BMJ Open

The journey to multimorbidity: a longitudinal study in an urban setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031649
Article Type:	Research
Date Submitted by the Author:	14-May-2019
Complete List of Authors:	Ashworth, Mark; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Durbaba, Stevo; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Whitney, David; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Crompton, James; London Borough of Lambeth, Public Health, Adults and Health Wright, Michael; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Dodhia, Hiten; London Borough of Lambeth, Public Health
Keywords:	PRIMARY CARE, Multimorbidity, Demographic factors, Risk factors

TITLE:

The journey to multimorbidity: a longitudinal study in an urban setting

AUTHORS:

Mark Ashworth[1], Stevo Durbaba[2], David Whitney[3], James Crompton[4], Michael Wright[5] Hiten Dodhia[6].

1. Reader in Primary Care, School of Population Health & Environmental Sciences, King's College London, London, UK
2. e-Resources Developer, School of Population Health & Environmental Sciences, King's College London, London, UK
3. Data Analyst, School of Population Health & Environmental Sciences, King's College London, London, UK
4. Head of Health Intelligence, Public Health, Adults and Health, London Borough of Lambeth, London, UK
5. Project Manager, School of Population Health & Environmental Sciences, King's College London, London, UK
6. Consultant in Public Health, London Borough of Lambeth, London, UK

Corresponding author email: mark.ashworth@kcl.ac.uk

Word count: 3187

ABSTRACT

Objective. To study the social determinants and risk factors for multimorbidity and the acquisition sequence of multimorbidity.

Design. Cross-sectional analysis and longitudinal study based on anonymised primary care data.

Setting. General practices in an urban multi-ethnic borough in London, UK.

Participants. 332,353 patients aged ≥18 years

Main outcome measures. Clinical and demographic characteristics of patients with multimorbidity, defined as ≥3 of 12 Long Term Conditions (LTCs) selected according to high predicted healthcare utilisation. Multilevel logistic regression was used to model the social determinants and risk factors for multimorbidity. Alluvial plots were constructed to illustrate differing multimorbidity acquisition sequences according to age, ethnicity and social deprivation.

Results. The commonest LTCs were diabetes (63.0%) and chronic pain (42.8%). Social deprivation and ethnicity were independent determinants of multimorbidity: for most compared to least

deprived quintile, Adjusted Odds Ratio (AOR) 1.56 (95%CIs, 1.41, 1.72); for South Asian compared to white ethnicity, AOR 1.44 (95%CIs, 1.29, 1.61); for black compared to white ethnicity, AOR 0.86 (95%CIs, 0.80, 0.92). The three included risk factors were relatively strong determinants of multimorbidity: hypertension, AOR 5.05 (95%CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95%CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45). The most common initial onset conditions were diabetes and depression; diabetes particularly in older and black ethnic groups; depression particularly in younger, more deprived and white ethnicity groups. Chronic pain was less common as an initial condition.

Conclusions Our findings confirm the importance of age, social deprivation and ethnicity as determinants of multimorbidity. However, the risk factors of smoking, obesity and hypertension were stronger determinants of multimorbidity than deprivation or ethnicity. The acquisition sequence of multimorbidity is patterned by demographic determinants. Understanding the onset conditions of multimorbidity and risk factors may lead to the development of interventions to slow the progression of multimorbidity.

Strengths and limitations

- This study defines multimorbidity based on the inclusion of Long Term Conditions (LTCs) with high predicted healthcare utilisation rates
- Multimorbidity is studied in a deprived, multi-ethnic community
- Few studies of multimorbidity have analysed longitudinal data; this study uses longitudinal data to identify the acquisition sequence of multimorbidity and how this is influenced by demographic determinants
- Difficulties gaining access to anonymised primary care data limited the sample size
- Coding validation and completeness restricted the analysis of available primary care data

BACKGROUND

Healthcare utilisation is increasingly driven by multimorbidity (1). Each long-term condition (LTC) included within a definition of multimorbidity is likely to generate multiple primary care consultations with general practitioners (GPs), practice nurses and other health care professionals, and may result in Accident and Emergency (A&E) attendances, referral to out-patient appointments and hospital admissions.

Estimates of healthcare utilisation attributable to multimorbidity vary according to the LTCs included within a definition of multimorbidity. There is no standard definition of multimorbidity. One systematic review noted that the number of included LTCs ranged from five to 185 with estimates of population prevalence ranging from 13.1% to 71.8% depending on the number of included conditions (2). In one recent UK study in which multimorbidity was defined as two or more of a selection of 36 LTCs, 27.2% of the population had multimorbidity, accounting for 52.9% of GP consultations and generating a median of nine annual GP consultations (3). LTCs also contribute to the development of frailty, itself a driver of healthcare utilisation (4).

In response to high demand for health and social care, many healthcare providers and commissioners have sought to identify those patients with greatest needs through a process termed 'risk stratification'. Several electronic tools have been developed to offer population based risk stratification (5). An alternative approach is the use of expert panels to define high-demand patient groups (6). Having identified the cohort of patients with the greatest requirement for health and social care services, the purpose of this process is to guide resource allocation on a needs basis, often with the implicit assumption that additional investment in a primary care setting may reduce demand for more expensive secondary care services. Although funding and healthcare need should align, there is little evidence that investment in additional community resources for those most at risk of hospital attendance results in overall reductions in secondary care utilisation (7).

Following a consultation exercise and report from an expert panel, two inner London boroughs and their commissioning groups made the decision to define multimorbidity based on predicted high healthcare and social care demand. Multimorbidity in this context consisted of three or more LTCs considered most likely to result in functional impairment and high service demand. This 'high service demand' definition was used to identify a cohort to receive a package of integrated care, termed 'care coordination'.

We aimed to study the characteristics of this multimorbidity cohort, defining both the demographic determinants and risk factors associated with multimorbidity acquisition. Then to determine the acquisition sequence of multimorbidity and the influence of demographic factors on this sequence.

METHODS

Study setting

Our study was set in Lambeth, one of the two inner London boroughs adopting the 'care coordination' definition of multimorbidity (8). The population sample consisted of all patients registered at all general practices in Lambeth, with the exception of patients who had opted out of anonymised data sharing for research purposes.

Study design

We conducted a cross-sectional analysis and longitudinal study based on anonymised coded primary care data extracted from electronic health records (EHR) held in primary care.

Study population

We included data on all patients aged 18 years and over registered with a general practice. For the population with multimorbidity, we included all those with LTCs recorded in the EHR and included in the 'care coordination' definition of multimorbidity: Atrial Fibrillation (AF), Chronic Obstructive

Pulmonary Disease (COPD), Chronic Pain (CP), Chronic Kidney Disease (CKD), Coronary Heart Disease (CHD), Diabetes (DM), Dementia, Depression, Heart Failure (HF), Serious Mental Illness, Stroke, Morbid Obesity. The definition and specified codes for each condition was that used by the Quality and Outcomes Framework (QOF), based on 'QOF38' definitions (9). Two of the selected conditions were not included within the QOF: chronic pain, defined on the basis of two or more repeat prescriptions for opioid analgesics (British National Formulary, chapter 4.7.2) or neuropathic pain medication (British National Formulary, chapter 4.7.3) (10); morbid obesity defined as a Body Mass Index ≥40 kg/m^2.

Demographic data consisted of gender, age in years and self-ascribed ethnicity obtained from the EHR. Social deprivation data derived from residency data was based on the Index of Multiple Deprivation 2010 classification at Lower Super Output Area (LSOA), stratified into locally based quintiles (11). Local deprivation quintiles were used in place of national quintiles since mean deprivation levels are high in Lambeth, the 22nd most deprived local authority (out of 326) in England (12).

Risk factors included in the analysis were: hypertension (defined as patients on the QOF Hypertension register); moderate obesity (defined as a Body Mass Index of 30.0-39.9; note that patients with BMI ≥40 were included as one of the LTCs within the definition of multimorbidity and therefore not included as a risk factor); smoking (patients with any record of being a smoker).

Data variables

The data consisted of 'real world', routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Data were extracted from the EHR into a secure data warehouse on a quarterly basis and contained information on patient demographic characteristics, LTCs, clinical values and medication. The data used in this study were extracted in May 2018.

Data analysis

We analysed demographic, social and risk factor data for the multimorbidity cohort and general population using univariable statistical methods applied at patient-level. Demographic and risk factor determinants of multimorbidity were analysed using mixed effects multilevel logistic regression models, with data analysed at patient level and general practice level, based on the registered general practice of each patient in the study sample. We also conducted a sensitivity analysis using mixed effects logistic regression models to allow for random effects, adjusted for clustering at the practice level. The sensitivity analysis allowed pseudo-r^2 values and Receiver Operating Characteristic (ROC) curves to be derived. Analysis was conducted using the statistical software package STATA IC 15.

The acquisition sequence for patients in the multimorbidity cohort was established by searching the EHR for date of onset of each LTC. For this analysis, we established the order of acquisition and tabulated the frequency of first, second and third LTCs. We displayed findings using alluvial plots, an infographic allowing representation of multiple pathways. These were constructed using the software R, and the packages 'ggplot2' and 'ggalluvial' (14).

Patient and public involvement

The borough based and statutory organisation, Lambeth HealthWatch, represented the interests of patients and public in this work; they contributed to the original protocol design and shared in dissemination of the findings.

RESULTS

Multimorbidity cohort characteristics

The final study population consisted of 332,353 patients aged ≥18 years. Data from 13,369 (4.0%) patients had been excluded because a data sharing opt-out code was recorded in their EHR. Patients were included in the final sample even though some demographic data were missing: 3289 (0.99%)

patients could not be linked to a LSOA and therefore had missing IMD-2015 score data; ≤10 patients had missing coded gender data.

In all, 5597 (1.7%) patients had a record of three or more of the selected LTCs, the 'multimorbidity cohort'. Most (n = 3542) of this cohort had three LTCs (63.3%); 1333 (23.8%) had four LTCs; 492 (8.8%) had 5 LTCs; the remaining 230 (4.1%) had more than 5 LTCs.

A summary of LTC frequencies within the multimorbidity cohort is displayed in Table 1. The most common LTCs within this cohort were: DM (63.0%) and chronic pain (42.8%). In contrast, the most common of the included LTCs in the adult general population were: depression (8.4%), DM (5.4%) and morbid obesity (3.2%).

Table 1. Frequencies of Long Term Conditions included in the multimorbidity cohort (n = 5597 patients) compared with the remainder of the registered population and with the general population aged ≥18 years.

Long Term Condition	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756	Population values: frequency (valid %) N = 332,353
DM	3525 (63.0%)	14,405 (4.4%)	17,930 (5.4%)
Chronic Pain	2397 (42.8%)	5813 (1.8%)	8210 (2.5%)
CKD	2101 (37.5%)	3470 (1.1%)	5571 (1.7%)
CHD	2099 (37.5%)	2668 (0.8%)	4767 (1.4%)
Depression	2086 (37.3%)	25,877 (7.9%)	27,963 (8.4%)
Morbid Obesity	1653 (29.5%)	8883 (2.7%)	10,486 (3.2%)
AF	1254 (22.4%)	1493 (0.5%)	2747 (0.8%)
COPD	1247 (22.3%)	2387 (0.7%)	3634 (1.1%)
Heart Failure	1186 (21.2%)	544 (0.2%)	1730 (0.5%)
Stroke	1095 (19.6%)	1571 (0.5%)	2666 (0.8%)
SMI	692 (12.4%)	4099 (1.3%)	4791 (1.4%)
Dementia	616 (11.0%)	738 (0.2%)	1354 (0.4%)

The demographic characteristics of the multimorbidity cohort are displayed in Table 2. Within the cohort, 33.9% were aged under 65 years (compared with 91.7% in the sample population); 27.7% were of a 'Black' ethnicity (18.0% in the sample population) and 46.0% were born in the UK (45.2% in the sample population). The mean age for multimorbid patients in least and most deprived quintiles was 73.0 and 69.3 years, respectively; and for the white, black and south Asian populations was 71.2, 69.4 and 71.9 years, respectively.

Table 2. Demographic characteristics of multimorbidity cohort, compared with remainder of registered population aged ≥18 years.

Demographic characteristic	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756
Female gender	3042 (54.4%)	161,445 (49.4%)
Age <65 years	1899 (33.9%)	299,742 (91.7%)
Age ≥65-74 years	1249 (22.3%)	15,992 (4.9%)
Age ≥75-84 years	1479 (26.4%)	8038 (2.5%)
Age ≥85 years	970 (17.3%)	2884 (0.9%)
White	3022 (54.0%)	179,859 (55.0%)
Black	1553 (27.7%)	58,939 (18.0%)
South Asian	469 (8.4%)	22,323 (6.8%)
Mixed	197 (3.5%)	15,177 (4.6%)
Other	100 (1.8%)	9804 (3.0%)
Unknown	256 (4.6%)	40,654 (12.4%)
Country of origin: UK*	1413 (46.0%)	69,675 (45.2%)
Language preference: English*	3755 (84.0%)	240,287 (73.8%)
Social deprivation: 1st quintile (most deprived)*	1500 (26.8%)	63,374 (19.4%)
Social deprivation: 2nd quintile*	1232 (22.0%)	63,073 (19.3%)
Social deprivation: 3rd quintile*	995 (17.8%)	66,409 (20.3%)

Social deprivation: 4th quintile*	1013 (18.1%)	66,334 (20.3%)
Social deprivation: 5th quintile (least deprived)*	828 (14.8%)	64,306 (19.7%)

*missing data with reduction in denominator number

Multimorbidity cohort demographic determinants

Demographic determinants of the multimorbidity cluster are displayed in Table 3. Based on the adjusted odds ratio (AOR) derived from the multilevel regression model, the strongest determinant for multimorbidity was related to age. After adjustment for age, both social deprivation (the more deprived quintiles) and ethnicity (Black and south Asian ethnicities) remained significant determinants, albeit with much smaller odds ratios. The AOR for black and South Asian compared with white ethnicity was 1.15 (95% CIs, 1.07, 1.23) and 1.19 (95% CIs, 1.07, 1.33), respectively; for most compared to least deprived quintile, the AOR was 1.83 (95% CIs, 1.66, 2.02).

Table 3. Demographic determinants of the multimorbidity cohort: adjusted odds ratios derived from mixed effects multi-level logistic regression modelling.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.99, 1.11
Age <65 years	(reference group)	
Age ≥65-74 years	11.81	10.96, 12.72
Age ≥75-84 years	27.14	25.22, 29.20
Age ≥85 years	48.70	44.60, 53.18
White	(reference group)	
Black	1.15	1.07, 1.23
South Asian	1.19	1.07, 1.33
Mixed	0.96	0.83, 1.12
Other	0.76	0.62, 0.93
Unknown	0.44	0.38, 0.50
Social deprivation: 1st quintile (most deprived)	1.83	1.66, 2.02

Social deprivation: 2nd quintile	1.58	1.43, 1.75
Social deprivation: 3rd quintile	1.27	1.15, 1.40
Social deprivation: 4th quintile	1.21	1.10, 1.33
Social deprivation: 5th quintile (least deprived)	(reference group)	

Multimorbidity cohort risk factor determinants

Addition of the three risk factors included in the study attenuated the odds ratios for multimorbidity related to age (Table 4). Social deprivation remained a determinant of multimorbidity: for most compared to least deprived quintile, the AOR was 1.56 (95% CIs, 1.41, 1.72). South Asian ethnicity remained a significant determinant of multimorbidity: AOR 1.44 (95% CIs, 1.29, 1.61) but black ethnicity was no longer a positive determinant: AOR 0.86 (95% CIs, 0.80, 0.92).

The three risk factors were significant determinants of multimorbidity: hypertension, AOR 5.05 (95% CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95% CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45).

Table 4. Demographic determinants of the multimorbidity cohort: odds ratios derived from mixed effects multi-level logistic regression modelling with addition of three risk factors: Hypertension, Obesity (moderate), Smoking (ever).

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.10	1.01, 1.19
Age <65 years	(reference group)	
Age ≥65-74 years	4.01	3.69, 4.36
Age ≥75-84 years	8.12	7.46, 8.83
Age ≥85 years	15.71	14.20, 17.38
White	(reference group)	
Black	0.86	0.80, 0.92
South Asian	1.44	1.29, 1.61

Mixed	0.95	0.81, 1.11
Other	0.83	0.67, 1.03
Unknown	0.60	0.52, 0.69
Social deprivation: 1st quintile (most deprived)	1.56	1.41, 1.72
Social deprivation: 2nd quintile	1.35	1.22, 1.50
Social deprivation: 3rd quintile	1.18	1.06, 1.31
Social deprivation: 4th quintile	1.18	1.06, 1.30
Social deprivation: 5th quintile (least deprived)	(reference group)	
Hypertension register	5.05	4.69, 5.44
Moderate obesity	3.41	3.21, 3.63
Smoker (ever)	2.30	2.16, 2.45

Sensitivity analyses

Re-analysis of the determinants of multimorbidity using regression modelling adjusted for clustering at practice level resulted in similar adjusted odds ratios to those obtained in the primary analyses (Table 5). We explored goodness-of-fit through derived pseudo-r^2 values for demographic determinants: pseudo-r^2 = 0.22, and for risk factor adjusted determinants (hypertension, moderate obesity, smoking): pseudo-r^2 = 0.32. The areas under the ROC curve for each model were 0.84 and 0.93, respectively (Supplementary file: Figures 1a, 1b).

Table 5. Demographic determinants of the multimorbidity cohort: odds ratios derived from mixed effects logistic regression modelling adjusted for clustering at practice level.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.97, 1.13
Age <65 years	(reference group)	
Age ≥65-74 years	12.08	10.53, 13.86
Age ≥75-84 years	27.74	24.28, 31.71

Age ≥85 years	49.84	43.00, 57.78
White	(reference group)	
Black	1.19	1.06, 1.33
South Asian	1.16	1.00, 1.34
Mixed	0.98	0.81, 1.20
Other	0.77	0.62, 0.96
Unknown	0.44	0.36, 0.55
Social deprivation: 1st quintile (most deprived)	1.96	1.69, 2.26
Social deprivation: 2nd quintile	1.66	1.43, 1.92
Social deprivation: 3rd quintile	1.31	1.13, 1.51
Social deprivation: 4th quintile	1.24	1.09, 1.42
Social deprivation: 5th quintile (least deprived)	(reference group)	

Multimorbidity acquisition sequence

Of the 5597 patients in the multimorbidity cohort, 5196 had three distinct dates of onset for each of their three or more component LTCs. The remaining 401 (7.2%) patients had identical dates of onset recorded for two or three of their first three LTCs and therefore could not be classified into a sequence. Alluvial plots were constructed displaying the acquisition sequence for all patients (Figure 1a) and edited to display dominant flows of patients in each category. Figure 1b displays the three most common starting conditions and subsequent most commonly acquired second and third LTCs.

The alluvial plots illustrate that diabetes and depression were the most common starting conditions for patients with multimorbidity; diabetes was also relatively common as the second or third acquired LTC whereas depression was predominantly a first-onset LTC (Figure 1b).

In the most deprived quintile, diabetes and depression were the most common starting conditions whereas in the least deprived quintile, diabetes and CHD were more common as starting conditions than depression (Figures 2b and 3b).

Multimorbidity in the white ethnic group was dominated by depression as the starting condition, whereas in the black ethnic group diabetes was the most common starting condition, with depression and SMI also relatively common (Figures 4b and 5b).

Multimorbidity in the under 65 year old cohort was dominated by depression as the starting condition, with SMI also relatively common; in the ≥65 year old cohort, diabetes and CHD were the most common starting conditions.

Chronic pain appeared to be more common as a second or third acquired LTC but less common as a first LTC. This sequence was apparent in the overall picture, in the pattern displayed by least and most deprived quintiles, for black and white ethnicities and for younger and older age cohorts.

Morbid obesity was among the more common starting conditions in the most deprived cohort and younger age cohort. However, in other demographic samples, morbid obesity was more common as a second and third acquired LTC.

DISCUSSION

We report on multimorbidity using a definition originating from a health service commissioning perspective, consisting of three or more of 12 LTCs selected because of likely high impact upon health service and social care utilisation. In total, 1.7% of the adult population in our study sample had multimorbidity according to these narrowly defined criteria. Diabetes and chronic pain were the most prevalent LTCs within this cohort. Independent of age, both ethnicity and social deprivation were significant determinants of multimorbidity. However, the risk factors of hypertension, obesity and smoking were more strongly associated with multimorbidity than social deprivation or ethnicity.

