available, (ii) person-years of data following a gap of one year or more in the observation record, even if the patient re-registered with an RCGP RSC practice.

Data from all participants constitutes the open cohort – i.e. a cohort which included entry of new registrants such as patients turning 50 and those entering RCGP RSC practices from other areas. A closed cohort was also defined, including only eligible participants present in the cohort index year of 2006. These two cohorts will be used for different analyses during the project and allow for exploration of cohort ageing and the impact of frailty on the overall population on service use.

Data sources

Electronic health records (EHR) were extracted from the RCGP RSC dataset. Publicly available datasets were linked to the primary care data, including the Income Deprivation Affecting Older People Index (IDAOPI) 2015 [14], geographical information from the geography portal of the Office for National Statistics (ONS) [15] and workforce data for GP practices [16].

Measures

a. Electronic Frailty Index

The electronic frailty index (eFI) was used to determine the presence of frailty. The eFI includes 36 deficits across disease states, symptoms/signs, abnormal laboratory values and indicators of disability. These are used to calculate an eFI score as the number of deficits present as an equally weighted proportion of the total possible [9]. An eFI score was calculated twice yearly, on the 1st January and 1st July for each calendar year for each participant, using the same Read codes (Clinical Terms Version 3 - CTV3) as in the original derivation of the score. The eFI was categorised into frailty states as: fit (0-0.12), mild (0.13-0.24), moderate (0.25-0.36) or severe (>0.36) [9]. As the eFI is designed to be a cumulative index, reversals (for example data artefacts due to change in GP practice during follow-up) were imputed to the previous

frailty state. In the case of reversal due to reduction in medicines use and change in the polypharmacy deficit, was re-calculated on a yearly basis from available prescription information.

b. Sociodemographic variables

Sociodemographic variables include age, sex, Indices of Multiple Deprivation (IMD) quintiles and IDAOPI quintiles. The IMD is a small-area measure of socioeconomic status based on postcode, ranked nationally, which includes seven domains: income, employment, education/skills/training, health and disability, crime, barriers to housing and services, and living environment [17]. The IDAOPI is a subset of the Income Deprivation Domain, and focusses specifically on the percentage of the population aged 60 and over who receive income support, income based job seekers allowance, pension credit or child tax credit claimants aged 60 and over and their partners (if also aged 60 or over). The 2015 deprivation indices were related to the last known patient address in the dataset or, where missing, were imputed using the GP indices (3.6% of patients). Ethnicity data was maximised using a customised ontology and coded into categories (Asian, Black, White, Mixed/other) [18]. The most recent ethnicity reported in the patient record was used as the baseline ethnicity value to reduce missing values in the year of entry to the cohort. Age was categorised into four groups, reflecting groupings reported in literature relating to older adults' healthcare, and cut-offs for services reported by the study Stakeholder Engagement Group (SEG): 50-64, 65-74, 75-84 and ≥85. Whether or not a patient had been in receipt of residential care during the cohort period was ascertained by using a combination of Read codes [8] and use of a household key (11 or more patients at the same address with a median age of 50 or above) for the patient's last known address at the date of data extraction (May 2019).

c. Clinical variables

Individual eFI deficits, and additional common long-term conditions not included in the eFI but present in the Quality Outcomes Framework (QOF) with dates of onset were generated on a yearly basis for each patient. Other annual data included smoking status, influenza and pneumococcal vaccinations. All body mass index (BMI) present in the patient record were provided. Due to the differences in availability of values, a baseline BMI value was defined as the first recording in a patient's cohort entry year, or, where missing, the first value in the nearest previous year to cohort entry (up to a maximum of 2 years) or the nearest year afterwards (up to 2 years).

d. General practice variables

General practice information included the geographical region, urban/rural indicators based on the 2011 rural/urban classification (RUC11) [19], IMD and IDAOPI for the practice postcode, number of patients registered in the practice, and total practice consultations per year. The total general practitioner (GP), nurse and overall practice staff full-time equivalent (FTE) for each general practice in 2013 (the first year this information is available to be linked on practice code) was included [20]. Each calendar year of participant data was linked to a general practice identifier and dates of the participant registering and leaving the RCGP RSC practice were provided.

e. Death

The month and year of death presented here were obtained from RCGP RSC primary care data.

f. Service use

Primary care service use outcomes include number of days in a calendar year with a consultation by consultation type (administrative, face-to-face, telephone, electronic consultation, or home visit), total number of GP prescriptions and number of unique prescriptions by British National Formulary (BNF) chapter [21].

Statistical Analysis

A description of the primary care data from the open cohort is presented in this paper. The characteristics of RCGP RSC practices with participants in the cohort are described for the calendar year 2006 (first year of cohort). Age category distributions for both the open and closed cohorts were analysed and presented graphically for the calendar years 2006-2017. The reasons for exit from the cohort were summarised.

Patient sociodemographic and clinical characteristics at the year of cohort entry (i.e. for the open cohort) are described according to the four age groups, and missing data quantified.

Planned linkage to additional data sources and further analyses

The primary care data will be linked to hospital episode statistics (HES) secondary care data from NHS Digital and mortality data from the ONS via NHS Digital using an irreversibly encrypted NHS number. Data on secondary care service use using HES data including outpatient visits, emergency department attendances, hospital admissions, intensive care admission, length of hospital stay and death in hospital will be analysed. Primary care recorded deaths will be verified by cross-checking against ONS data, although primary care death data for the calendar period of this cohort has been shown to be accurate [22, 23].

The incidence of frailty per 1,000 person-years at risk (PYAR) both overall and by covariates of interest will be calculated. The prevalence of the four frailty categories in each calendar year of the cohort and the transition rates between frailty categories will also be calculated. Multistate modelling will be used to estimate transition rates between the frailty categories, time spent in each frailty state and which characteristics are associated with these transitions (using both univariable and multivariable analysis). Time-dependent Cox models will be used to examine the relationship between frailty state and binary clinical outcomes e.g. mortality, and mixed-effects negative binomial models for count-based clinical outcomes, e.g. unplanned hospitalisations and service use.

Primary and secondary care service use and prescription use will be described for participants in each of the age groups according to frailty categories, calculating total use, means and medians for each group. Accumulated costs for each person will be calculated over the study period with appropriate annual

inflation adjustment. Generalised linear mixed models will be used to estimate costs, adjusting for baseline characteristics, using bootstrap methods for sensitivity analyses. Mean costs per person for different frailty categories and other patient groupings will be estimated.

All estimates will be presented with 95% confidence intervals.

Results

The open cohort comprised 2,177,656 patients, contributing 15,552,946 person-years of data (Figure 1).

Practice characteristics

419 primary care practices distributed across England contributed to the cohort between 2006-2017 inclusive. Practices varied widely in their patient numbers and consequently their totals of yearly consultations (Table 1). Practices reflected population distributions throughout England, with 78% in urban areas and an even spread across IMD quintiles.

Table 1 Primary care practice characteristics in 2006 (n=419)

Practice characteristic	n (%)
Geographic region (n, %)	
London	64 (15.3%)
Midlands and East	90 (21.5%)
North	128 (30.6%)
South	137 (32.7%)
Rural/urban Classification	
Urban: Major conurbation	150 (35.8%)
Urban: Minor conurbation	15 (3.6%)
Urban: City and Town	163 (38.9%)
Rural: Town and fringe	69 (16.5%)
Rural: Village and dispersed	22 (5.3%)
Practice size (patients)	
median	6,858
Upper: lower quartile	4,110: 9,819
Practice size (staff - FTEs)[1]	
GPs (mean, SD)	6 (3)
Nurses[2] (mean, SD)	2 (2)
Total staff[2] (mean, SD)	14 (8)
Consultations (median, upper: lower quartile)	
Face to face	42,661 [22,426: 72,362]
Clinical administration	20,474 [5,811: 73,050]
e-consultations (data from 2017)	4 [0: 43]
Telephone	1,713 [261: 4383]
Home visits	669 [50: 1,992]
Practice Index of Multiple Deprivation quintile	
Most deprived	93 (22.2%)
2nd quintile	84 (20.1%)

3rd quintile	83 (19.8%)
4th quintile	85 (20.3%)
Least deprived	74 (17.7%)

[1] Data not available for 2 practices

[2] Not specified in 30 practices

Patient characteristics at cohort entry

The mean age of participants at cohort entry was 61 years (SD 12) (Table 2). Demographic trends with increasing age were observed, including a higher proportion of female sex, lower ethnic diversity and rural residence in the older age groups. Ethnicity data was more likely to be missing with increasing age, decreasing deprivation, male sex, urban location and for people in residential care. Patterns of indices of deprivation appeared similar across age groups, with half the cohort located in the two least deprived quintiles. Long-term conditions were more prevalent in older age groups at baseline, with the exception of depression and obesity which were more common in younger age groups (Table 3). The eFI score increased with age, as did the proportion of participants in the mild, moderate and severe frailty categories. The proportion of people with frailty at cohort entry increased from 10% in the 50-64 age group to 69% in people aged ≥85.

Table 2 Sociodemographic characteristics of participants at cohort entry

-	Age Group				Total (%)
	50-64	**65-74**	**75-84**	**≥85**	
Age group[1]	1,413,576 (64.9%)	385,474 (17.7%)	259,125 (11.9%)	119,481 (5.5%)	2,177,656
Female	698,158 (49.4%)	199,914 (51.9%)	151,462 (58.5%)	84,437 (70.7%)	1,133,971 (52.1%)
Ethnicity[2]					
Asian	52,703 (5.1%)	11,419 (4.1%)	4,521 (2.7%)	916 (1.4%)	69,559 (4.5%)
Black	29,387 (2.8%)	5,577 (2.0%)	2,350 (1.4%)	440 (0.7%)	37,754 (2.4%)
Mixed/Other	15,461 (1.5%)	2,480 (0.9%)	1,110 (0.6%)	277 (0.4%)	19,328 (1.3%)
White	937,135 (90.6%)	260,473 (93.0%)	160,063 (95.3%)	63,054 (97.5%)	1,420,725 (91.8%)
Missing[3]	*378,890 (26.8%)*	*105,525 (27.4%)*	*91,081 (35.2%)*	*54,794 (45.9%)*	*630,290 (28.9%)*
Urban	1,102,809 (78.0%)	294,247 (76.3%)	200,358 (77.3%)	91,492 (76.6%)	1,688,906 (77.6%)
Residential care	1,019 (0.1%)	1,708 (0.4%)	5,371 (2.1%)	9,121 (7.6%)	17,219 (0.8%)
IMD quintile					
1 (Most deprived)	193,552 (13.7%)	49,320 (12.8%)	34,151 (13.2%)	14,894 (12.5%)	291,917 (13.4%)
2	220,674 (15.6%)	60,287 (15.6%)	41,887 (16.2%)	19,592 (16.4%)	342,440 (15.7%)
3	280,969 (19.9%)	79,288 (20.6%)	54,244 (20.9%)	25,806 (21.6%)	440,307 (20.2%)
4	340,796	93,998	62,573	28.815	526,182