The acquisition sequence of multimorbidity differed substantially according to age, ethnicity and social deprivation. Diabetes and depression were the most common starting conditions overall. Diabetes as a starting condition was notably more common in the older and black ethnic group. Depression as a starting condition was notably more common in patients who were younger, more deprived and in the white ethnic group. Differences in acquisition sequence between most and least deprived areas and white and black ethnicities (Figures 3-5) are unlikely to have been strongly influenced by age differences since mean age was similar for each of these cohorts.

Strengths and Limitations

Data access was a limitation to this analysis. We were only able to obtain data from one of the two boroughs adopting this approach to multimorbidity, the other lacked a data extraction system preventing us from analysing large datasets of patient-level data. Had we gained access to the data, this would have approximately doubled our sample size and enabled further analysis of multimorbidity in deprived, multi-ethnic populations. This difficulty in accessing anonymised data hampers the analysis of patient-level data in many areas of the UK (15). Data coding constrains the analysis of primary care data and we were only able to study the association of multimorbidity with a limited range of risk factors while other known risk factors such as exercise and diet could not be captured. We aimed to display the acquisition sequence of LTCs using alluvial plots. However, these plots do not readily display time data resulting in a lack of clarity in the rate of progression of multimorbidity. A time-to-event analysis is required for identifying those patients who progress rapidly from first LTC into multimorbidity, which is the subject of further study. Similarly, the acquisition sequence could not be determined for a small minority of patients with identical LTC date-of-onset recording by GPs whereas in reality, it is unlikely that the LTC onset dates were simultaneous. Furthermore, as with all studies based on primary care data, there may be coding anomalies which introduce bias into the estimates of LTC prevalence. QOF coding criteria were used for 10 of the included LTCs, standardising the definition. However, the prevalence of conditions such

as depression may be underestimated using QOF criteria (16). For the two LTCs not included in the QOF, the definition is dependent on GP coding. Thus 'morbid obesity' was only included in our study if there was a BMI recording which may have resulted in an under-estimate of prevalence. 'Chronic pain' was defined based on medication consumption whereas many patients with chronic pain may have sought alternatives to analgesic medication resulting in an underestimate of prevalence; conversely, our inclusion criteria of 'two or more prescriptions over the preceding year' may have resulted in an overestimate of prevalence, with some patients recovering from chronic pain during the course of the year. Finally, the richness of locally based data covering a whole borough with unique demographic characteristics has to be offset against possible loss of generalisability to other areas with very different social deprivation and ethnicity characteristics.

Comparison with the literature

Our cross-sectional data, although conducted in a deprived, multi-ethnic population, are similar to the findings of others reporting on increased multimorbidity prevalence associated with age, social deprivation and ethnic minority status (17). Comparison with other multimorbidity studies is difficult because of the highly restricted definition of multimorbidity used in the current study. Nevertheless, other studies have reported the high prevalence of both diabetes and depression in multimorbidity cohorts (3, 17). A higher prevalence of mental health and physical health LTC combinations has been noted in deprived areas, although in our own study we did not conduct an analysis to identify specific LTC combinations (18). Certain conditions have been found to be more prevalent in deprived communities, also contributing to higher prevalence of multimorbidity in these areas, such as depression and addiction issues in younger deprived populations (19), or more generally, depression, drugs, anxiety, dyspepsia, chronic pain, CHD, DM (20). These findings are aligned to our own reporting of high proportions of multimorbid patients in deprived areas with depression and diabetes. However, the inner city population in our study is characterised by a much younger overall age profile compared with the national population: 8% of the population of Lambeth is aged 65

years or more compared to a mean of 18% in England (21). Added to this, the most socially deprived and black ethnic minority groups in our sample were somewhat younger. Both factors are likely to have reduced the overall study prevalence of conditions associated with ageing such as diabetes and CHD. Many multimorbidity studies do not include morbid obesity within their definition (3, 17); we found morbid obesity was particularly associated with social deprivation.

Several publications have reported on combinations and clusters of LTCs but few longitudinal analyses have been reported (22). One study from Australia reported the order of appearance for eight LTCs, reporting in detail for asthma and mood related disorders, the two LTCs most strongly associated with the risk of developing a second LTC. For those with baseline asthma, there was a higher subsequent risk of developing COPD and hypercholesterolaemia; for those with baseline mood disorders, there risk of subsequent asthma, diabetes and other mental disorders was increased (23). In a study of cardiometabolic conditions in Australia, nearly one-quarter of women initially diagnosed with stroke subsequently progressed to other conditions which was a much larger proportion progressing to other conditions than in those initially diagnosed with diabetes (9.9%) or heart disease (11.4%) (24). A further US study explored the acquisition sequence of 20 LTCs describing dyads and triads of conditions and reporting, for example, that the most common triad sequence in 20-39 year olds was depression, asthma and substance misuse whereas in 50-59 year olds it was hyperlipidaemia, hypertension and diabetes (25). They concluded that combinations of LTCs vary extensively by age and sex. Our own findings confirm variation by age and sex, with ethnicity adding to the pattern of variation. Some authors have suggested that the study of acquisition sequence may suggest potential interventions to prevent, minimise or delay progression toward multimorbidity (23). Our findings that three risk factors are more strongly associated with multimorbidity than deprivation and ethnicity, suggest that interventions to reduce the impact of these risk factors may contribute to a reduction in the prevalence of multimorbidity.

CONCLUSION

We have confirmed the role of age, social deprivation and ethnicity as determinants of multimorbidity in an inner city multi-ethnic population and have extended previous findings demonstrating the way in which the acquisition sequence of multimorbidity is patterned by these determinants. Three risk factors, hypertension, obesity and smoking were stronger determinants of multimorbidity than either deprivation or ethnicity. The strength of these risk factors as determinants suggests interventions which may be effective in reducing the prevalence, delaying the onset or slowing the progression of multimorbidity.

What is already known on this topic

- Multimorbidity is known to be associated with social deprivation and ethnicity
- Exposure to cardiovascular risk factors is known to increase the prevalence of multimorbidity

What this study adds

- Ethnicity has differing associations with multimorbidity. After adjustment for risk factors, South Asian ethnicity was positively associated with the prevalence of multimorbidity whereas black ethnicity was negatively associated.
- The most common initial onset multimorbidity conditions are diabetes and depression
- The acquisition sequence of multimorbidity is patterned by age, ethnicity, social deprivation and gender
- Three risk factors, hypertension, obesity and smoking, are stronger determinants of multimorbidity than deprivation and ethnicity.
- Understanding the risk factors and acquisition sequence of multimorbidity may lead to the development of interventions to reduce the prevalence, delay the onset or slow the progression of multimorbidity.

Contributors: MA, HD and MW contributed to the idea and design of the study. SD, DW and JC led on data extraction and preparation. Statistical analysis was conducted by MA and HD. MA produced the first draft of the paper; all co-authors contributed and approved the final draft. MA is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This work was supported by the Guy's and St Thomas' Charity with a grant awarded November 2017. The views expressed are those of the authors and not necessarily those of the

funders. The findings presented are independent from the funders who have had no role in data collection, data analysis or writing of the paper.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval

All data were extracted under the terms of a signed Data Sharing Agreement (DSA) with each practice and with project-specific approval following submission of a Data Privacy Impact Assessment (DPIA), approved by Lambeth Clinical Commissioning Group (CCG) on 2/11/17. Information Governance approval required 'low number suppression', ensuring that data could not be displayed if the patient number was 10 or less in any given category; in these circumstances, data reporting would state: '≤10 patients'. Separate Ethical Committee approval was not required (Health Research Authority, personal correspondence, 29/9/17) since all data were fully anonymised for the purposes of research access, and all Patient Identifiable Data (PID) had been removed.

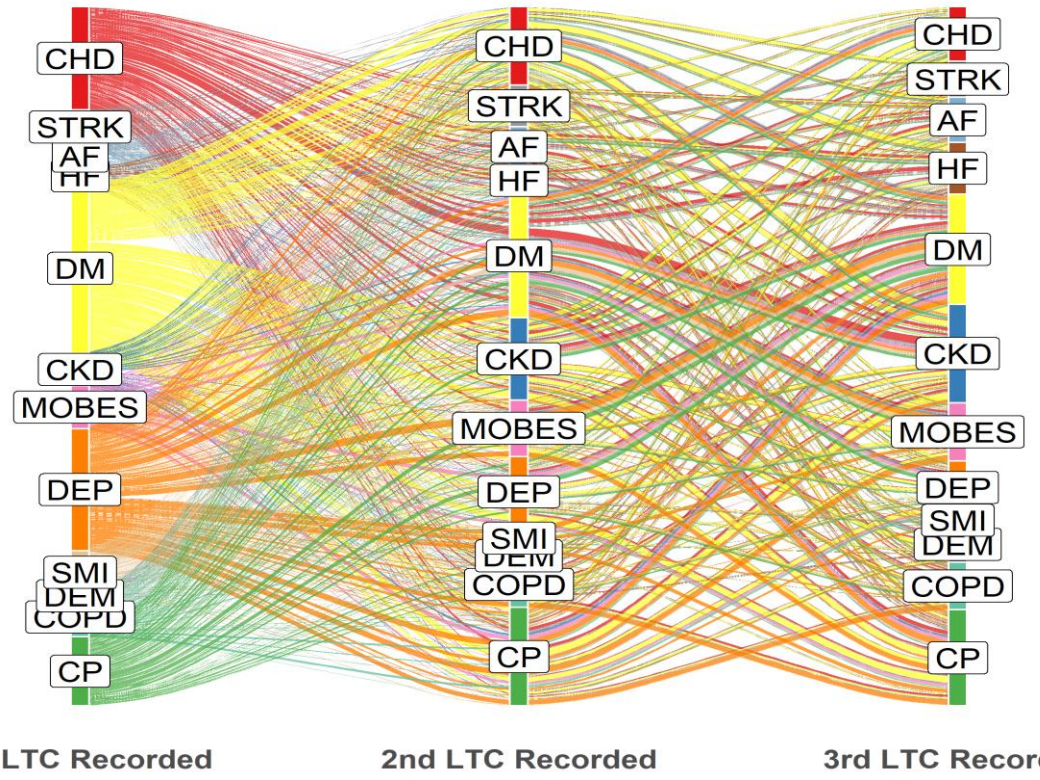
Data sharing: No additional data available.

REFERENCES:

1. Palladino R, Tayu Lee J, Ashworth M, *et al*. Associations between multimorbidity, healthcare utilisation and health status: evidence from 16 European countries. *Age Ageing* 2016;45:431-5. https://dx.doi.org/10.11606%2FS1518-8787.2018052000639
2. Fortin M, Stewart M, Poitras ME, *et al*. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142-51. https://doi.org/10.1370/afm.1337
3. Cassell A, Edwards D, Harshfield A, *et al*. The epidemiology of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2018;68:e245-e51. https://doi.org/10.3399/bjgp18X695465
4. Hanlon P, Nicholl BI, Jani BD, *et al*. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. *Lancet Public Health* 2018;3:e323-e32. https://doi.org/10.1016/S2468-2667(18)30091-4
5. Billings J, Blunt I, Steventon A, *et al*.. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2. http://dx.doi.org/10.1136/bmjopen-2012-001667
6. Huntley AL, Johnson R, Purdy S, *et al*. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med*. 2012;10:134-41. https://doi.org/10.1370/afm.1363
7. Roland M, Abel G. Reducing emergency admissions: are we on the right track? *Br Med J* 2012;345:e6017. https://doi.org/10.1136/bmj.e6017
8. From one to many. Exploring people's progression to multiple long-term conditions in an urban environment. Guy's and St Thomas' Charity. London, 2018. https://www.gsttcharity.org.uk/sites/default/files/GSTTC_MLTC_Report_2018.pdf.
9. Quality and Outcomes Framework (QOF) business rules v 38 2017-2018 October code release.NHS Digital. Leeds, 2018. https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof/quality-and-outcome-framework-qof-business-rules/quality-and-outcomes-framework-qof-business-rules-v-38-2017-2018-october-code-release.
10. Joint Formulary Committee: British National Formulary (online). London: BMJ Group and Pharmaceutical Press,http://www.medicinescomplete.com
11. English Indices of Deprivation 2015 - LSOA Level. Ministry of Housing, Communities and Local Government. London, 2018. https://data.gov.uk/dataset/8f601edb-6974-417e-9c9d-85832dd2bbf2/english-indices-of-deprivation-2015-lsoa-level
12. Lambeth, State of the borough, 2016. Lambeth Council. London, 2017. https://www.lambeth.gov.uk/sites/default/files/State%20of%20Borough%202016%20-%20v3.pdf
13. Statacorp. Stata Statistical Software: Release 14 2015. Available from: http://www.stata.com/
14. Alluvial Diagrams in ggplot2. Brunson, JC, 2018. https://cran.r-project.org/web/packages/ggalluvial/vignettes/ggalluvial.html
15. Harwich E, Lasko-Skinner R. Making NHS data work for everyone. Reform. London, 2018. https://reform.uk/sites/default/files/2018-12/Making%20NHS%20data%20work%20for%20everyone%20WEB_1.pdf
16. Kendrick T, Stuart B, Newell C, *et al*. Changes in rates of recorded depression in English primary care 2003-2013: Time trend analyses of effects of the economic recession, and the GP contract quality outcomes framework (QOF). *J Affect Disord* 2015;180:68-78. https://doi.org/10.1016/j.jad.2015.03.040
17. Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37-43. https://doi.org/10.1016/S0140-6736(12)60240-2

18. McLean G, Guthrie B, Mercer SW, *et al*. General practice funding underpins the persistence of the inverse care law: cross-sectional study in Scotland. *Br J Gen Pract* 2015;65:e799-805. https://doi.org/10.3399/bjgp15X687829
19. Leiser S, Deruaz-Luyet A, N'Goran AA, *et al*. Determinants associated with deprivation in multimorbid patients in primary care-A cross-sectional study in Switzerland. *PLoS One* 2017;12:e0181534. https://doi.org/10.1371/journal.pone.0181534
20. McLean G, Gunn J, Wyke S, *et al*. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. *Br J Gen Prac*t 2014;64:e440-7. https://doi.org/10.3399/bjgp14X680545
21. Lambeth. Demographic factsheet, May 2015. Lambeth & Southwark's Public Health Intelligence Team, 2015. https://www.lambeth.gov.uk/sites/default/files/ssh-lambeth-demography-2015.pdf
22. Multimorbidity: a priority for global health research. Academy of Medical Sciences. London, 2018. https://acmedsci.ac.uk/file-download/82222577
23. Ruel G, Levesque JF, Stocks N, *et al*. Understanding the evolution of multimorbidity: evidences from the North West Adelaide Health Longitudinal Study (NWAHS). *PLoS One* 2014;9:e96291. https://doi.org/10.1371/journal.pone.0096291
24. Xu X, Mishra GD, Dobson AJ, *et al*. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. *PLoS Med* 2018;15:e1002516. https://doi.org/10.1371/journal.pmed.1002516
25. St Sauver JL, Boyd CM, Grossardt BR, *et al*. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open* 2015;5:e006413. https://doi.org/10.1136/bmjopen-2014-006413

Figure 1a: Acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 5196); data excluded if simultaneous onset dates.



Long Term Condition label abbreviations (n = 12): CHD = Coronary Heart Disease; STRK = Stroke; AF = Atrial Fibrillation; HF = Heart Failure; DM = Diabetes; CKD = Chronic Kidney Disease; MOBES = Morbid Obesity; Dep = Depression; SMI = Serious Mental Illness; DEM = Dementia; COPD = Chronic Obstructive Pulmonary Disease; CP = Chronic Pain.

Figure 1b: Acquisition sequence of Long Term Conditions; dominant pathways displayed with patient flows ≥ 35 (n = 769).

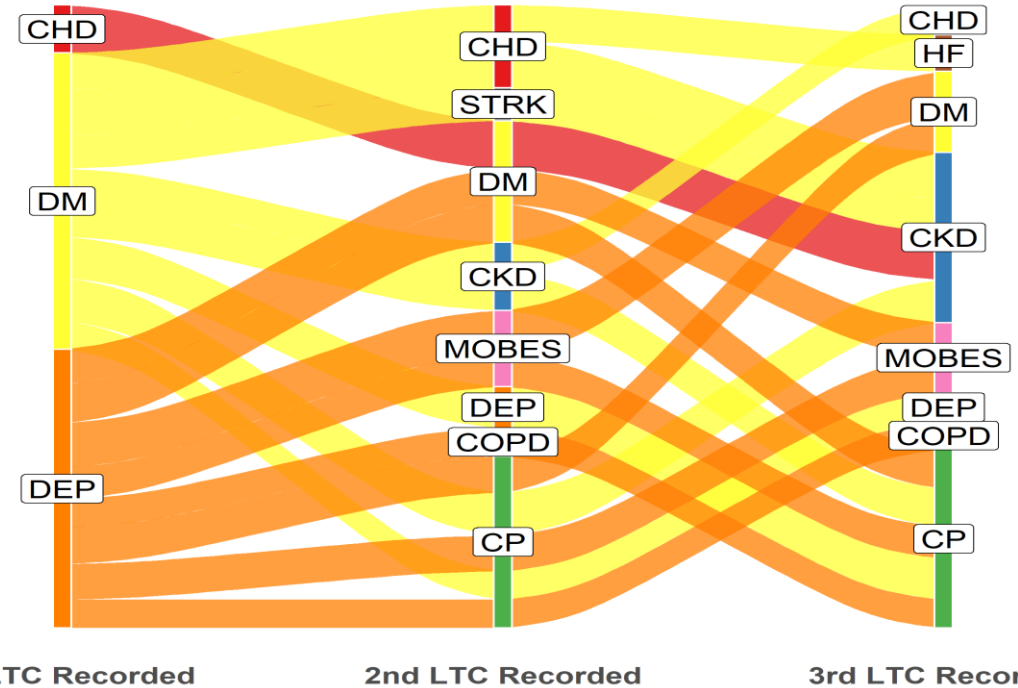


Figure 2a: Most deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1394).

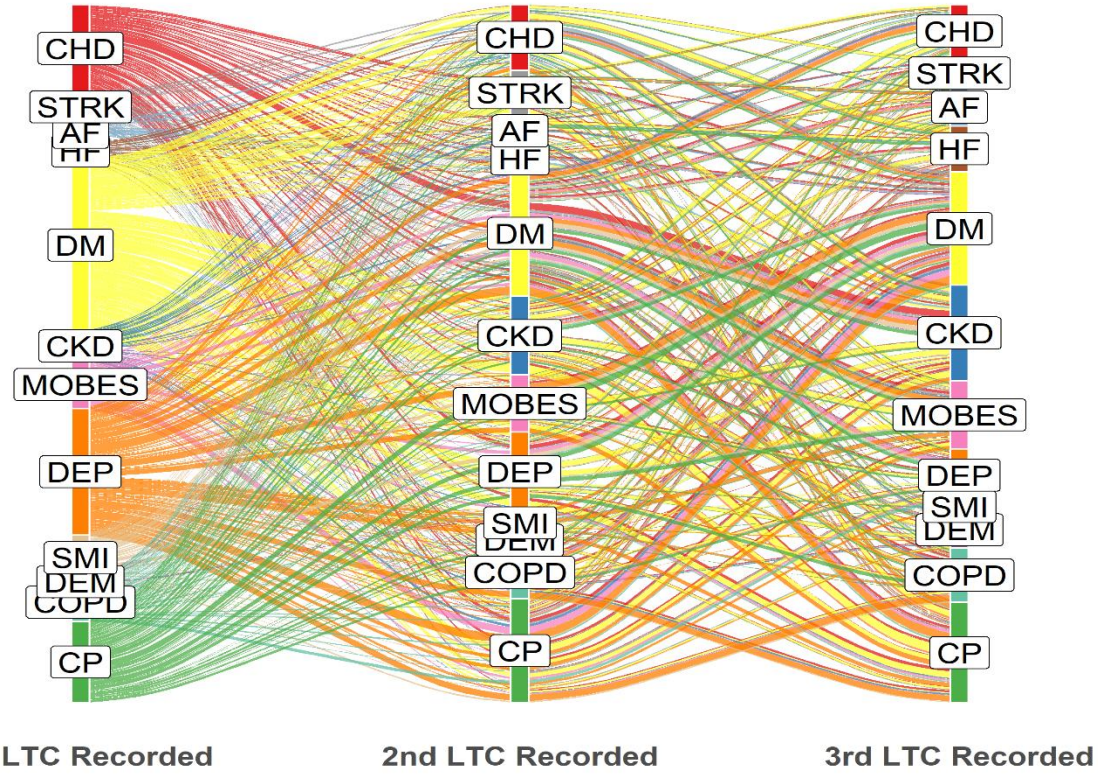


Figure 2b: Most deprived quintile: dominant pathways displayed with patient flows ≥ 13 (n = 145).

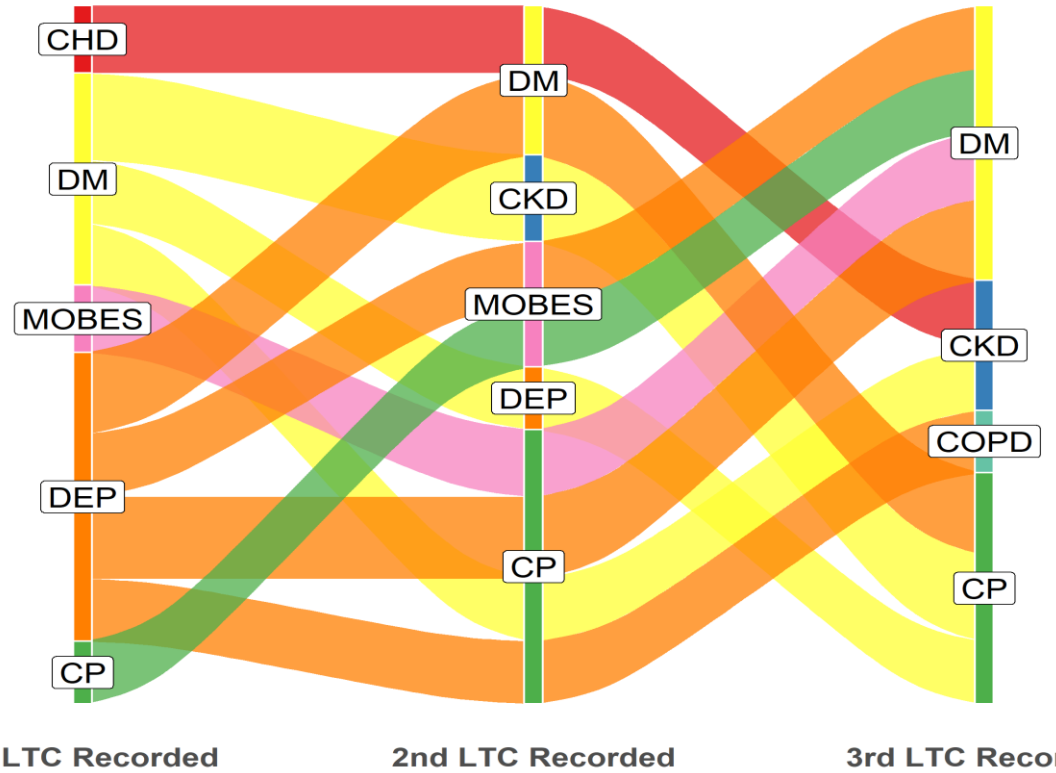


Figure 3a: Least deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 763).

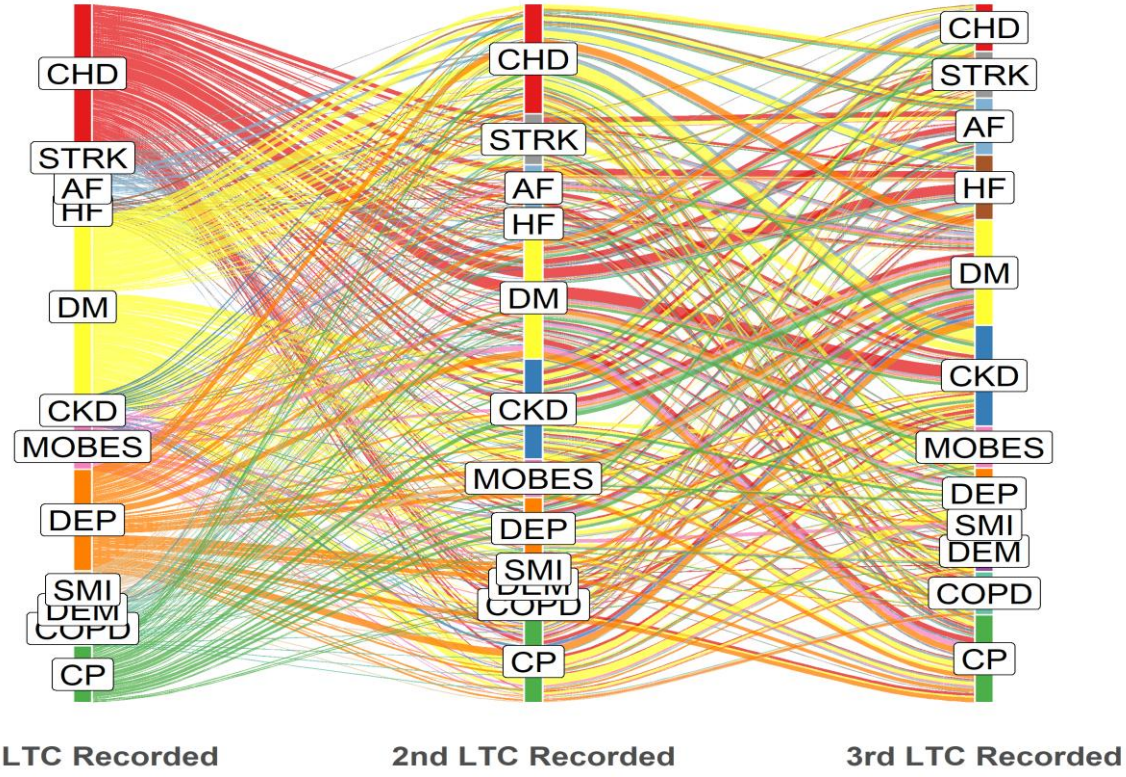


Figure 3b: Least deprived quintile: dominant pathways displayed with patient flows ≥ 8 (n = 89).

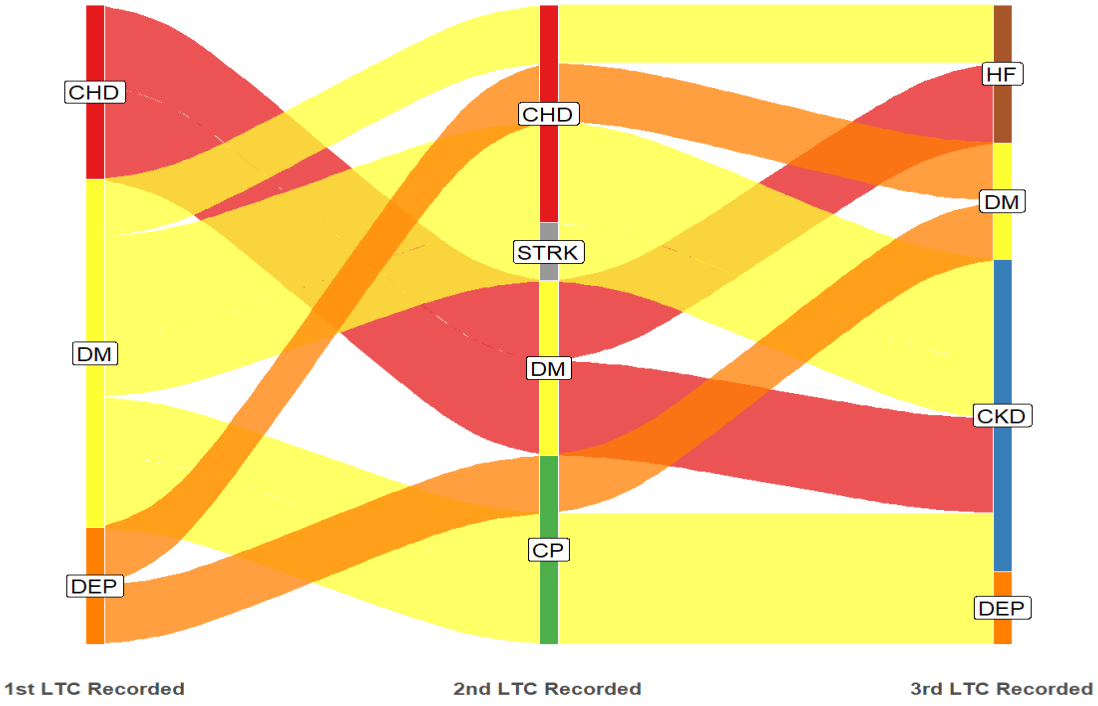


Figure 4a: 'White' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 2788).

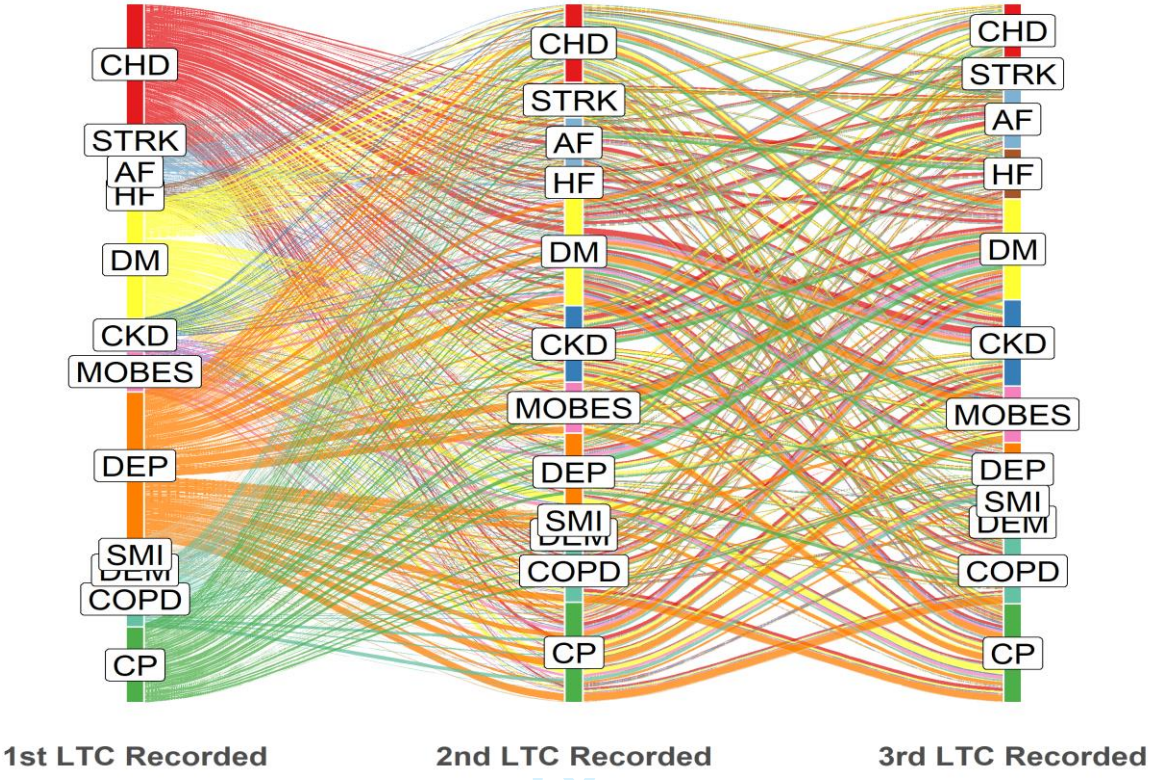


Figure 4b: 'White' ethnic group: dominant pathways displayed with patient flows ≥ 18 (n = 287).

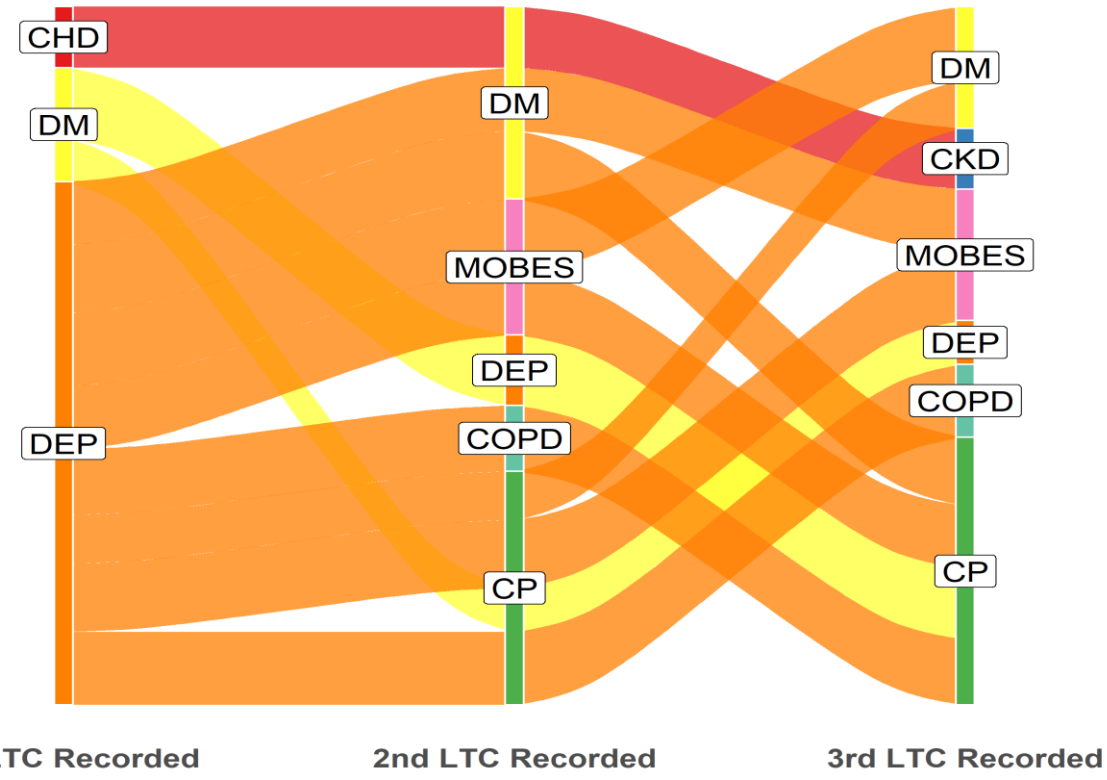


Figure 5a: 'Black' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1448).

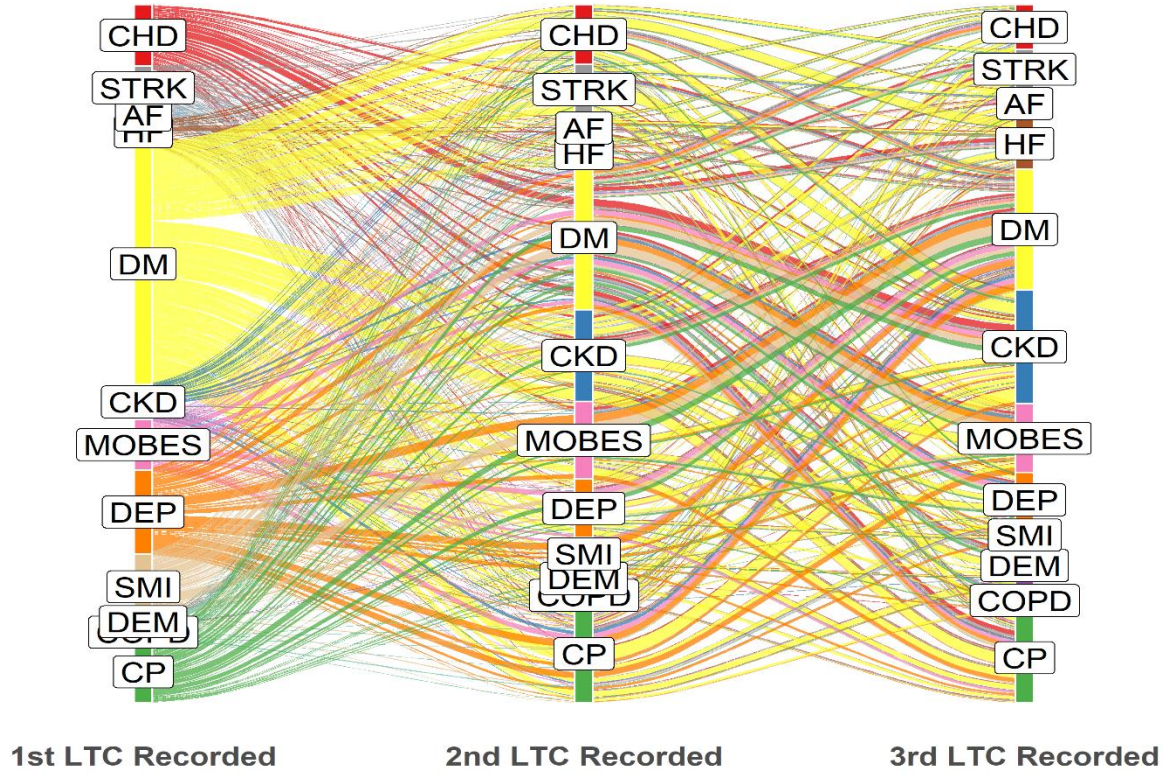


Figure 5b: 'Black' ethnic group: dominant pathways displayed with patient flows ≥ 15 (n = 227).

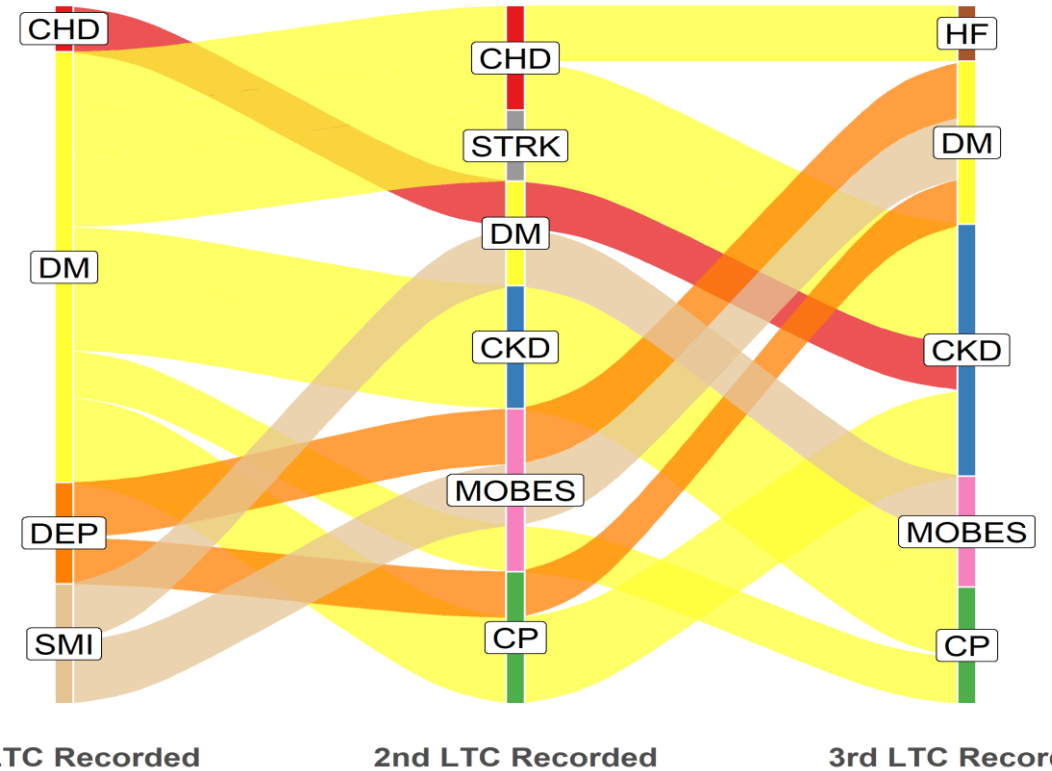


Figure 6a: Age under 65 years: dominant pathways displayed with patient flows ≥ 25 (n = 343).