	(24.1%)	(24.4%)	(24.2%)	(24.1%)	(24.2%)
5 (Least deprived)	377,585 (26.7%)	102,581 (26.6%)	66,270 (25.6%)	30,374 (25.4%)	576,810 (26.5%)
IDAOPI quintile					
1 (Most deprived)	199,722 (14.1%)	50,167 (13.0%)	34,440 (13.3%)	15,493 (13.0%)	299,822 (13.8%)
2	217,183 (15.4%)	58,934 (15.3%)	42,894 (16.6%)	19,930 (16.7%)	338,941 (15.6%)
3	269,450 (19.1%)	76,828 (19.9%)	55,166 (21.3%)	27,233 (22.8%)	428,677 (19.7%)
4	336,857 (23.8%)	93,684 (24.3%)	62,160 (24.0%)	29,063 (24.3%)	521,764 (24.0%)
5 (Least deprived)	390,364 (27.6%)	105,861 (27.5%)	64,465 (24.9%)	27,762 (23.2%)	588,452 (27.0%)

[1] % as proportion of total cohort

[2] % as proportion of known values

[3] missing values as % of cohort

Table 3 Clinical characteristics of participants in year of cohort entry

	Age Group				Total (%)
	50-64	65-74	75-84	≥85	
eFI score					
median	0.028	0.083	0.139	0.167	0.056
Upper: lower quartile	[0: 0.083]	[0.028: 0.139]	[0.083: 0.194]	[0.111: 0.250]	[0.028: 0.111]
Frailty category					
Fit	1,273,304 (90.1%)	272,694 (70.7%)	120,357 (46.5%)	37,243 (31.2%)	1,703,598 (78.2%)
Mild	127,029 (9.0%)	94,558 (24.5%)	99,154 (38.3%)	49,192 (41.2%)	369,933 (17.0%)
Moderate	12,055 (0.9%)	16,167 (4.2%)	32,732 (12.6%)	25,360 (21.2%)	86,214 (3.4%)
Severe	1,188 (0.1%)	2,055 (0.5%)	6,882 (2.7%)	7,686 (6.4%)	17,811 (0.8%)
Long-term conditions					
Atrial fibrillation	11,359 (0.8%)	15,381 (4.0%)	23,978 (9.3%)	17,553 (14.7%)	68,271 (3.1%)
Coronary Artery Disease	16,176 (1.1%)	16,017 (4.2%)	12,015 (4.6%)	2,419 (2.0%)	46,627 (2.1%)
Dementia	7,705 (0.6%)	7,812 (2.0%)	18,748 (7.2%)	19,328 (16.2%)	53,593 (2.5%)
Depression	271,343 (19.2%)	55,438 (14.4%)	37,418 (14.4%)	18,220 (15.3%)	382,419 (17.6%)
Haemorrhagic Stroke	3,938 (0.3%)	1,959 (0.5%)	1,733 (0.7%)	867 (0.7%)	8,497 (0.4%)
Heart Failure	6,219 (0.4%)	8,736 (2.3%)	14,976 (5.8%)	12,583 (10.5%)	42,514 (2.0%)
Hypertension	265,702 (18.8%)	161,622 (41.9%)	136,905 (52.8%)	60,133 (50.3%)	624,362 (28.7%)