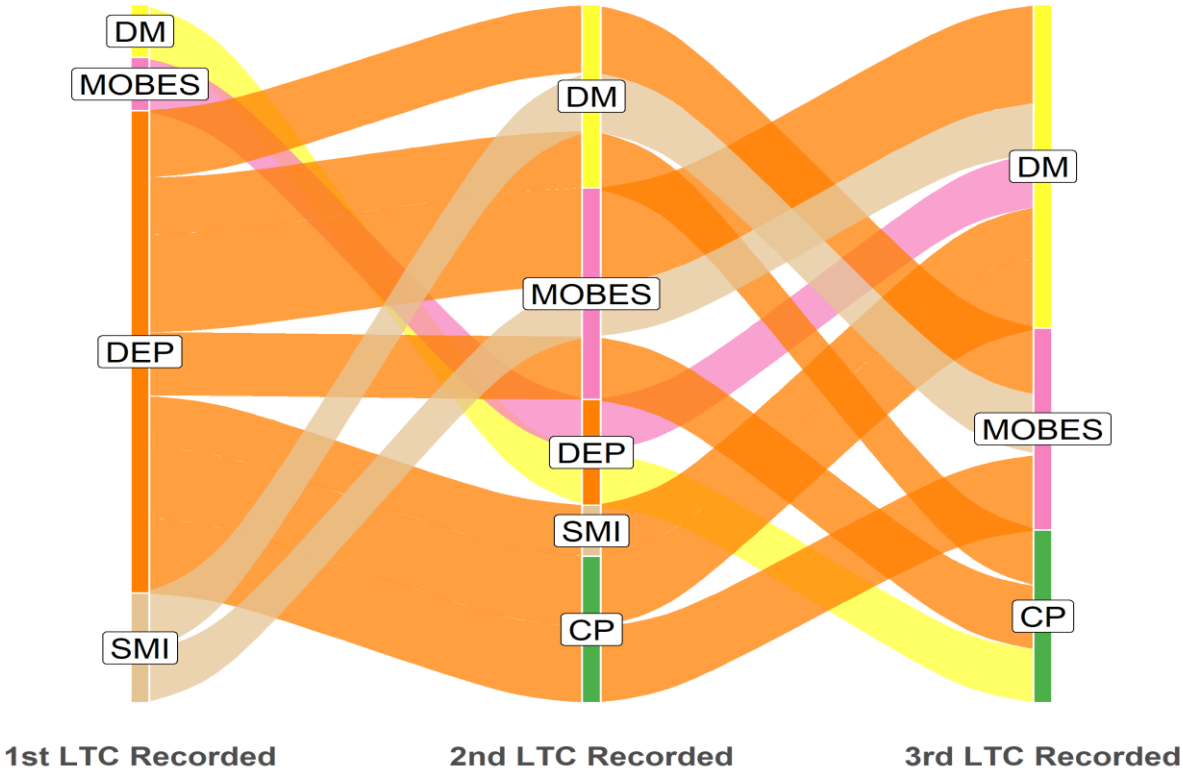
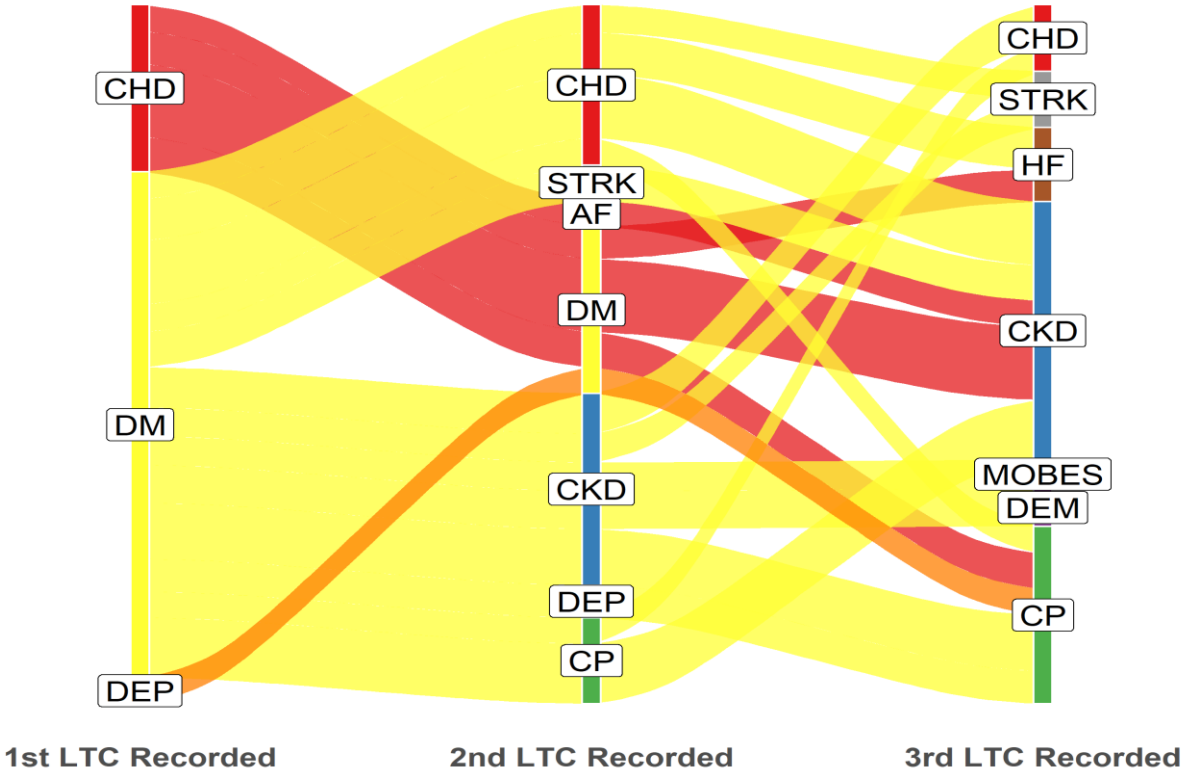


Figure 6b: Age 65 years and over: dominant pathways displayed with patient flows ≥ 20 (n = 536).



Supplementary File

Figure 1a: Area under Receiver Operating Characteristic (ROC) curve = 0.84; based on regression model presented in Table 5.

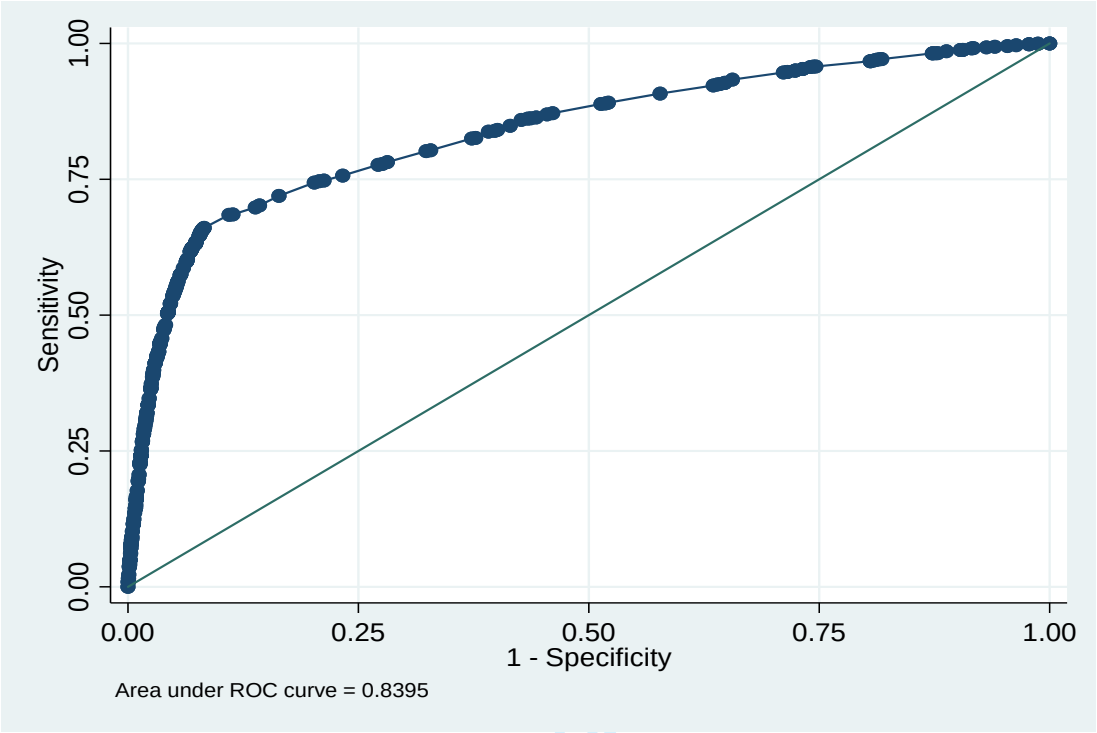
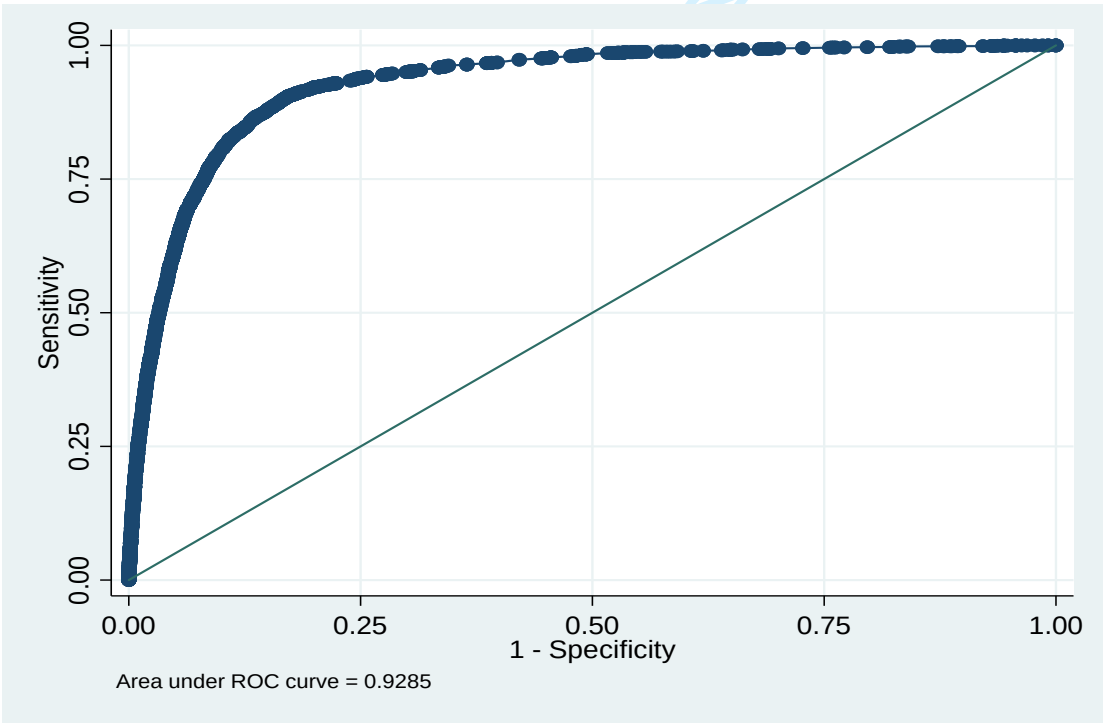


Figure 1b: Area under Receiver Operating Characteristic (ROC) curve = 0.93; based on regression model presented in Table 5 with the addition of three risk factors: Hypertension, Smoking (ever), Obesity (moderate).



STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	"Cross-sectional analysis and longitudinal study"
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1	Summarised in Abstract: Design, Setting, Participants, Main outcome measures, Results
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4	The association between multimorbidity and social determinants and risk factors has not previously been determined in a deprived multi-ethnic community; identifying both the determinants of multimorbidity and the acquisition sequence may suggest interventions to prevent multimorbidity of slow its progression
Objectives	3	State specific objectives, including any prespecified hypotheses	1	"To study the social determinants and risk factors for multimorbidity and the acquisition sequence of multimorbidity"
Methods				
Study design	4	Present key elements of study design early in the paper	5	Presented under 'Data Variables' and 'Data Analysis'.

Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3	Presented under 'Study Setting', 'Study Design', 'Study population'.
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	4-5	Cross-sectional study: all patients registered at general practices in one south London borough (Lambeth) with the exception of those with an 'informed dissent' code in their case-notes. Longitudinal component: the same cohort of patients studied from onset of first Long Term Condition (LTC) to acquisition of 3 or more LTCs.
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5	Summarised under the heading, 'Data Variables': "….routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Data were extracted from the EHR into a secure data warehouse …..and contained information on patient demographic characteristics,

				LTCs, clinical values and medication."
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5	All data was derived from the Electronic Health Record of patients registered at GP practices in the sample area.
Bias	9	Describe any efforts to address potential sources of bias	5-6	Patient records were all included in the analysis, reducing sampling bias. However, 4.0% of patients had an 'informed dissent' code in their case-notes, prohibiting any access to data. We were therefore unable to determine if the omission of these patients may have introduced bias
Study size	10	Explain how the study size was arrived at	5-6	As above – the sample of 332,353 patients represented all patients registered at GP practices in the study area, with the exception of those with informed dissent codes. All Odds Ratios were presented with 95% CI's so that the effect of small numbers on the CI could be seen

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5/6	Quantitative variables analysed according to description in section headed, 'Data Analysis'.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	5/6	Regression modelling – see 'Data Analysis'
		(*b*) Describe any methods used to examine subgroups and interactions	6	Sub-groups of the key predictor variables were stratified (into age bands, deprivation quintiles, major ethnic groups)
		(*c*) Explain how missing data were addressed	5/6	See section on missing demographic codes, missing clinical codes, 'informed dissent codes'.
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a	The study was not a sample
		(*e*) Describe any sensitivity analyses	8	See heading 'Sensitivity analyses'
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6/7	Summarised under 'Multimorbidity cohort characteristics'.
		(b) Give reasons for non-participation at each stage	n/a	n/a (exclusion criteria stated above)
		(c) Consider use of a flow diagram	n/a	Exclusion criteria summarised on pg4-6 without use of a Flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Tables 1,2	Table 1 summarises, 'Frequencies of Long Term Conditions included in the multimorbidity cohort'; Table 2 summarises, 'Demographic characteristics of multimorbidity cohort'.

		(b) Indicate number of participants with missing data for each variable of interest	5/6	Missing data numbers summarised under, 'Multimorbidity cohort characteristics'
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a	
		Cross-sectional study—Report numbers of outcome events or summary measures	13	Summary measures in Table 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Tables 1-5	Univariable analysis presented in Tables 1, 2. Multivariable analysis presented in Tables 3,4,5. Table 3 includes different confounder variables to Tables 4,5. All confounder variables are presented in the Tables.
		(*b*) Report category boundaries when continuous variables were categorized	Tables 2-5	Continuous variables were categorised: 'age' into 3 age bands; social deprivation into quintiles
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a	The study was about the high risk of multimorbidity in specified groups, comparing the power of risk factors.

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Table 5; Supplementary file 1	Sensitivity analyses summarised: logistic regression adjusted for clustering; Receiver Operating Characteristic curve
Discussion				
Key results	18	Summarise key results with reference to study objectives	13/14	Key finding summarised at opening of Discussion. The demographic and risk factors for multimorbidity are defined. The acquisition sequence of multimorbidity is described.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14/15	Discussed under heading, 'Strengths and Limitations'
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-17	Cautious interpretation summarised with reference to the literature. Cautious conclusion reached about possible interventions to prevent or delay multimorbidity onset.
Generalisability	21	Discuss the generalisability (external validity) of the study results	14/15	Summarised under, 'Strengths and Limitations'. Findings were derived from one deprived, multi-ethnic community and may not generalise to less deprived, less diverse populations.
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18/19	The study was funded by the Guy's and St Thomas' Charity

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

The journey to multimorbidity: a longitudinal analysis exploring cardiovascular risk factors and determinants in an urban setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031649.R1
Article Type:	Original research
Date Submitted by the Author:	29-Aug-2019
Complete List of Authors:	Ashworth, Mark; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Durbaba, Stevo; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Whitney, David; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Crompton, James; London Borough of Lambeth, Public Health, Adults and Health Wright, Michael; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Dodhia, Hiten; London Borough of Lambeth, Public Health
Primary Subject Heading:	General practice / Family practice
Secondary Subject Heading:	Public health
Keywords:	PRIMARY CARE, Multimorbidity, Demographic factors, Risk factors

TITLE:

The journey to multimorbidity: a longitudinal analysis exploring cardiovascular risk factors and determinants in an urban setting

AUTHORS:

Mark Ashworth[1], Stevo Durbaba[2], David Whitney[3], James Crompton[4], Michael Wright[5] Hiten Dodhia[6].

1. Reader in Primary Care, School of Population Health & Environmental Sciences, King's College London, London, UK
2. e-Resources Developer, School of Population Health & Environmental Sciences, King's College London, London, UK
3. Data Analyst, School of Population Health & Environmental Sciences, King's College London, London, UK
4. Head of Health Intelligence, Public Health, Adults and Health, London Borough of Lambeth, London, UK
5. Project Manager, School of Population Health & Environmental Sciences, King's College London, London, UK
6. Consultant in Public Health, London Borough of Lambeth, London, UK

Corresponding author email: mark.ashworth@kcl.ac.uk

Word count: 3621

ABSTRACT

Objective. To study the social determinants and cardiovascular risk factors for multimorbidity and the acquisition sequence of multimorbidity.

Design. Longitudinal study based on anonymised primary care data.

Setting. General practices in an urban multi-ethnic borough in London, UK.

Participants. 332,353 patients aged ≥18 years

Main outcome measures. Clinical and socio-demographic characteristics of patients with multimorbidity, defined as ≥3 of 12 Long Term Conditions (LTCs) selected according to high predicted healthcare utilisation. Multilevel logistic regression was used to model the social determinants and risk factors for multimorbidity. Alluvial plots were constructed to illustrate multimorbidity acquisition sequences according to age, ethnicity and social deprivation.

Results. 5597 (1.7%) patients had ≥3 selected LTCs, the 'multimorbidity cohort'. The commonest LTCs were diabetes (63.0%) and chronic pain (42.8%). Social deprivation and ethnicity were independent determinants of multimorbidity: for most compared to least deprived quintile, Adjusted Odds Ratio (AOR) 1.56 (95%CIs, 1.41, 1.72); for South Asian compared to white ethnicity, AOR 1.44 (95%CIs, 1.29, 1.61); for black compared to white ethnicity, AOR 0.86 (95%CIs, 0.80, 0.92). The included risk factors were relatively strong determinants of multimorbidity: hypertension, AOR 5.05 (95%CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95%CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45). The most common initial onset conditions were diabetes and depression; diabetes particularly in older and black ethnic groups; depression particularly in younger, more deprived and white ethnicity groups. Chronic pain was less common as an initial condition.

Conclusions Our findings confirm the importance of age, social deprivation and ethnicity as determinants of multimorbidity. Smoking, obesity and hypertension as risk factors were stronger determinants of multimorbidity than deprivation or ethnicity. The acquisition sequence of multimorbidity is patterned by socio-demographic determinants. Understanding onset conditions of multimorbidity and cardiovascular risk factors may lead to the development of interventions to slow the progression of multimorbidity.

Strengths and limitations

- This study uses a definition of multimorbidity based on Long Term Conditions (LTCs) with high predicted healthcare utilisation rates
- Multimorbidity is studied in a deprived, multi-ethnic community
- Longitudinal data is used to identify the acquisition sequence of multimorbidity and how this is influenced by socio-demographic determinants
- Difficulties gaining access to anonymised primary care data limited the sample size and may have contributed to selection bias
- Coding validation and completeness restricted the analysis of available primary care data

BACKGROUND

Healthcare utilisation is increasingly driven by multimorbidity (1). Each long-term condition (LTC) included within a definition of multimorbidity is likely to generate multiple primary care consultations with general practitioners (GPs), practice nurses and other health care professionals, and may result in Accident and Emergency (A&E) attendances, referral to out-patient appointments and hospital admissions.

Estimates of healthcare utilisation attributable to multimorbidity vary according to the LTCs included within a definition of multimorbidity. There is no standard definition of multimorbidity. One systematic review noted that the number of included LTCs ranged from five to 185 with estimates of population prevalence ranging from 13.1% to 71.8% depending on the number of included conditions (2). In one recent UK study in which multimorbidity was defined as two or more of a selection of 36 LTCs, 27.2% of the population had multimorbidity, accounting for 52.9% of GP consultations and generating a median of nine annual GP consultations (3). LTCs also contribute to the development of frailty, itself a driver of healthcare utilisation (4).

In response to high demand for health and social care, many healthcare providers and commissioners have sought to identify those patients with greatest needs through a process termed 'risk stratification'. Several electronic tools have been developed to offer population based risk stratification (5). An alternative approach is the use of expert panels to define high-demand patient groups (6). Having identified the cohort of patients with the greatest requirement for health and social care services, the purpose of this process is to guide resource allocation on a needs basis, often with the implicit assumption that additional investment in a primary care setting may reduce demand for more expensive secondary care services. Although funding and healthcare need should align, there is little evidence that investment in additional community resources for those most at risk of hospital attendance results in overall reductions in secondary care utilisation (7).