Ischaemic Stroke	9,833 (0.7%)	11,097 (2.9%)	15,836 (6.1%)	10,617 (8.9%)	47,383 (2.2%)
Malignancy	48,115 (3.4%)	32,230 (8.4%)	29,796 (11.5%)	14,646 (12.3%)	124,787 (5.7%)
Peripheral Arterial Disease	8,144 (0.6%)	9,541 (2.5%)	11,073 (4.3%)	4,992 (4.2%)	33,750 (1.6%)
Rheumatoid Arthritis	11,149 (0.8%)	6,244 (1.6%)	5,236 (2.0%)	2,169 (1.8%)	24,798 (1.1%)
Transient Ischaemic Attack	8,065 (0.6%)	10,774 (2.8%)	15,795 (6.1%)	10,916 (9.1%)	45,550 (2.1%)
Diabetes	89,567 (6.3%)	49,954 (13.0%)	37,755 (14.6%)	13,514 (11.3%)	190,790 (8.8%)
Chronic Obstructive Pulmonary Disease	28,352 (2.0%)	22,538 (5.9%)	20,399 (7.9%)	7,395 (6.2%)	78,684 (3.6%)
Chronic Kidney Disease	83,821 (5.9%)	42,059 (10.9%)	36,783 (14.2%)	19,404 (16.2%)	182,067 (8.4%)
Asthma	95,438 (6.8%)	31,682 (8.2%)	20,747 (8.0%)	6,365 (5.3%)	154,232 (7.1%)
Osteoporosis	26,939 (1.9%)	21,884 (5.7%)	24,155 (9.3%)	14,096 (11.8%)	87,074 (4.0%)
Morbid obesity risk group	46,465 (3.3%)	9,516 (2.5%)	3,799 (1.5%)	697 (0.6%)	60,477 (2.8%)
BMI category[1]					
Underweight	10,660 (1.2%)	4,749 (1.6%)	6,520 (3.5%)	5,547 (8.8%)	27,476 (1.9%)
Normal	270,394 (29.3%)	88,178 (29.9%)	70,979 (37.9%)	31,659 (50.4%)	461,210 (31.4%)
Overweight	350,099 (38.0%)	119,969 (40.7%)	72,079 (38.4%)	18,858 (30.3%)	561,005 (38.2%)
Obese	290,704	82,017	37,970	6,743	417,434

	(31.5%)	(27.8%)	(20.3%)	(10.7%)	(28.5%)
Missing[2]	*491,719 (34.8%)*	*90,561 (23.5%)*	*71,577 (27.6%)*	*56,674 (47.4%)*	*710,531 (32.6%)*
Vaccinations					
Flu vaccination	248,157 (17.6%)	269,364 (69.9%)	187,976 (72.5%)	71,906 (60.2%)	777,403 (35.7%)
Pneumococcal vaccination	119,926 (8.5%)	231,908 (60.2%)	184,638 (71.3%)	77,218 (64.6%)	613,690 (28.2%)
Smoking status[1]					
Non-smoker	539,051 (40.7%)	138,073 (37.9%)	94,660 (39.7%)	51,381 (52.4%)	823,165 (40.6%)
Ex-smoker	437,970 (33.0%)	157,393 (43.2%)	109,868 (46.1%)	37,801 (38.5%)	743,032 (36.7%)
Active smoker	348,396 (26.3%)	68,858 (18.9%)	33,807 (14.2%)	8,904 (9.1%)	459,965 (22.7%)
Missing[2]	*88,159 (6.2%)*	*21,150 (5.5%)*	*20,790 (8.0%)*	*21,395 (17.9%)*	*151,494 (7.0%)*
Prescriptions					
median	4	18	32	39	8
Upper: lower quartile	[0: 15]	[4: 42]	[12: 60]	[16: 71]	[1: 29]

[1] % as proportion of known values

[2] missing values as % of cohort

Follow-up

Participant data was extracted for the twelve-year period from 2006 to 2017, inclusive. There were 1,107,481 eligible patients in the first year of the cohort (2006), increasing to 1,491,954 at the beginning of 2017, with a total of 1,070,175 new participants joining the cohort between 2007-2017. Patients contributed a mean of 5 years of data, with 647,239 patients (58.4%) who were present in the first cohort year (2006) having the full 12 years of data. Patients present in 2006 comprised 50.9% of the cohort and contributed 67.0% of the total person-years.

Between 2006 and 2017, 137,481 patients died (6.3% of cohort) and 635,400 patients moved out of an RCGP RSC practice (29.2% of the cohort). The full details of entry and exit to the cohort by calendar year according to age groups and frailty category at cohort entry is given in Supplemental Tables 1 and 2. There was an inflow of new participants over the cohort period, across all age groups and frailty categories, which was more notable in younger age groups. The mean follow-up period increased from 4.8 years in people categorised as fit at cohort entry to 7.4 years in people categorised as severely frail (Supplemental Table 3). The age distribution over the cohort period for the closed cohort (participants who were present in 2006 onwards, showing attrition due to death and leaving RCGP RSC practices) and the open cohort (participants present in 2006 plus those moving into an RCGP RSC practice and people turning 50) is given in Figure 2.

Discussion

There is an urgent need for a better understanding of current and future care requirements for people with frailty. This cohort of approximately 2.1 million adults with long-term follow-up data on frailty status using the electronic Frailty Index (eFI) applied to routinely collected primary care data is the largest cohort so far that will be used for longitudinal analysis to explore frailty dynamics and its impact using linked hospital and mortality data. The dataset includes adults aged 50 and over from around 5% of general practices from a single country (England), thus meeting the project aims to provide a whole-system analysis of representative population-level data. The general practices from which the cohort is derived vary in their characteristics, with a range of geographical locations, urban/rural mix, practice sizes and areas of differing deprivation, further demonstrating the representativeness and generalisability of the RCGP RSC dataset [13]. This diversity will reflect a variety of care settings and approaches to managing people with frailty, so that the subsequent simulation models can reflect population and care heterogeneity.

Characteristics of patients at cohort entry show expected patterns in sociodemographic variables and trends in clinical conditions and lifestyle factors, further demonstrating representativeness of the data and suitability for planned analyses. The observed trends in increasing eFI with age group and the greater proportions of moderate and severe frailty categories in older age groups reflect current knowledge, and the observed presence of frailty in 10% of adults aged 50–64 in our cohort highlights the importance of examining frailty and its trajectories earlier in life. The study design of an open cohort allows entry of younger patients throughout its duration, representing a dynamically ageing population in which the overall mean age increases over the cohort period. The cohort therefore includes substantial data on middle-aged adults which is novel, and also crucial for observing potential earlier manifestations of frailty and its progression over time, as it is likely that future interventions to reduce incidence and progression to earlier stages of frailty may be targeted at this age group.

Study strengths and limitations

This dataset is derived from a single national health service (NHS), in which registered patients are managed in accordance with specific guidelines and broadly similar care pathways even though living in different regions of the country and under different primary healthcare practices. Primary care is free at the point of delivery, as is hospital care. However, the cohort does not include information from private healthcare, which is most commonly available via private medical insurance and accessed by around 11% of the UK population, although schemes have limited cover for general practice [24]. Information on adult social care, which is means tested, is also not available as it is organised via a mix of state, private and voluntary providers.

The large size of the dataset enables the stratification of the cohort by key characteristics whilst maintaining sufficiently sized subgroups to provide precise estimates, which will enable the project to meet the research objectives. The dataset includes a wide range of covariates which have been identified in other studies as being associated with either frailty onset or outcomes following frailty occurrence. However, covariates such as social factors (e.g. loneliness, living situation) and contemporaneous information on residential care status is not available. To address this, information on the impact of such factors on outcomes will be sourced from the literature and included within ongoing analysis supporting the development of the simulation model.

Approximately one third of participants have no ethnicity data, which seemed to be related to a variety of patient characteristics. Participants identifying as ‘white’ comprise 92% of the given values, as compared to 86% in the 2011 Census (data for England and Wales) [25]. It is possible that the under-representation of people from ethnic minorities could also be due to recognised issues with lower primary healthcare usage, rather than practices not reflecting their catchment populations, or due to a reporting bias in the primary care data [26]. However, inclusion of ethnicity data from HES in future planned analyses should increase the proportion of the cohort with recorded ethnicity [27].

During the cohort period, there was significant movement of participants both into and out of the cohort, reflecting real-life population flows which are essential for modelling population health needs. Frailty data and patient and service use outcomes are being collected for each year that the patient is registered, thus providing full outcome ascertainment for each year of participation. The 65–74 group had the lowest number of exits, and higher numbers in other age groups could be a consequence of greater mobility in the working age population or moves related to higher levels of support in older age groups, for example following a health or social care crisis [28]. This could lead to an underestimation of incidence and progression of frailty in our cohort, although participants of older age and greater frailty severity also had the longest follow-up periods. Although the location of practices seemed equally distributed across the deprivation quintiles, there was a trend across all age groups for greater representation of patients from the least deprived quintiles, perhaps reflecting movement to wealthier retirement areas and the suburbs.

The 36 eFI deficits and other long-term conditions were defined using Read codes. For future data extractions, migration of clinical term definitions from Read codes to the Systemised Nomenclature of

Medicine – Clinical Terms (SNOMED CT®) will be necessary to reflect national harmonisation of coding tools across the healthcare pathway [29].

Implications for planned analyses

The long follow-up period, averaging 5 years and with more than 600,000 participants over 12 years, is essential for being able to move beyond merely exploring impact of frailty states and associated outcomes at a defined timepoint. The time frame allows for many transitions between frailty states over the cohort period to be observed and gives ample scope for investigating the trajectories of adults with frailty. The use of multistate methods to look at transitions between and time spent in any given state is a key strength of the proposed analyses.

Current evidence suggests that socioeconomic and educational deprivation is associated with higher frailty scores [30]. Analysis of this cohort will enable an in-depth analysis of the patterns of frailty onset and transitions according to social deprivation, which will inform decisions on whether this potentially high-risk population may benefit from targeted earlier intervention to improve outcomes.

The geographical range of the data will enable parameterisation of the simulation model with more localised data, permitting adaptation to reflect local situations and facilitating planning on whichever scale is the most appropriate. If there is interest in looking at the data for particular GP characteristics, for example focussing on larger practices or practices positioned in more deprived areas, the primary analyses can be re-run for the required subgroups adjust parameters for the simulation model.

Conclusions

This cohort is the largest dataset compiling frailty transitions and outcomes and will provide unique information on frailty within middle-aged to young-old populations. Data presented here show that frailty is already common in middle-age and continues to increase across the later life course. The cohort will allow exploration of the impact of morbidities, socioeconomic and lifestyle factors on frailty onset, trajectories and outcomes over time. Strengths of this cohort are the use of large-scale routine primary care data with linkage to secondary care and healthcare costs, and a dynamically ageing population with lengthy follow-up, enabling novel insights into the onset and progression of frailty.