Following a consultation exercise and report from an expert panel, two inner London boroughs and their commissioning groups made the decision to define multimorbidity based on predicted high healthcare and social care demand. Multimorbidity in this context consisted of three or more LTCs considered most likely to result in functional impairment and high service demand. This 'high service demand' definition was used to identify a cohort to receive a package of integrated care, termed 'care coordination'. A focus on this narrowly defined category of multimorbidity would inevitably mean that the proportion of patients defined as 'multimorbid' would be lower than reported in studies based on broader definitions of multimorbidity (2,3).

We aimed to study the characteristics of this multimorbidity cohort, defining both the socio-demographic determinants and cardiovascular risk factors associated with multimorbidity acquisition. Then to determine the acquisition sequence of multimorbidity and the influence of socio-demographic factors on this sequence. The aim was to study the characteristics of the multimorbidity cohort. The main objectives were to define both the socio-demographic determinants and cardiovascular risk factors associated with multimorbidity acquisition; also to determine the acquisition sequence of multimorbidity and the influence of demographic factors on this sequence.

METHODS

Study setting

Our study was set in Lambeth, one of the two inner London UK boroughs adopting the 'care coordination' definition of multimorbidity (8). The population sample consisted of all patients registered at all general practices in Lambeth, with the exception of patients who had opted out of anonymised data sharing for research purposes.

Study design

We conducted a longitudinal analysis based on anonymised coded primary care data extracted from electronic health records (EHR) held in primary care.

Study population

We included data on all patients aged 18 years and over registered with a general practice. For the population with multimorbidity, we included all those with LTCs recorded in the EHR and included in the 'care coordination' definition of multimorbidity: Atrial Fibrillation (AF), Chronic Obstructive Pulmonary Disease (COPD), Chronic Pain (CP), Chronic Kidney Disease (CKD), Coronary Heart Disease (CHD), Diabetes (DM), Dementia, Depression, Heart Failure (HF), Serious Mental Illness (SMI), Stroke, Morbid Obesity. The definition and specified codes for each condition was that used by the Quality and Outcomes Framework (QOF), based on 'QOF38' definitions (9). Two of the conditions selected for inclusion within the definition of multimorbidity were not included within the QOF: chronic pain, defined on the basis of two or more repeat prescriptions for opioid analgesics (British National Formulary, chapter 4.7.2) or neuropathic pain medication (British National Formulary, chapter 4.7.3) (10); morbid obesity defined as a Body Mass Index ≥40 kg/m^2. For each LTC, the date of onset was obtained from the EHR and used in the longitudinal analysis.

Demographic data consisted of gender, age in years and self-ascribed ethnicity obtained from the EHR. Social deprivation data derived from residency data was based on the Index of Multiple Deprivation 2010 classification at Lower Super Output Area (LSOA), stratified into locally based quintiles (11). Local deprivation quintiles were used in place of national quintiles since mean deprivation levels are high in Lambeth, the 22nd most deprived local authority (out of 326) in England (12).

Cardiovascular risk factors included in the analysis were: hypertension (defined as patients on the QOF Hypertension register); moderate obesity (defined as a Body Mass Index of 30.0-39.9; note that patients with BMI ≥40 were included as one of the LTCs within the definition of multimorbidity and therefore not included as a risk factor); smoking (patients with any record of being a smoker).

Data variables

The data consisted of 'real world', routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Data were extracted from the EHR into a secure data warehouse and contained information on patient demographic characteristics, LTCs, clinical values and medication. The data used in this study were extracted in May 2018.

Data analysis

We analysed socio-demographic (age, gender, ethnicity), social (area level deprivation) and cardiovascular risk factor (hypertension, moderate obesity, smoking status) data for the multimorbidity cohort and general population using univariable statistical methods applied at patient-level. Socio-demographic and risk factor determinants of multimorbidity were analysed at patient level using mixed effects multilevel logistic regression models, adding fixed effects to describe patient characteristics and random effects to describe general practice level, based on the registered general practice of each patient in the study sample. We also conducted a sensitivity analysis using mixed effects logistic regression models to allow for random effects, adjusted for clustering at the practice level. The sensitivity analysis allowed pseudo-r^2 values and Receiver Operating Characteristic (ROC) curves to be derived. Analysis was conducted using the statistical software package STATA IC 15 (13).

The acquisition sequence for patients in the multimorbidity cohort was established by searching the EHR for date of onset of each LTC. For this analysis, we established the order of acquisition and tabulated the frequency of first, second and third LTCs. We displayed findings using alluvial plots, an infographic allowing representation of multiple pathways. These were constructed using the software R, and the packages 'ggplot2' and 'ggalluvial' (14).

Patient and public involvement

The borough based and statutory organisation, Lambeth HealthWatch, represented the interests of patients and public in this work; they contributed to the original protocol design and shared in dissemination of the findings.

RESULTS

Multimorbidity cohort characteristics

The final study population consisted of 332,353 patients aged ≥18 years. Data from 13,369 (4.0%) patients had been excluded because a data sharing opt-out code was recorded in their EHR. Patients were included in the final sample even though some socio-demographic data were missing: 3289 (0.99%) patients could not be linked to a LSOA and therefore had missing IMD-2015 score data; ≤10 patients had missing coded gender data; ≤10 patients had missing coded age data. Patients with any category of missing data (n = 3301) were excluded from the multivariable analysis.

In all, 5597 (1.7%) patients had a record of three or more of the selected LTCs, the 'multimorbidity cohort'. Most (n = 3542) of this cohort had three LTCs (63.3%); 1333 (23.8%) had four LTCs; 492 (8.8%) had 5 LTCs; the remaining 230 (4.1%) had more than 5 LTCs. Of the remaining population, 45,241 (13.6%) had one LTC and 10,992 (3.3%) had two LTCs.

A summary of LTC frequencies within the multimorbidity cohort is displayed in Table 1. The most common LTCs within this cohort were: DM (63.0%) and chronic pain (42.8%). In contrast, the most common of the included LTCs in the adult general population were: depression (8.4%), DM (5.4%) and morbid obesity (3.2%).

The socio-demographic characteristics of the multimorbidity cohort are displayed in Table 2. Within the cohort, 33.9% were aged under 65 years (compared with 91.7% in the sample population); 27.7% were of a 'Black' ethnicity (18.0% in the sample population) and 46.0% were born in the UK (45.2% in the sample population). The mean age for multimorbid patients in least and most deprived

quintiles was 73.0 (standard deviation (SD), 13.5) and 69.3 (SD, 12.9) years, respectively; and for the white, black and south Asian populations was 71.2 (SD, 13.3), 69.4 (SD, 14.3) and 71.9 (SD, 11.8) years, respectively.

Multimorbidity cohort socio-demographic determinants

Socio-demographic determinants of the multimorbidity cohort are displayed in Table 3. Based on the adjusted odds ratio (AOR) derived from the multilevel regression model, the strongest determinant for multimorbidity was related to age. After adjustment for age, both social deprivation (the more deprived quintiles) and ethnicity (Black and south Asian ethnicities) remained significant determinants, albeit with much smaller odds ratios. The AOR for black and South Asian compared with white ethnicity was 1.15 (95% CIs, 1.07, 1.23) and 1.19 (95% CIs, 1.07, 1.33), respectively; for most compared to least deprived quintile, the AOR was 1.83 (95% CIs, 1.66, 2.02).

Multimorbidity cohort risk factor determinants

Addition of the three risk factors included in the study attenuated the odds ratios for multimorbidity related to age (Table 4). Social deprivation remained a determinant of multimorbidity: for most compared to least deprived quintile, the AOR was 1.56 (95% CIs, 1.41, 1.72). South Asian ethnicity remained a significant determinant of multimorbidity: AOR 1.44 (95% CIs, 1.29, 1.61) but black ethnicity was no longer a positive determinant: AOR 0.86 (95% CIs, 0.80, 0.92).

The three risk factors were significant determinants of multimorbidity: hypertension, AOR 5.05 (95% CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95% CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45).

Sensitivity analyses

Re-analysis of the determinants of multimorbidity using regression modelling adjusted for clustering at practice level resulted in similar adjusted odds ratios to those obtained in the primary analyses (Table 5). We explored goodness-of-fit through derived pseudo-r^2 values for demographic

determinants: pseudo-r^2 = 0.22, and for risk factor adjusted determinants (hypertension, moderate obesity, smoking): pseudo-r^2 = 0.32. The areas under the ROC curve for each model were 0.84 and 0.93, respectively (Supplementary file: Figures 1, 2).

Multimorbidity acquisition sequence

Of the 5597 patients in the multimorbidity cohort, 5196 had three distinct dates of onset for each of their three or more component LTCs. The remaining 401 (7.2%) patients had identical dates of onset recorded for two or three of their first three LTCs and therefore could not be classified into a sequence. Alluvial plots were constructed displaying the acquisition sequence for LTCs and edited to display dominant flows of patients in each category. Figure 1 displays the three most common starting conditions and subsequent most commonly acquired second and third LTCs. The unedited alluvial plot displaying all patient flows is shown in the Supplementary file (Supplementary File, Figure 3).

The alluvial plots illustrate that diabetes and depression were the most common starting conditions for patients with multimorbidity; diabetes was also relatively common as the second or third acquired LTC whereas depression was predominantly a first-onset LTC (Figure 1).

In the most deprived quintile, diabetes and depression were the most common starting conditions whereas in the least deprived quintile, diabetes and CHD were more common as starting conditions than depression (Figures 2, 3). Unedited alluvial plots comparing most and least deprived quintiles are shown in the Supplementary file (Supplementary File, Figures 4, 5).

Multimorbidity in the white ethnic group was dominated by depression as the starting condition, whereas in the black ethnic group diabetes was the most common starting condition, with depression and SMI also relatively common (Figures 4, 5). Relatively small numbers resulted in poor definition of the alluvial plot in the South Asian group and this figure is not presented. Unedited

alluvial plots comparing black and white ethnic groups are shown in the Supplementary file (Supplementary File, Figures 6, 7).

Multimorbidity in the under 65 year old cohort was dominated by depression as the starting condition, with SMI also relatively common; in the ≥65 year old cohort, diabetes and CHD were the most common starting conditions (Figures 6, 7).

Chronic pain appeared to be more common as a second or third acquired LTC but less common as a first LTC. This sequence was apparent in the overall picture, in the pattern displayed by least and most deprived quintiles, for black and white ethnicities and for younger and older age cohorts.

Morbid obesity was among the more common starting conditions in the most deprived cohort and younger age cohort. However, in other socio-demographic samples, morbid obesity was more common as a second and third acquired LTC.

DISCUSSION

We report on multimorbidity using a definition originating from a health service commissioning perspective, consisting of three or more of 12 LTCs selected because of likely high impact upon health service and social care utilisation. In total, 1.7% of the adult population in our study sample had multimorbidity according to these narrowly defined criteria. Diabetes and chronic pain were the most prevalent LTCs within this cohort. Independent of age, both ethnicity and social deprivation were significant determinants of multimorbidity. However, the cardiovascular risk factors of hypertension, obesity and smoking were more strongly associated with multimorbidity than social deprivation or ethnicity.

The acquisition sequence of multimorbidity differed substantially according to age, ethnicity and social deprivation. Diabetes and depression were the most common starting conditions overall. Diabetes as a starting condition was notably more common in the older and black ethnic group.

Depression as a starting condition was notably more common in patients who were younger, more deprived and in the white ethnic group. Differences in acquisition sequence between most and least deprived areas and white and black ethnicities (Figures 2 - 5) are unlikely to have been strongly influenced by age differences since mean age was similar for each of these cohorts.

Strengths and Limitations

Data access was a limitation to this analysis. We were only able to obtain data from one of the two boroughs adopting this approach to multimorbidity, the other lacked a data extraction system preventing us from analysing large datasets of patient-level data. Had we gained access to the data, this would have approximately doubled our sample size and enabled further analysis of multimorbidity in deprived, multi-ethnic populations. This difficulty in accessing anonymised data hampers the analysis of patient-level data in many areas of the UK (15). Data coding constrains the analysis of primary care data and we were only able to study the association of multimorbidity with a limited range of risk factors while other known risk factors such as exercise and diet could not be captured. We aimed to display the acquisition sequence of LTCs using alluvial plots. However, these plots do not readily display time data resulting in a lack of clarity in the rate of progression of multimorbidity. A time-to-event analysis is required for identifying those patients who progress rapidly from first LTC into multimorbidity, which is the subject of further study. Similarly, the acquisition sequence could not be determined for a small minority of patients with identical LTC date-of-onset recording by GPs whereas in reality, it is unlikely that the LTC onset dates were simultaneous. Furthermore, as with all studies based on primary care data, there may be coding anomalies which introduce bias into the estimates of LTC prevalence. QOF coding criteria were used for 10 of the included LTCs, standardising the definition. However, the prevalence of conditions such as depression may be underestimated using QOF criteria (16). For the two LTCs not included in the QOF, the definition is dependent on GP coding. Thus 'morbid obesity' was only included in our study if there was a BMI recording which may have resulted in an under-estimate of prevalence. 'Chronic

pain' was defined based on medication consumption whereas many patients with chronic pain may have sought alternatives to analgesic medication resulting in an underestimate of prevalence; conversely, our inclusion criteria of 'two or more prescriptions over the preceding year' may have resulted in an overestimate of prevalence, with some patients recovering from chronic pain during the course of the year. In common with other observational studies, significant associations between multimorbidity and socio-demographic or risk factor determinants may imply, but cannot prove, causality. Whilst interventional studies are required to obtain stronger evidence of causality, causal inference may be derived by time series analyses and further study of potential confounding and residual variance. Finally, the richness of locally based data covering a whole borough with unique socio-demographic characteristics has to be offset against possible loss of generalisability to other areas with very different social deprivation and ethnicity characteristics.

Comparison with the literature

Our cross-sectional data, although conducted in a deprived, multi-ethnic population, are similar to the findings of others reporting on increased multimorbidity prevalence associated with age, social deprivation and ethnic minority status (17). Comparison with other multimorbidity studies is difficult because of the highly restricted definition of multimorbidity used in the current study. Nevertheless, other studies have reported the high prevalence of both diabetes and depression in multimorbidity cohorts (3, 17). A higher prevalence of mental health and physical health LTC combinations has been noted in deprived areas, although in our own study we did not conduct an analysis to identify specific LTC combinations (18). Certain conditions have been found to be more prevalent in deprived communities, also contributing to higher prevalence of multimorbidity in these areas, such as depression and addiction issues in younger deprived populations (19), or more generally, depression, drugs, anxiety, dyspepsia, chronic pain, CHD, DM (20). These findings are aligned to our own reporting of high proportions of multimorbid patients in deprived areas with depression and diabetes. The known influence of population age profiles on the demography of multimorbidity (19)

is illustrated by our finding of markedly differing alluvial plot profiles describing multimorbidity acquisition in a younger cohort dominated by mental health related conditions (Figure 6) and an older cohort dominated by CHD and DM (Figure 7). The inner city population in our study is characterised by a much younger overall age profile compared with the national population: 8% of the population of Lambeth is aged 65 years or more compared to a mean of 18% in England (21). Added to this, the most socially deprived and black ethnic minority groups in our sample were somewhat younger. Both factors are likely to have reduced the overall study prevalence of conditions associated with ageing such as diabetes and CHD. Many multimorbidity studies do not include morbid obesity within their definition (3, 17); we found morbid obesity was particularly associated with social deprivation.

Several publications have reported on combinations and clusters of LTCs but few longitudinal analyses have been reported (22). One study from Australia reported the order of appearance for eight LTCs, reporting in detail for asthma and mood related disorders, the two LTCs most strongly associated with the risk of developing a second LTC. For those with baseline asthma, there was a higher subsequent risk of developing COPD and hypercholesterolaemia; for those with baseline mood disorders, there risk of subsequent asthma, diabetes and other mental disorders was increased (23). In a study of cardiometabolic conditions in Australia, nearly one-quarter of women initially diagnosed with stroke subsequently progressed to other conditions which was a much larger proportion progressing to other conditions than in those initially diagnosed with diabetes (9.9%) or heart disease (11.4%) (24). A further US study explored the acquisition sequence of 20 LTCs describing dyads and triads of conditions and reporting, for example, that the most common triad sequence in 20-39 year olds was depression, asthma and substance misuse whereas in 50-59 year olds it was hyperlipidaemia, hypertension and diabetes (25). They concluded that combinations of LTCs vary extensively by age and sex. Our own findings confirm variation by age and sex, with ethnicity adding to the pattern of variation. Some authors have suggested that the study of acquisition sequence may imply potential interventions to prevent, minimise or delay progression

toward multimorbidity (23). Our findings that three cardiovascular risk factors are more strongly associated with multimorbidity than deprivation and ethnicity, suggest that interventions to reduce the impact of these risk factors may contribute to a reduction in the prevalence of multimorbidity. The control of hypertension, smoking and obesity are often perceived in terms of primary cardiovascular disease prevention or of secondary prevention of single LTCs but may also be conceptualised in terms of multimorbidity prevention. However, a focus on risk factors should not detract from 'the causes of the causes', since social conditions themselves generate causal pathways leading from socio-economic determinants to risk behaviours (26) and interventions which address health behaviours while failing to engage in social determinants may paradoxically result in increased health inequalities (27).

CONCLUSION

We have confirmed the role of age, social deprivation and ethnicity as determinants of multimorbidity in an inner city multi-ethnic population and have extended previous findings demonstrating the way in which the acquisition sequence of multimorbidity is patterned by these determinants. Three cardiovascular risk factors, hypertension, obesity and smoking were stronger determinants of multimorbidity than either deprivation or ethnicity. The strength of these risk factors as determinants suggests interventions which may be effective in reducing the prevalence, delaying the onset or slowing the progression of multimorbidity.

Contributors: MA, HD and MW contributed to the idea and design of the study. SD, DW and JC led on data extraction and preparation. Statistical analysis was conducted by MA and HD. MA produced the first draft of the paper; all co-authors contributed and approved the final draft. MA is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This work was supported by the Guy's and St Thomas' Charity with a grant awarded November 2017. The views expressed are those of the authors and not necessarily those of the funders. The findings presented are independent from the funders who have had no role in data collection, data analysis or writing of the paper.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval

All data were extracted under the terms of a signed Data Sharing Agreement (DSA) with each practice and with project-specific approval following submission of a Data Privacy Impact Assessment (DPIA), approved by Lambeth Clinical Commissioning Group (CCG) on 2/11/17. Information Governance approval required 'low number suppression', ensuring that data could not be displayed if the patient number was 10 or less in any given category; in these circumstances, data reporting would state: '≤10 patients'. Separate Ethical Committee approval was not required (Health Research Authority, personal correspondence, 29/9/17) since all data were fully anonymised for the purposes of research access, and all Patient Identifiable Data (PID) had been removed.

Data sharing: No additional data available.

REFERENCES:

1. Palladino R, Tayu Lee J, Ashworth M, *et al*. Associations between multimorbidity, healthcare utilisation and health status: evidence from 16 European countries. *Age Ageing* 2016;45:431-5. https://dx.doi.org/10.11606%2FS1518-8787.2018052000639
2. Fortin M, Stewart M, Poitras ME, *et al*. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142-51. https://doi.org/10.1370/afm.1337
3. Cassell A, Edwards D, Harshfield A, *et al*. The epidemiology of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2018;68:e245-e51. https://doi.org/10.3399/bjgp18X695465
4. Hanlon P, Nicholl BI, Jani BD, *et al*. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. *Lancet Public Health* 2018;3:e323-e32. https://doi.org/10.1016/S2468-2667(18)30091-4
5. Billings J, Blunt I, Steventon A, *et al*.. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2. http://dx.doi.org/10.1136/bmjopen-2012-001667
6. Huntley AL, Johnson R, Purdy S, *et al*. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med*. 2012;10:134-41. https://doi.org/10.1370/afm.1363
7. Roland M, Abel G. Reducing emergency admissions: are we on the right track? *Br Med J* 2012;345:e6017. https://doi.org/10.1136/bmj.e6017
8. From one to many. Exploring people's progression to multiple long-term conditions in an urban environment. Guy's and St Thomas' Charity. London, 2018. https://www.gsttcharity.org.uk/sites/default/files/GSTTC_MLTC_Report_2018.pdf.
9. Quality and Outcomes Framework (QOF) business rules v 38 2017-2018 October code release.NHS Digital. Leeds, 2018. https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof/quality-and-outcome-framework-qof-business-rules/quality-and-outcomes-framework-qof-business-rules-v-38-2017-2018-october-code-release.
10. Joint Formulary Committee: British National Formulary (online). London: BMJ Group and Pharmaceutical Press,http://www.medicinescomplete.com
11. English Indices of Deprivation 2015 - LSOA Level. Ministry of Housing, Communities and Local Government. London, 2018. https://data.gov.uk/dataset/8f601edb-6974-417e-9c9d-85832dd2bbf2/english-indices-of-deprivation-2015-lsoa-level
12. Lambeth, State of the borough, 2016. Lambeth Council. London, 2017. https://www.lambeth.gov.uk/sites/default/files/State%20of%20Borough%202016%20-%20v3.pdf
13. Statacorp. Stata Statistical Software: Release 14 2015. Available from: http://www.stata.com/
14. Alluvial Diagrams in ggplot2. Brunson, JC, 2018. https://cran.r-project.org/web/packages/ggalluvial/vignettes/ggalluvial.html
15. Harwich E, Lasko-Skinner R. Making NHS data work for everyone. Reform. London, 2018. https://reform.uk/sites/default/files/2018-12/Making%20NHS%20data%20work%20for%20everyone%20WEB_1.pdf
16. Kendrick T, Stuart B, Newell C, *et al*. Changes in rates of recorded depression in English primary care 2003-2013: Time trend analyses of effects of the economic recession, and the GP contract quality outcomes framework (QOF). *J Affect Disord* 2015;180:68-78. https://doi.org/10.1016/j.jad.2015.03.040
17. Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37-43. https://doi.org/10.1016/S0140-6736(12)60240-2

18. McLean G, Guthrie B, Mercer SW, *et al*. General practice funding underpins the persistence of the inverse care law: cross-sectional study in Scotland. *Br J Gen Pract* 2015;65:e799-805. https://doi.org/10.3399/bjgp15X687829
19. Leiser S, Deruaz-Luyet A, N'Goran AA, *et al*. Determinants associated with deprivation in multimorbid patients in primary care-A cross-sectional study in Switzerland. *PLoS One* 2017;12:e0181534. https://doi.org/10.1371/journal.pone.0181534
20. McLean G, Gunn J, Wyke S, *et al*. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. *Br J Gen Prac*t 2014;64:e440-7. https://doi.org/10.3399/bjgp14X680545
21. Lambeth. Demographic factsheet, May 2015. Lambeth & Southwark's Public Health Intelligence Team, 2015. https://www.lambeth.gov.uk/sites/default/files/ssh-lambeth-demography-2015.pdf
22. Multimorbidity: a priority for global health research. Academy of Medical Sciences. London, 2018. https://acmedsci.ac.uk/file-download/82222577
23. Ruel G, Levesque JF, Stocks N, *et al*. Understanding the evolution of multimorbidity: evidences from the North West Adelaide Health Longitudinal Study (NWAHS). *PLoS One* 2014;9:e96291. https://doi.org/10.1371/journal.pone.0096291
24. Xu X, Mishra GD, Dobson AJ, *et al*. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. *PLoS Med* 2018;15:e1002516. https://doi.org/10.1371/journal.pmed.1002516
25. St Sauver JL, Boyd CM, Grossardt BR, *et al*. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open* 2015;5:e006413. https://doi.org/10.1136/bmjopen-2014-006413
26. Ferraro KF, Schaffer MH, Wilkinson LR. Childhood Disadvantage and Health Problems in Middle and Later Life: Early Imprints on Physical Health? *Am Sociol Rev* 2016;81:107-133. https://doi.org/10.1177/0003122415619617
27. Link BG, Phelan J. Social Conditions As Fundamental Causes of Disease. *J Health Soc Behavior* 1995;(Extra Issue): 80-94. DOI: 10.2307/2626958

Table 1. Frequencies of Long Term Conditions included in the multimorbidity cohort (n = 5597 patients) compared with the remainder of the registered population and with the general population aged ≥18 years.

Long Term Condition	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756	Population values: frequency (valid %) N = 332,353
DM	3525 (63.0%)	14,405 (4.4%)	17,930 (5.4%)
Chronic Pain	2397 (42.8%)	5813 (1.8%)	8210 (2.5%)
CKD	2101 (37.5%)	3470 (1.1%)	5571 (1.7%)
CHD	2099 (37.5%)	2668 (0.8%)	4767 (1.4%)
Depression	2086 (37.3%)	25,877 (7.9%)	27,963 (8.4%)
Morbid Obesity	1653 (29.5%)	8883 (2.7%)	10,486 (3.2%)
AF	1254 (22.4%)	1493 (0.5%)	2747 (0.8%)
COPD	1247 (22.3%)	2387 (0.7%)	3634 (1.1%)
Heart Failure	1186 (21.2%)	544 (0.2%)	1730 (0.5%)
Stroke	1095 (19.6%)	1571 (0.5%)	2666 (0.8%)
SMI	692 (12.4%)	4099 (1.3%)	4791 (1.4%)
Dementia	616 (11.0%)	738 (0.2%)	1354 (0.4%)

Table 2. Socio-demographic characteristics of multimorbidity cohort, compared with remainder of registered population aged ≥18 years.

Demographic characteristic	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756
Female gender	3042 (54.4%)	161,445 (49.4%)
Age <65 years	1899 (33.9%)	299,742 (91.7%)
Age ≥65-74 years	1249 (22.3%)	15,992 (4.9%)
Age ≥75-84 years	1479 (26.4%)	8038 (2.5%)
Age ≥85 years	970 (17.3%)	2884 (0.9%)
White	3022 (54.0%)	179,859 (55.0%)
Black	1553 (27.7%)	58,939 (18.0%)

South Asian	469 (8.4%)	22,323 (6.8%)
Mixed	197 (3.5%)	15,177 (4.6%)
Other	100 (1.8%)	9804 (3.0%)
Unknown	256 (4.6%)	40,654 (12.4%)
Country of origin: UK*	1413 (46.0%)	69,675 (45.2%)
Language preference: English*	3755 (84.0%)	240,287 (73.8%)
Social deprivation: 1st quintile (most deprived)*	1500 (26.8%)	63,374 (19.4%)
Social deprivation: 2nd quintile*	1232 (22.0%)	63,073 (19.3%)
Social deprivation: 3rd quintile*	995 (17.8%)	66,409 (20.3%)
Social deprivation: 4th quintile*	1013 (18.1%)	66,334 (20.3%)
Social deprivation: 5th quintile (least deprived)*	828 (14.8%)	64,306 (19.7%)

*missing data with reduction in denominator number.

Table 3. Socio-demographic determinants of the multimorbidity cohort: adjusted odds ratios derived from mixed effects multi-level logistic regression modelling.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.99, 1.11
Age <65 years	(reference group)	
Age ≥65-74 years	11.81	10.96, 12.72
Age ≥75-84 years	27.14	25.22, 29.20
Age ≥85 years	48.70	44.60, 53.18
White	(reference group)	
Black	1.15	1.07, 1.23
South Asian	1.19	1.07, 1.33
Mixed	0.96	0.83, 1.12
Other	0.76	0.62, 0.93
Unknown	0.44	0.38, 0.50
Social deprivation: 1st quintile (most deprived)	1.83	1.66, 2.02
Social deprivation: 2nd quintile	1.58	1.43, 1.75
Social deprivation: 3rd quintile	1.27	1.15, 1.40
Social deprivation: 4th quintile	1.21	1.10, 1.33
Social deprivation: 5th quintile (least deprived)	(reference group)	

Table 4. Socio-demographic determinants of the multimorbidity cohort: odds ratios derived from mixed effects multi-level logistic regression modelling with addition of three risk factors: Hypertension, Obesity (moderate), Smoking (ever).

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.10	1.01, 1.19
Age <65 years	(reference group)	
Age ≥65-74 years	4.01	3.69, 4.36
Age ≥75-84 years	8.12	7.46, 8.83
Age ≥85 years	15.71	14.20, 17.38
White	(reference group)	
Black	0.86	0.80, 0.92
South Asian	1.44	1.29, 1.61
Mixed	0.95	0.81, 1.11
Other	0.83	0.67, 1.03
Unknown	0.60	0.52, 0.69
Social deprivation: 1st quintile (most deprived)	1.56	1.41, 1.72
Social deprivation: 2nd quintile	1.35	1.22, 1.50
Social deprivation: 3rd quintile	1.18	1.06, 1.31
Social deprivation: 4th quintile	1.18	1.06, 1.30
Social deprivation: 5th quintile (least deprived)	(reference group)	
Hypertension register	5.05	4.69, 5.44
Moderate obesity	3.41	3.21, 3.63
Smoker (ever)	2.30	2.16, 2.45

Table 5. Socio-demographic determinants of the multimorbidity cohort: odds ratios derived from mixed effects logistic regression modelling adjusted for clustering at practice level.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.97, 1.13
Age <65 years	(reference group)	
Age ≥65-74 years	12.08	10.53, 13.86
Age ≥75-84 years	27.74	24.28, 31.71
Age ≥85 years	49.84	43.00, 57.78
White	(reference group)	
Black	1.19	1.06, 1.33
South Asian	1.16	1.00, 1.34
Mixed	0.98	0.81, 1.20
Other	0.77	0.62, 0.96
Unknown	0.44	0.36, 0.55
Social deprivation: 1st quintile (most deprived)	1.96	1.69, 2.26
Social deprivation: 2nd quintile	1.66	1.43, 1.92
Social deprivation: 3rd quintile	1.31	1.13, 1.51
Social deprivation: 4th quintile	1.24	1.09, 1.42
Social deprivation: 5th quintile (least deprived)	(reference group)	

Captions and Legends to Figures:

Figures 1 - 7 display edited alluvial plots showing dominant patient pathways from acquisition of first to second and third Long Term Conditions.

Legend for all Figures:
Long Term Condition label abbreviations (n = 12): CHD = Coronary Heart Disease; STRK = Stroke; AF = Atrial Fibrillation; HF = Heart Failure; DM = Diabetes; CKD = Chronic Kidney Disease; MOBES = Morbid Obesity; Dep = Depression; SMI = Serious Mental Illness; DEM = Dementia; COPD = Chronic Obstructive Pulmonary Disease; CP = Chronic Pain.

Figure 1: Acquisition sequence of Long Term Conditions; dominant pathways displayed with patient flows ≥ 35 (n = 769).

Figure 2: Most deprived quintile: dominant pathways displayed with patient flows ≥ 13 (n = 145).

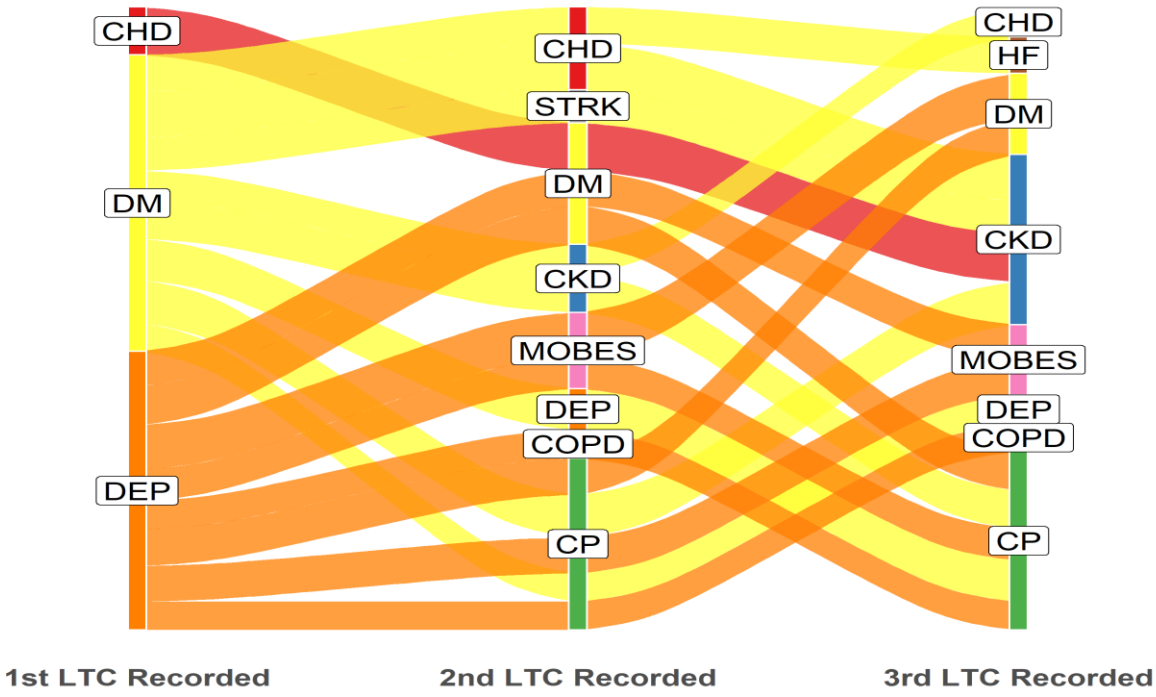
Figure 3: Least deprived quintile: dominant pathways displayed with patient flows ≥ 8 (n = 89).

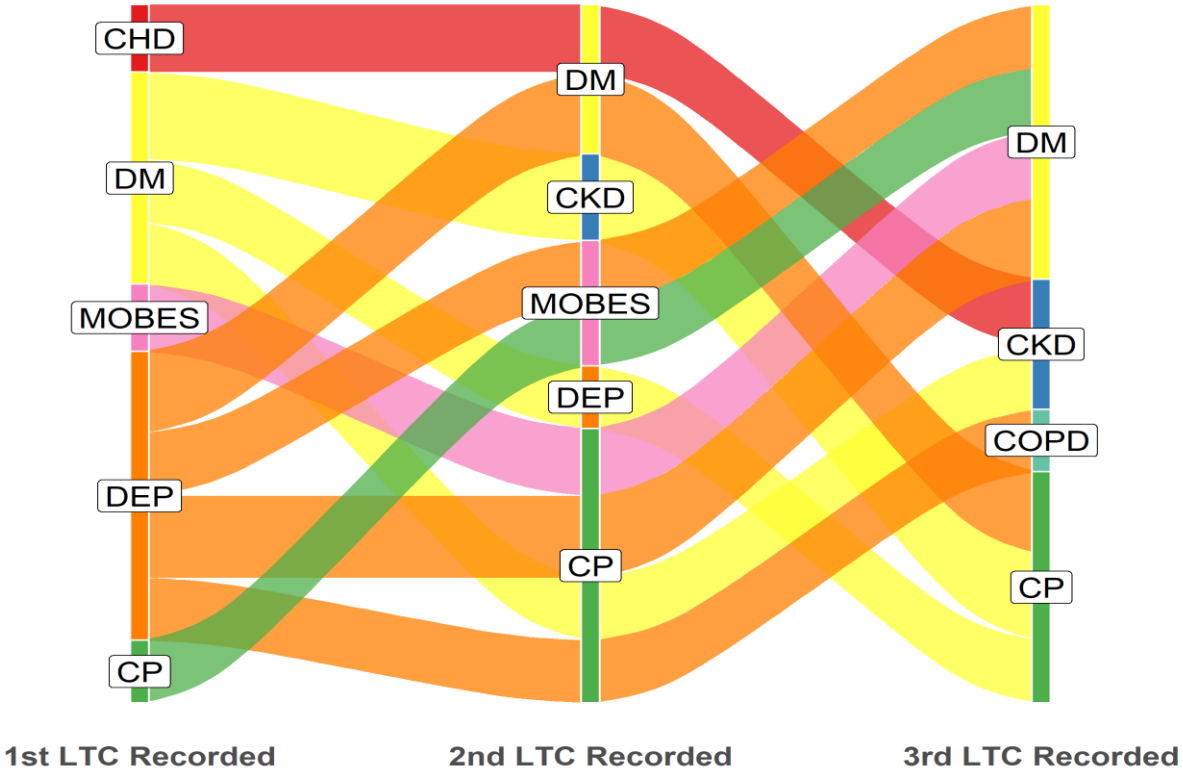
Figure 4: 'White' ethnic group: dominant pathways displayed with patient flows ≥ 18 (n = 287)

Figure 5: 'Black' ethnic group: dominant pathways displayed with patient flows ≥ 15 (n = 227)

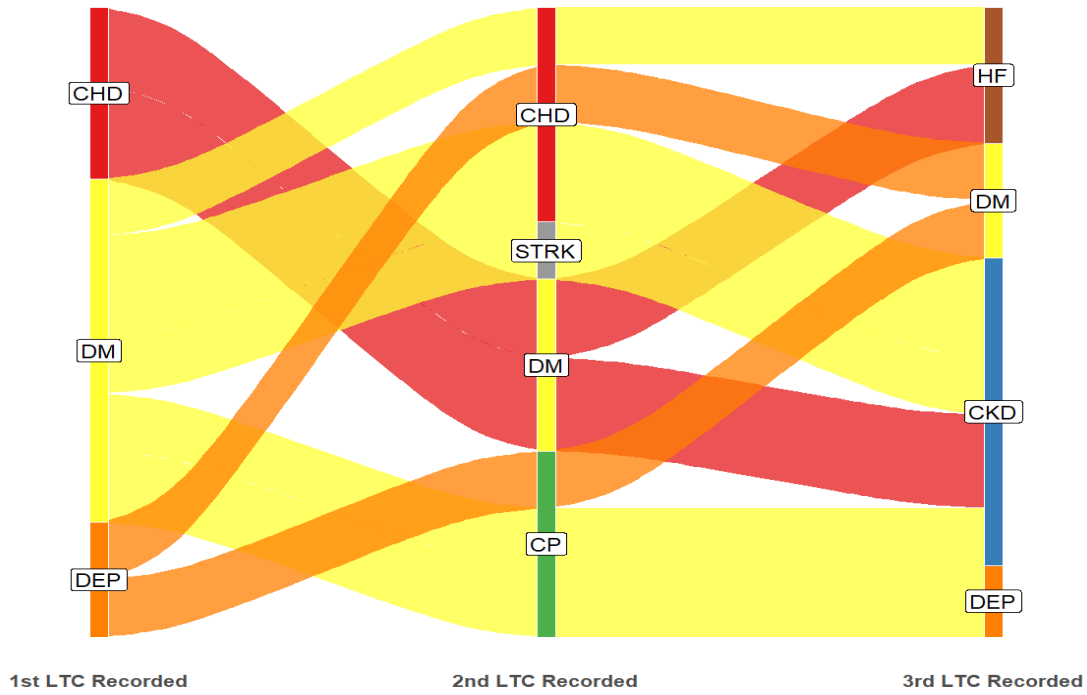
Figure 6: Age under 65 years: dominant pathways displayed with patient flows ≥ 25 (n = 343).

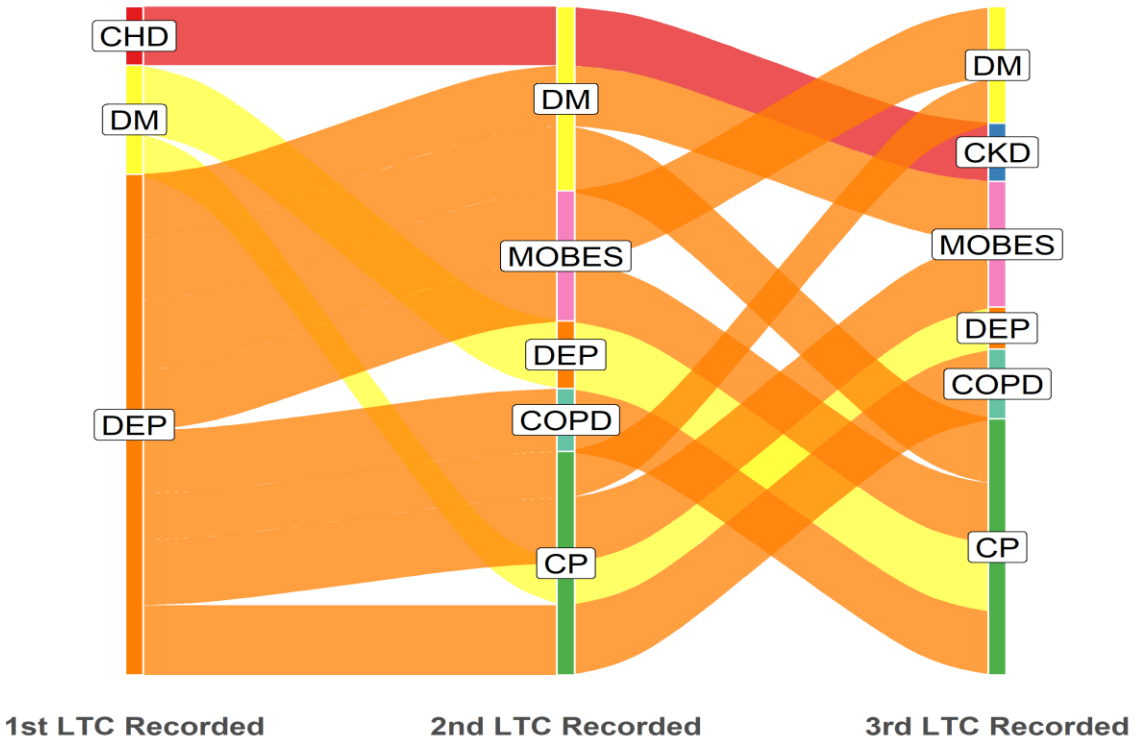
Figure 7: Age 65 years and over: dominant pathways displayed with patient flows ≥ 20 (n = 536).