Abbreviations

BMI	Body mass index
BNF	British National Formulary
CTV3	Clinical Terms Version 3
DARS	Digital Access Request Service
eFI	electronic Frailty Index

EHR	Electronic Health Records
FTE	Full-time equivalent
GP	General Practitioner
HES	Hospital Episode Statistics
HSDR	Health Services & Delivery Research
IDAOPI	Income Deprivation Affecting Older People Index
IGARD	Independent Group Advising on the Release of Data
IMD	Indices of Multiple Deprivation
NHS	National Health Service
NIHR	National Institute of Health Research
ONS	Office for National Statistics
PYAR	Person-years at risk
QOF	Quality Outcomes Framework
RCGP RSC	Royal College of General Practitioners Research Surveillance Centre
RUC11	Rural urban classification 11
SD	Standard deviation
SEG	Stakeholder Engagement Group
SNOMED CT®	Systemised Nomenclature of Medicine – Clinical Terms

Declarations

Ethics approval and consent to participate

The study was approved by the University of Southampton Research Ethics Committee (ref 46313) on 6/2/2019 and the RCGP RSC Information Governance Panel on 24/1/2019. Approval for release of secondary care and ONS data through the NHS Digital Data Access Request Service (DARS) was given by the Independent Group Advising on the Release of Data (IGARD) on 19/2/2021 (ref NIC 353126).

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from the RCGP RSC and NHS Digital, but restrictions apply to the availability of these data, which were used following approvals and data sharing agreements for the current study, and so are not publicly available.

Competing interests

The authors have no competing interests to declare.

Funding

This project is funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research (HSDR) programme, award HS&DR 16/116/43. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Author's contributions

BW is the principal investigator for this study. S DS F, PR, S de L, AC, SB, AB, HPP, SH and SZ are study co-applicants, contributing to the overall project objectives and study design. VW is the PPI contributor for the study and has informed study objectives and proposed analyses. TE, DE and FL have contributed to data specifications and analysis plans and co-ordinated data extraction. CF performed the data cleaning and analyses and drafted the paper. All authors have approved the final version.

Acknowledgements

The authors would like to thank patients and practices who are members of the RCGP RSC and who allow data sharing for this and other research, EMIS, TPP, In Practice and Wellbeing for their assistance with the RSC's pseudonymised data extract, and Filipa Ferreira, John Williams and Julian Sherlock from the RCGP RSC team for their assistance with the data specification and data extraction.

AC is part-funded by the National Institute for Health Research Applied Research Collaboration Yorkshire & Humber, and Health Data Research UK, an initiative funded by UK Research and Innovation Councils, National Institute for Health Research and the UK devolved administrations, and leading medical research charities.

The views expressed are those of the authors, and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K: **Frailty in elderly people**. *Lancet* 2013, **381**(9868):752-762.
2. Campbell AJ, Buchner DM: **Unstable disability and the fluctuations of frailty**. *Age Ageing* 1997, **26**(4):315-318.
3. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G *et al*: **Frailty in older adults: evidence for a phenotype**. *J Gerontol A Biol Sci Med Sci* 2001, **56**(3):M146-156.
4. Aguayo GA, Donneau AF, Vaillant MT, Schritz A, Franco OH, Stranges S, Malisoux L, Guillaume M, Witte DR: **Agreement Between 35 Published Frailty Scores in the General Population**. *Am J Epidemiol* 2017, **186**(4):420-434.
5. O'Caoimh R, Galluzzo L, Rodriguez-Laso A, Van der Heyden J, Ranhoff AH, Lamprini-Koula M, Ciutan M, Lopez-Samaniego L, Carcaillon-Bentata L, Kennelly S *et al*: **Prevalence of frailty at population level in European ADVANTAGE Joint Action Member States: a systematic review and meta-analysis**. *Ann Ist Super Sanita* 2018, **54**(3):226-238.
6. Buchman AS, Wilson RS, Bienias JL, Bennett DA: **Change in frailty and risk of death in older persons**. *Exp Aging Res* 2009, **35**(1):61-82.
7. Rohrmann S: **Epidemiology of Frailty in Older People**. In., 2020.
8. Hollinghurst J, Fry R, Akbari A, Clegg A, Lyons RA, Watkins A, Rodgers SE: **External validation of the electronic Frailty Index using the population of Wales within the Secure Anonymised Information Linkage Databank**. *Age Ageing* 2019.
9. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, Mohammed MA, Parry J, Marshall T: **Development and validation of an electronic frailty index using routine primary care electronic health record data**. *Age Ageing* 2016, **45**(3):353-360.
10. NHS England: **NHS standard general medical services contract 2017/18**. In., 2018.
11. NHS England: **The NHS Long Term Plan**. In., 2019.
12. Walsh B: **Study protocol: The dynamics of frailty in older people: modelling impact on health care demand and outcomes to inform service planning and commissioning.** In., 2018.
13. Correa A, Hinton W, McGovern A, van Vlymen J, Yonova I, Jones S, de Lusignan S: **Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile**. *BMJ Open* 2016, **6**(4):e011092.
14. HM Government: **File 3: supplementary indices - income deprivation affecting children index and income deprivation affected older people index**. In., 2015.
15. Office for National Statistics: **Output Area to LSOA to MSOA to Local Authority District (December 2017) Lookup with Area Classifications in Great Britain** In., 2017.
16. NHS Digital: **General Practice Workforce**. In.
17. Department for Communities and Local Government: **The English Index of Multiple Deprivation (IMD) 2015 – Guidance** In. Edited by Department for Communities and Local Government, 2015.

18. Tippu Z, Correa A, Liyanage H, Van Vlymen J, Burleigh D, McGovern A, Jones S, de Lusignan S: **Ethnicity recording in primary care computerised medical record systems: an ontological approach**. *BMJ Health &amp,amp, Care Informatics* 2016, **23**(4):799.
19. Office for National Statistics: **2011 rural/urban classification (RUC11)**. In., 2011.
20. NHS Digital: **General and Personal Medical Services, practice level dataset** In. Edited by Digital, N, 2013.
21. Joint Formulary Committee: **British National Formulary**, 81 edn, 2020.
22. Maguire A, Blak BT, Thompson M: **The importance of defining periods of complete mortality reporting for research using automated data from primary care**. *Pharmacoepidemiol Drug Saf* 2009, **18**(1):76-83.
23. Joy M, Hobbs FDR, McGagh D, Akinyemi O, de Lusignan S: **Excess mortality from COVID-19 in an English sentinel network population**. *Lancet Infect Dis* 2020.
24. The King's Fund: **The UK private health market**. 2014.
25. **Ethnicity facts and figures** [https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest]
26. Lakhani M: **No patient left behind: how can we ensure world class primary care for black and minority ethnic people?** In. London, 2008.
27. Mathur R, Bhaskaran K, Chaturvedi N, Leon DA, vanStaa T, Grundy E, Smeeth L: **Completeness and usability of ethnicity data in UK-based primary care and hospital databases**. *J Public Health (Oxf)* 2014, **36**(4):684-692.
28. Scheibl F, Farquhar M, Buck J, Barclay S, Brayne C, Fleming J, Collaboration obotCCo-sCS: **When Frail Older People Relocate in Very Old Age, Who Makes the Decision?** *Innovation in Aging* 2019, **3**(4).
29. HM Government: **Personalised Health and Care 2020. Using Data and Technology to Transform Outcomes for Patients and Citizens. A Framework for Action.** . In. Edited by Board NI, 2014.
30. Franse CB, van Grieken A, Qin L, Melis RJF, Rietjens JAC, Raat H: **Socioeconomic inequalities in frailty and frailty components among community-dwelling older citizens**. *PLoS One* 2017, **12**(11):e0187946.