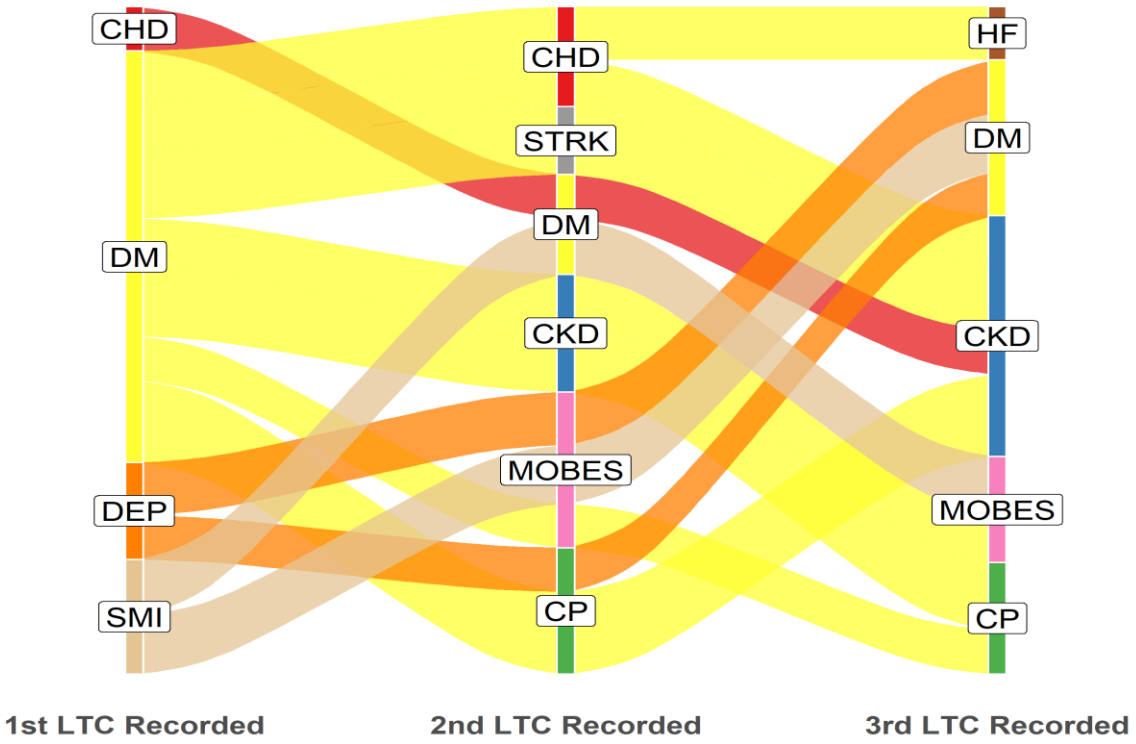


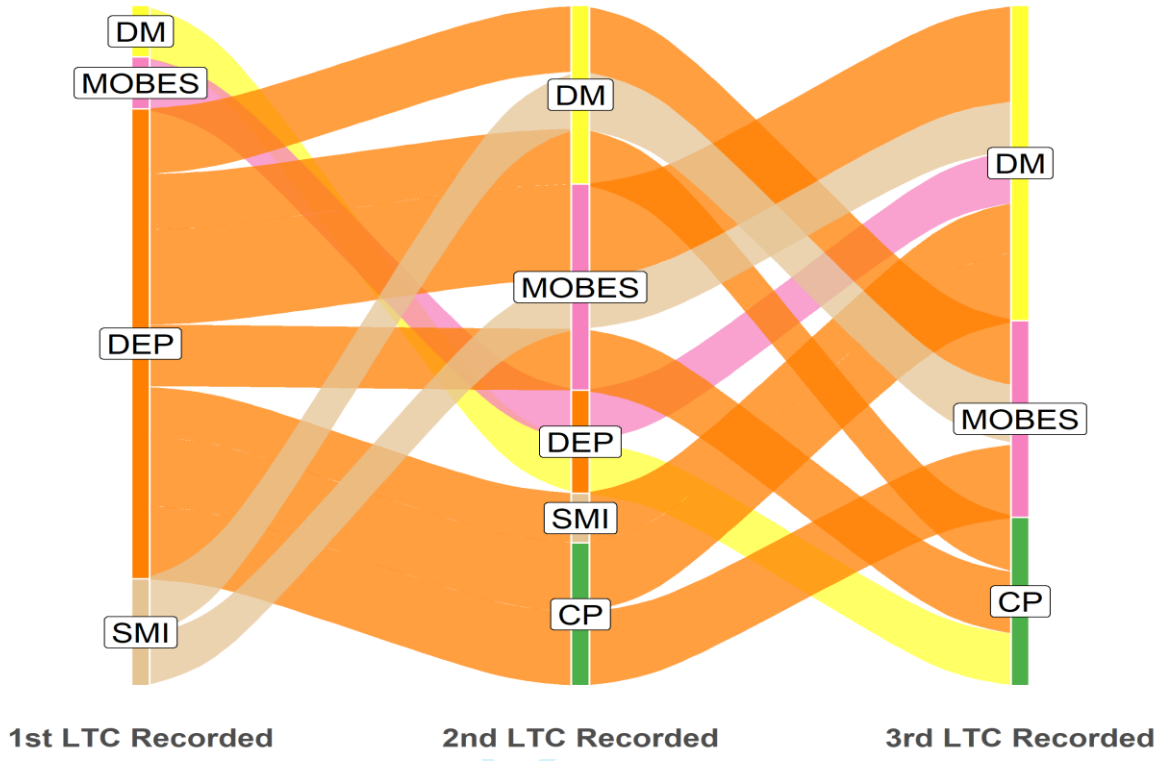


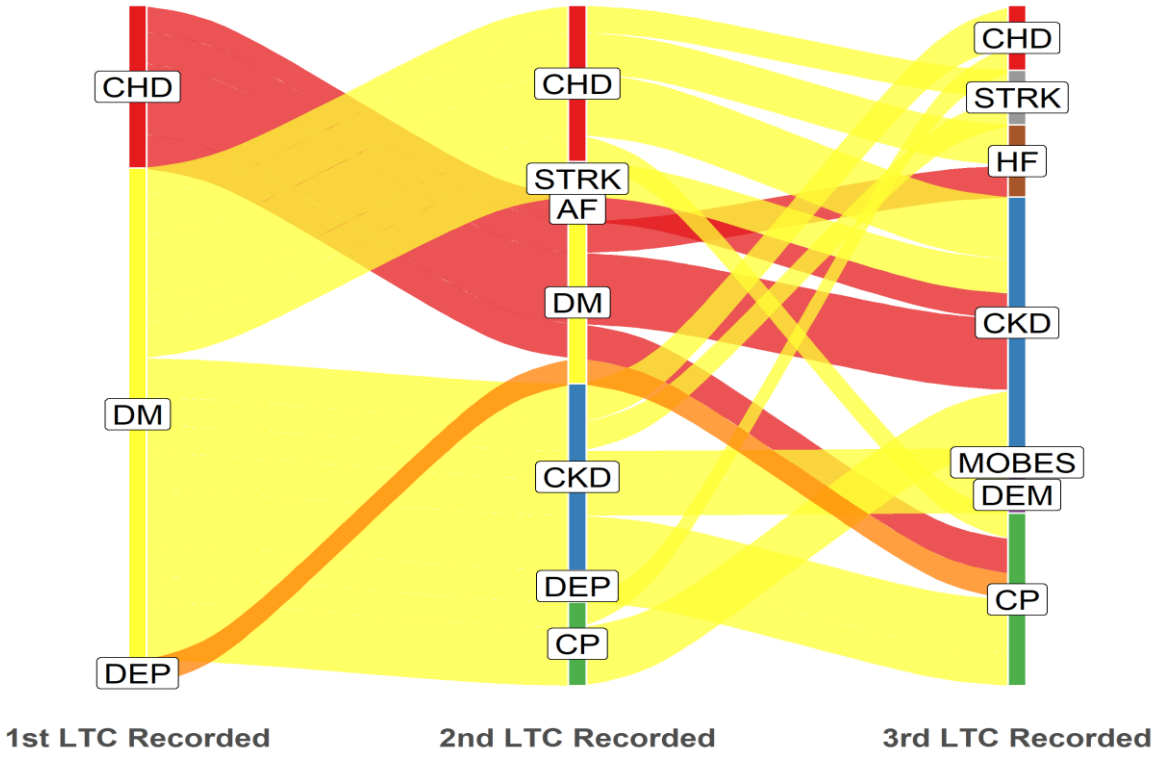
.











Supplementary File

Figure 1: Area under Receiver Operating Characteristic (ROC) curve = 0.84; based on regression model presented in Table 5.

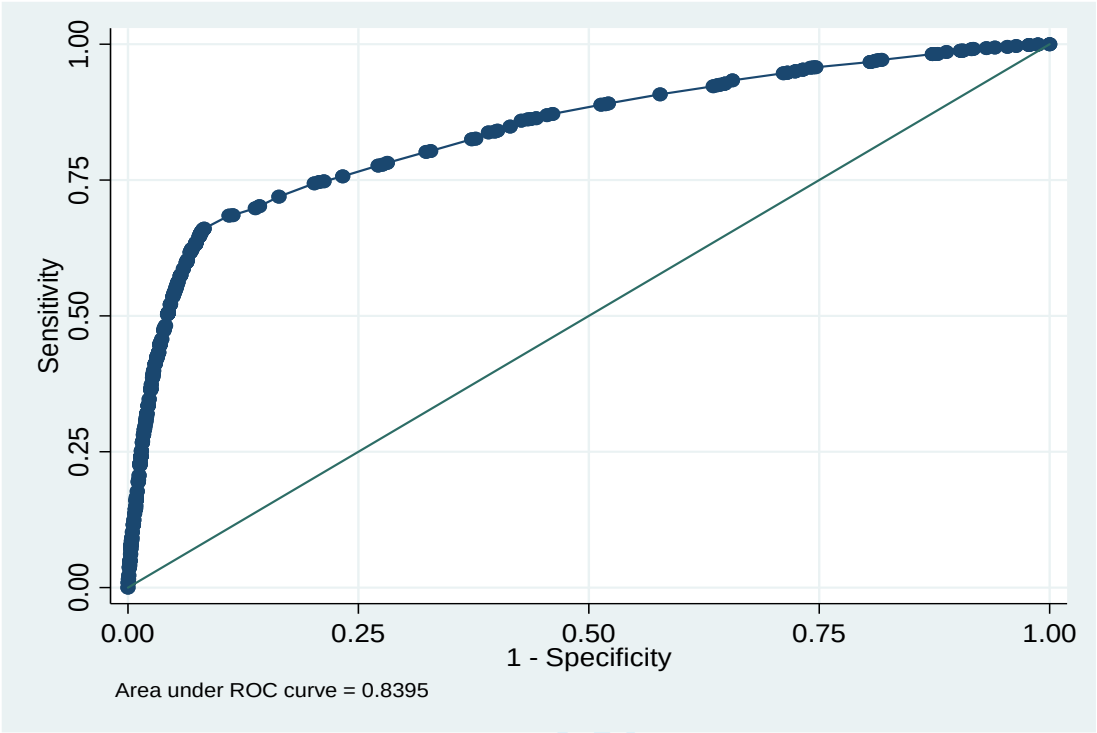
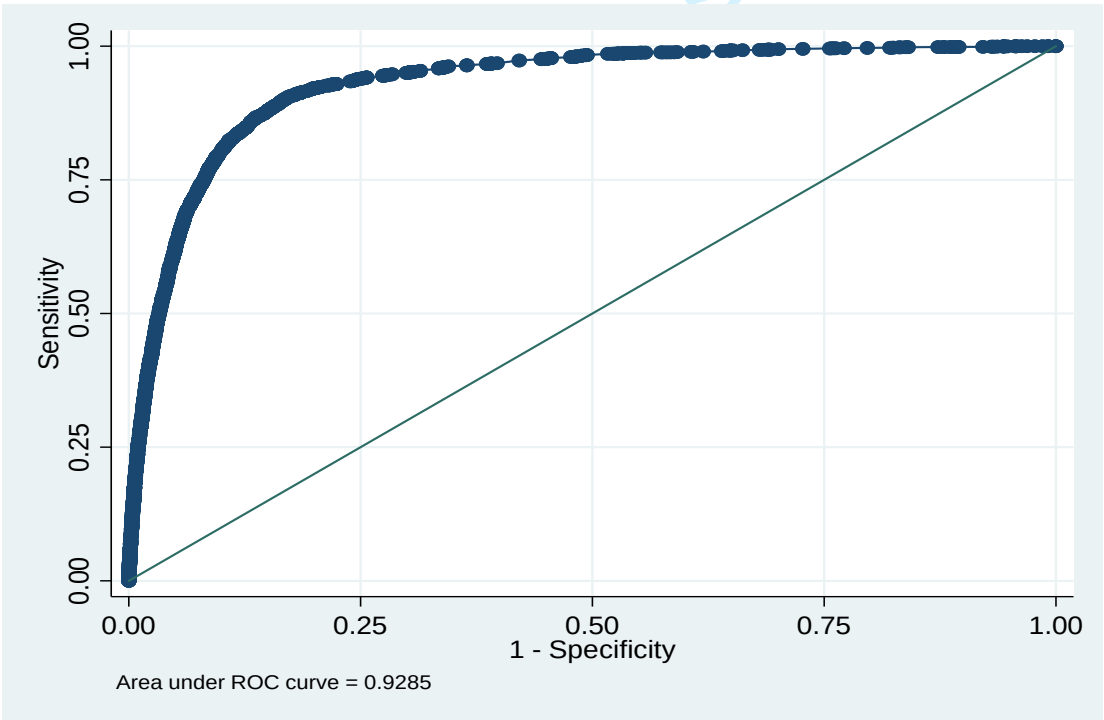
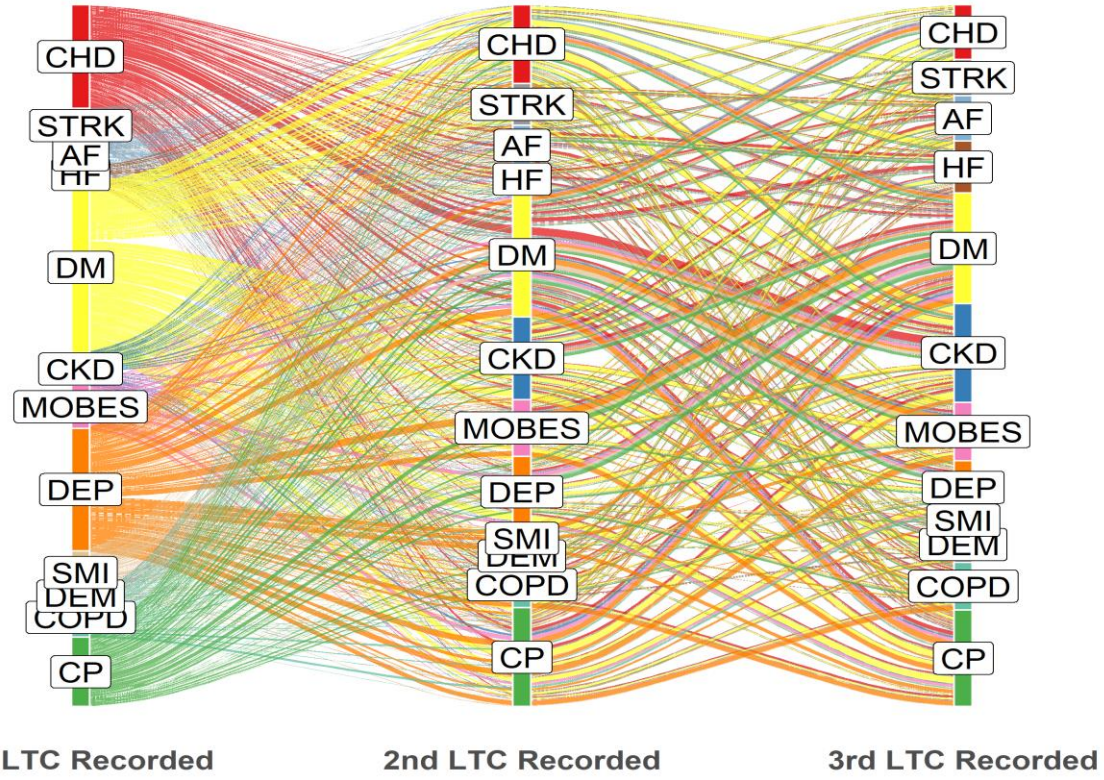


Figure 2: Area under Receiver Operating Characteristic (ROC) curve = 0.93; based on regression model presented in Table 5 with the addition of three risk factors: Hypertension, Smoking (ever), Obesity (moderate).



Figures 3-7 display un-edited alluvial plots for all available data. The figures in the main paper display alluvial plots edited to show dominant pathways.

Figure 3: Acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 5196); data excluded if simultaneous onset dates.



Long Term Condition label abbreviations (n = 12): CHD = Coronary Heart Disease; STRK = Stroke; AF = Atrial Fibrillation; HF = Heart Failure; DM = Diabetes; CKD = Chronic Kidney Disease; MOBES = Morbid Obesity; Dep = Depression; SMI = Serious Mental Illness; DEM = Dementia; COPD = Chronic Obstructive Pulmonary Disease; CP = Chronic Pain.

Figure 4: Most deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1394).

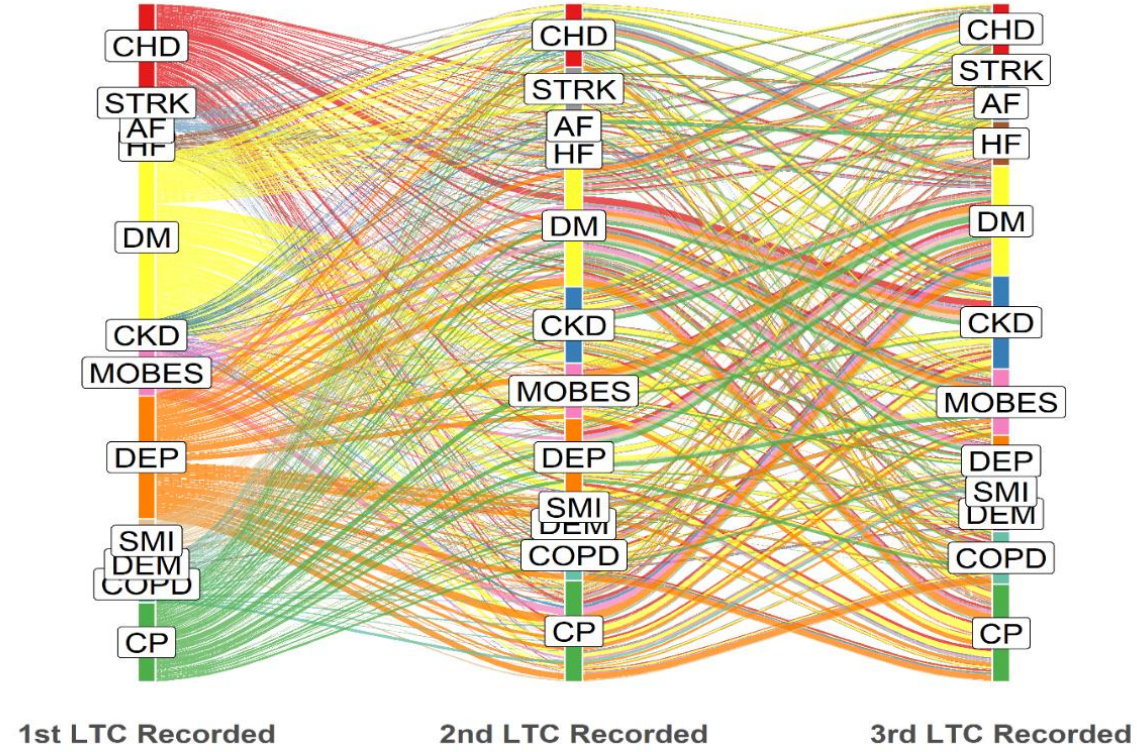


Figure 5: Least deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 763).