Figures

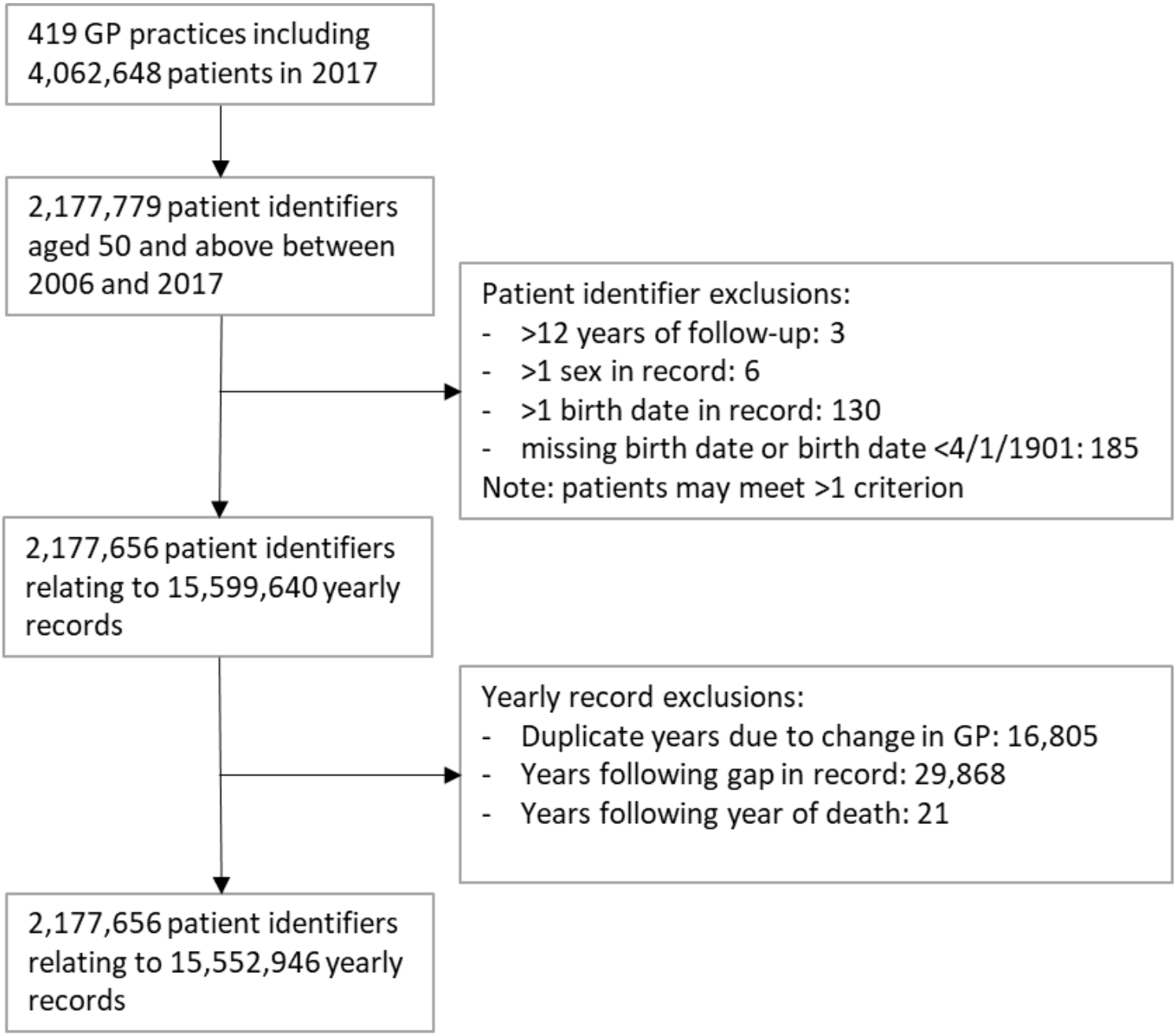

Figure 1

Cohort definition

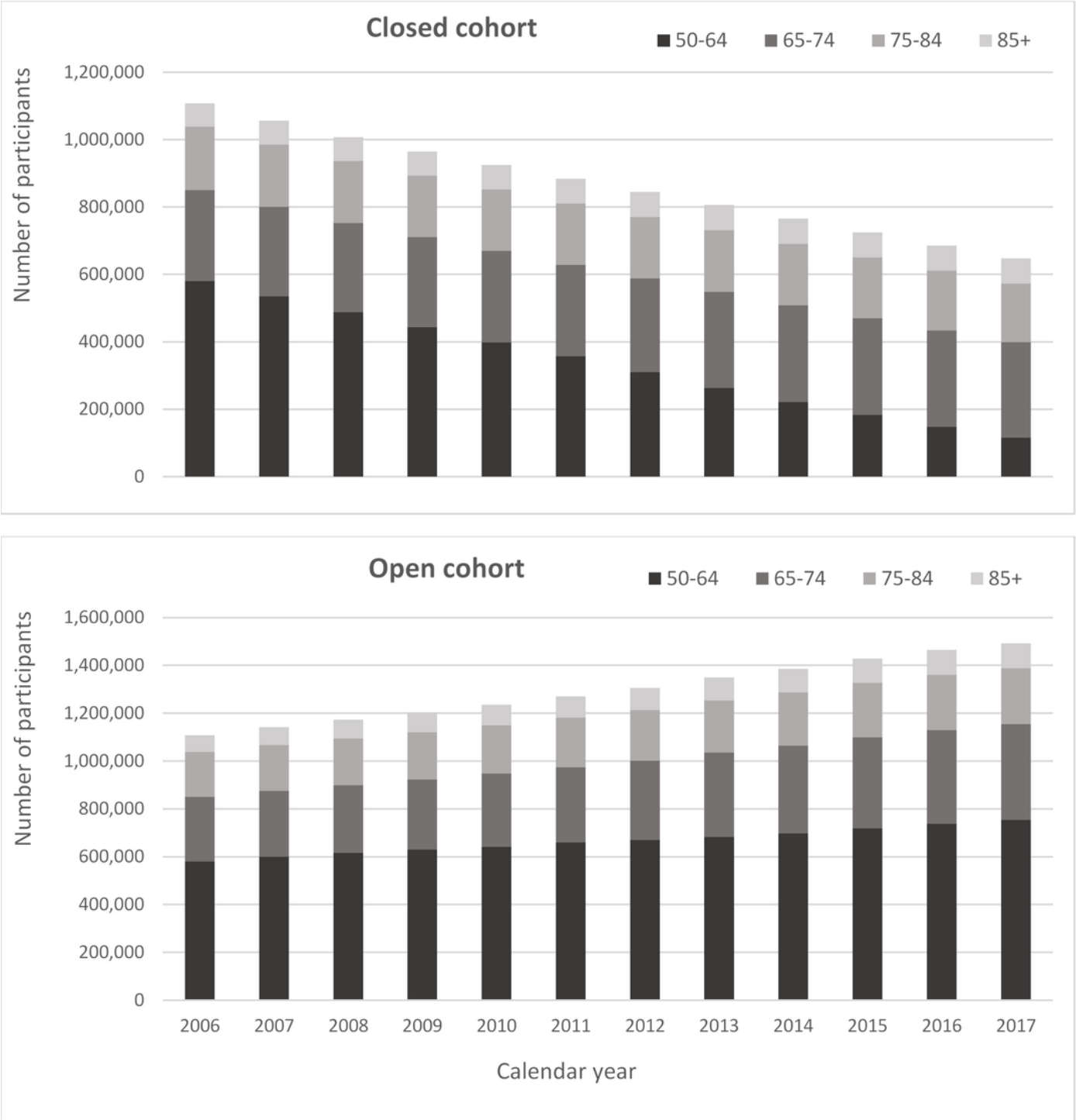

Figure 2

Age group distribution over cohort period

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaltables.docx