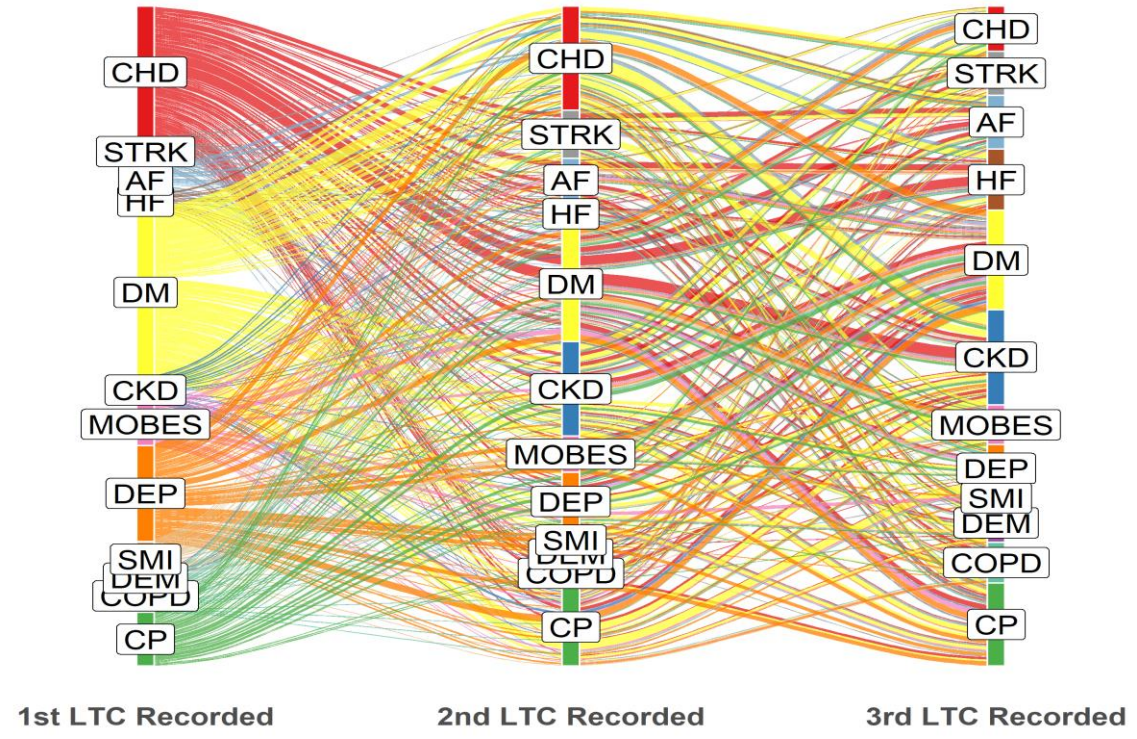


Figure 6: 'White' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 2788).

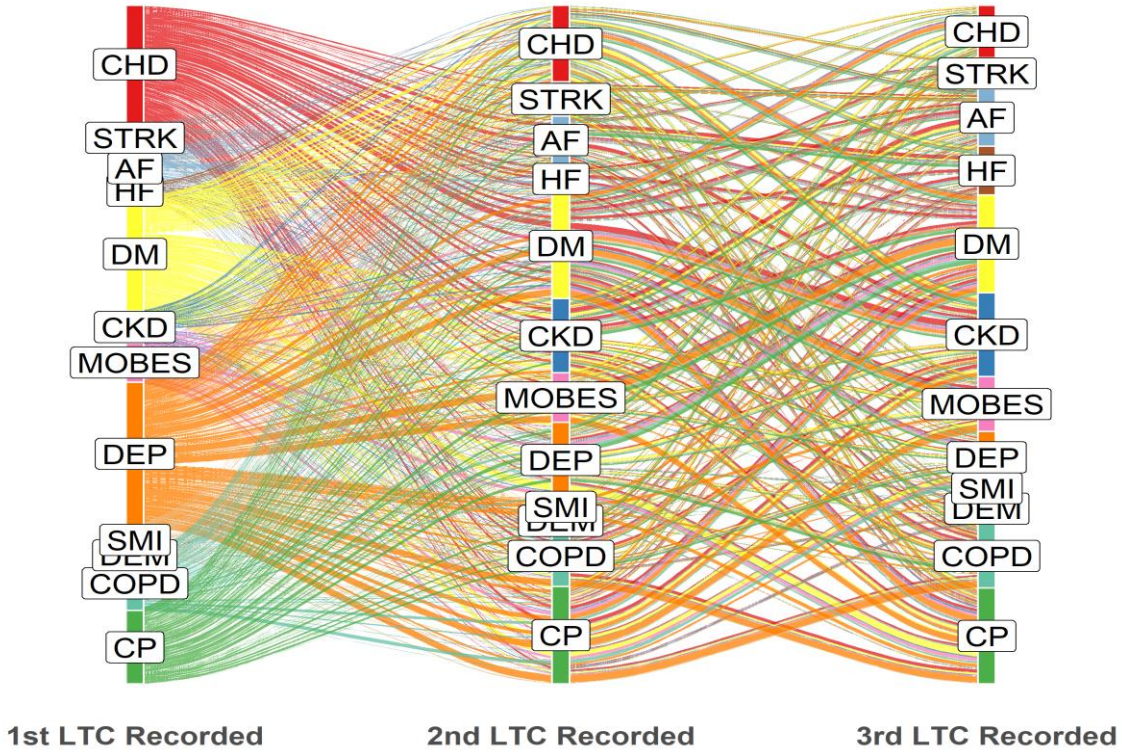
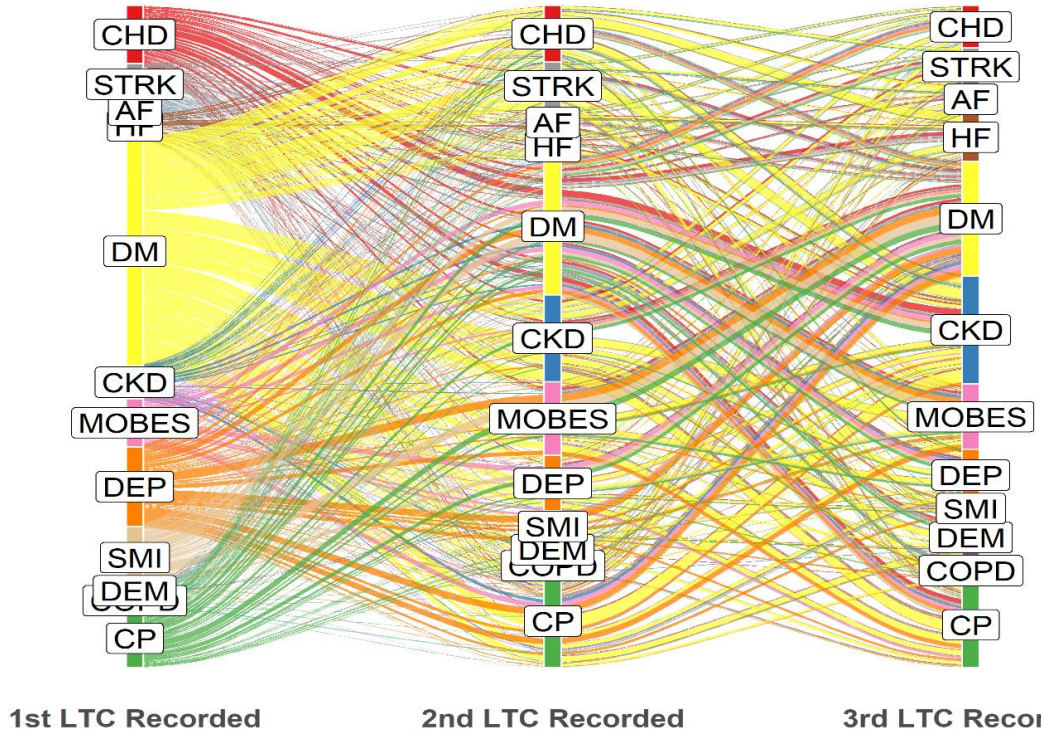


Figure 7: 'Black' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1448).



STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	"Cross-sectional analysis and longitudinal study"
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1	Summarised in Abstract: Design, Setting, Participants, Main outcome measures, Results
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4	The association between multimorbidity and social determinants and risk factors has not previously been determined in a deprived multi-ethnic community; identifying both the determinants of multimorbidity and the acquisition sequence may suggest interventions to prevent multimorbidity of slow its progression. For the purposes of this study, a locally derived definition of 'multimorbidity' has been used based on predicted high healthcare and social care demand. In contrast, most previously reported studies of multimorbidity have more

				inclusive definitions of multimorbidity.
Objectives	3	State specific objectives, including any prespecified hypotheses	1	"To study the social determinants and risk factors for multimorbidity and the acquisition sequence of multimorbidity"
Methods				
Study design	4	Present key elements of study design early in the paper	5	Presented under 'Data Variables' and 'Data Analysis'.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3	Presented under 'Study Setting', 'Study Design', 'Study population'.
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	4-5	Cross-sectional study: all patients registered at general practices in one south London borough (Lambeth) with the exception of those with an 'informed dissent' code in their case-notes. Longitudinal component: the same cohort of patients studied from onset of first Long Term Condition (LTC) to acquisition of 3 or more LTCs.
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case		

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5	Summarised under the heading, 'Data Variables': "….routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Data were extracted from the EHR into a secure data warehouse …..and contained information on patient demographic characteristics, LTCs, clinical values and medication."
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5	All data was derived from the Electronic Health Record of patients registered at GP practices in the sample area.
Bias	9	Describe any efforts to address potential sources of bias	5-6	Patient records were all included in the analysis, reducing sampling bias. However, 4.0% of patients had an 'informed dissent' code in their case-notes, prohibiting any access to data. We were therefore unable to determine if the omission of these patients may have introduced bias
Study size	10	Explain how the study size was arrived at	5-6	As above – the sample of 332,353 patients represented all patients registered at GP practices in the study area, with the exception of those with

	informed dissent codes. All Odds Ratios were presented with 95% CI's so that the effect of small numbers on the CI could be seen

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5/6	Quantitative variables analysed according to description in section headed, 'Data Analysis'.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	5/6	Regression modelling – see 'Data Analysis'
		(*b*) Describe any methods used to examine subgroups and interactions	6	Sub-groups of the key predictor variables were stratified (into age bands, deprivation quintiles, major ethnic groups)
		(*c*) Explain how missing data were addressed	5/6	See section on missing demographic codes, missing clinical codes, 'informed dissent codes'.
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a	The study was not a sample
		(*e*) Describe any sensitivity analyses	8	See heading 'Sensitivity analyses'
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6/7	Summarised under 'Multimorbidity cohort characteristics'.
		(b) Give reasons for non-participation at each stage	n/a	n/a (exclusion criteria stated above)
		(c) Consider use of a flow diagram	n/a	Exclusion criteria summarised on pg4-6 without use of a Flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Tables 1,2	Table 1 summarises, 'Frequencies of Long Term Conditions included in the multimorbidity cohort'; Table 2 summarises, 'Demographic characteristics of multimorbidity cohort'.

		(b) Indicate number of participants with missing data for each variable of interest	5/6	Missing data numbers summarised under, 'Multimorbidity cohort characteristics'
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a	
		Cross-sectional study—Report numbers of outcome events or summary measures	13	Summary measures in Table 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Tables 1-5	Univariable analysis presented in Tables 1, 2. Multivariable analysis presented in Tables 3,4,5. Table 3 includes different confounder variables to Tables 4,5. All confounder variables are presented in the Tables.
		(*b*) Report category boundaries when continuous variables were categorized	Tables 2-5	Continuous variables were categorised: 'age' into 3 age bands; social deprivation into quintiles
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a	The study was about the high risk of multimorbidity in specified groups, comparing the power of risk factors.

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Table 5; Supplementary file 1	Sensitivity analyses summarised: logistic regression adjusted for clustering; Receiver Operating Characteristic curve
Discussion				
Key results	18	Summarise key results with reference to study objectives	13/14	Key finding summarised at opening of Discussion. The demographic and risk factors for multimorbidity are defined. The acquisition sequence of multimorbidity is described.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14/15	Discussed under heading, 'Strengths and Limitations'
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-17	Cautious interpretation summarised with reference to the literature. Cautious conclusion reached about possible interventions to prevent or delay multimorbidity onset.
Generalisability	21	Discuss the generalisability (external validity) of the study results	14/15	Summarised under, 'Strengths and Limitations'. Findings were derived from one deprived, multi-ethnic community and may not generalise to less deprived, less diverse populations.
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18/19	The study was funded by the Guy's and St Thomas' Charity

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

The journey to multimorbidity: a longitudinal analysis exploring cardiovascular risk factors and socio-demographic determinants in an urban setting.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031649.R2
Article Type:	Original research
Date Submitted by the Author:	01-Nov-2019
Complete List of Authors:	Ashworth, Mark; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Durbaba, Stevo; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Whitney, David; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Crompton, James; London Borough of Lambeth, Public Health, Adults and Health Wright, Michael; King's College London Faculty of Life Sciences and Medicine, School of Population Health & Environmental Sciences Dodhia, Hiten; London Borough of Lambeth, Public Health
Primary Subject Heading:	General practice / Family practice
Secondary Subject Heading:	Public health
Keywords:	PRIMARY CARE, Multimorbidity, Demographic factors, Risk factors

TITLE:

The journey to multimorbidity: a longitudinal analysis exploring cardiovascular risk factors and socio-demographic determinants in an urban setting

AUTHORS:

Mark Ashworth[1], Stevo Durbaba[2], David Whitney[3], James Crompton[4], Michael Wright[5] Hiten Dodhia[6].

1. Reader in Primary Care, School of Population Health & Environmental Sciences, King's College London, London, UK
2. e-Resources Developer, School of Population Health & Environmental Sciences, King's College London, London, UK
3. Data Analyst, School of Population Health & Environmental Sciences, King's College London, London, UK
4. Head of Health Intelligence, Public Health, Adults and Health, London Borough of Lambeth, London, UK
5. Project Manager, School of Population Health & Environmental Sciences, King's College London, London, UK
6. Consultant in Public Health, London Borough of Lambeth, London, UK

Corresponding author email: mark.ashworth@kcl.ac.uk

WORD COUNT: 3657

ABSTRACT

Objective. To study the social determinants and cardiovascular risk factors for multimorbidity and the acquisition sequence of multimorbidity.

Design. Longitudinal study based on anonymised primary care data.

Setting. General practices in an urban multi-ethnic borough in London, UK.

Participants. 332,353 patients aged ≥18 years.

Main outcome measures. Clinical and socio-demographic characteristics of patients with multimorbidity, defined as ≥3 of 12 Long Term Conditions (LTCs) selected according to high predicted healthcare utilisation. Multilevel logistic regression was used to model the social determinants and cardiovascular risk factors. Alluvial plots were constructed to illustrate multimorbidity acquisition sequences according to age, ethnicity and social deprivation.

Results. 5597 (1.7%) patients had ≥3 selected LTCs, the 'multimorbidity cohort'. The commonest LTCs were diabetes (63.0%) and chronic pain (42.8%). Social deprivation and ethnicity were independent determinants of multimorbidity: most compared to least deprived quintile, Adjusted Odds Ratio (AOR) 1.56 (95%CIs, 1.41, 1.72); South Asian compared to white ethnicity, AOR 1.44 (95%CIs, 1.29, 1.61); black compared to white ethnicity, AOR 0.86 (95%CIs, 0.80, 0.92). The included cardiovascular risk factors were relatively strong determinants of multimorbidity: hypertension, AOR 5.05 (95%CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95%CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45). The most common initial onset conditions were diabetes and depression; diabetes particularly in older and black ethnic groups; depression particularly in younger, more deprived and white ethnicity groups. Chronic pain was less common as an initial condition.

Conclusions Our findings confirm the importance of age, social deprivation and ethnicity as determinants of multimorbidity. Smoking, obesity and hypertension as cardiovascular risk factors were stronger determinants of multimorbidity than deprivation or ethnicity. The acquisition sequence of multimorbidity is patterned by socio-demographic determinants. Understanding onset conditions of multimorbidity and cardiovascular cardiovascular risk factors may lead to the development of interventions to slow the progression of multimorbidity.

Strengths and limitations

- This study uses a definition of multimorbidity based on Long Term Conditions (LTCs) with high predicted healthcare utilisation rates
- Multimorbidity is studied in a deprived, multi-ethnic community
- Longitudinal data is used to identify the acquisition sequence of multimorbidity and how this is influenced by socio-demographic determinants
- Difficulties gaining access to anonymised primary care data limited the sample size and may have contributed to selection bias
- Coding validation and completeness restricted the analysis of available primary care data

BACKGROUND

Healthcare utilisation is increasingly driven by multimorbidity (1). Each long-term condition (LTC) included within a definition of multimorbidity is likely to generate multiple primary care consultations with general practitioners (GPs), practice nurses and other health care professionals, and may result in Accident and Emergency (A&E) attendances, referral to out-patient appointments and hospital admissions.

Estimates of healthcare utilisation attributable to multimorbidity vary according to the LTCs included within a definition of multimorbidity. There is no standard definition of multimorbidity. One systematic review noted that the number of included LTCs ranged from five to 185 with estimates of population prevalence ranging from 13.1% to 71.8% depending on the number of included conditions (2). In one recent UK study in which multimorbidity was defined as two or more of a selection of 36 LTCs, 27.2% of the population had multimorbidity, accounting for 52.9% of GP consultations and generating a median of nine annual GP consultations (3). LTCs also contribute to the development of frailty, itself a driver of healthcare utilisation (4).

In response to high demand for health and social care, many healthcare providers and commissioners have sought to identify those patients with greatest needs through a process termed

'risk stratification'. Several electronic tools have been developed to offer population based risk stratification (5). An alternative approach is the use of expert panels to define high-demand patient groups (6). Having identified the cohort of patients with the greatest requirement for health and social care services, the purpose of this process is to guide resource allocation on a needs basis, often with the implicit assumption that additional investment in a primary care setting may reduce demand for more expensive secondary care services. Although funding and healthcare need should align, there is little evidence that investment in additional community resources for those most at risk of hospital attendance results in overall reductions in secondary care utilisation (7).

Following a consultation exercise and report from an expert panel, two inner London boroughs and their commissioning groups made the decision to define multimorbidity based on predicted high healthcare and social care demand. Multimorbidity in this context consisted of three or more LTCs considered most likely to result in functional impairment and high service demand. This 'high service demand' definition was used to identify a cohort to receive a package of integrated care, termed 'care coordination'. A focus on this narrowly defined category of multimorbidity would inevitably mean that the proportion of patients defined as 'multimorbid' would be lower than reported in studies based on broader definitions of multimorbidity (2,3).

The aim was to study the characteristics of this multimorbidity cohort. The main objectives were to define both the socio-demographic determinants and cardiovascular risk factors associated with multimorbidity acquisition; also to determine the acquisition sequence of multimorbidity and the influence of demographic factors on this sequence.

METHODS

Study setting

Our study was set in Lambeth, one of the two inner London UK boroughs adopting the 'care coordination' definition of multimorbidity (8). The population sample consisted of all patients

registered at all general practices (n = 44) in Lambeth, with the exception of patients who had opted out of anonymised data sharing for research purposes.

Study design

We conducted a longitudinal analysis based on anonymised coded primary care data extracted from electronic health records (EHR) held in primary care.

Study population

We included data on all patients aged 18 years and over registered with a general practice. For the population with multimorbidity, we included all those with LTCs recorded in the EHR and included in the 'care coordination' definition of multimorbidity: Atrial Fibrillation (AF), Chronic Obstructive Pulmonary Disease (COPD), Chronic Pain (CP), Chronic Kidney Disease (CKD), Coronary Heart Disease (CHD), Diabetes (DM), Dementia, Depression, Heart Failure (HF), Serious Mental Illness (SMI), Stroke, Morbid Obesity. The definition and specified codes for each condition was that used by the Quality and Outcomes Framework (QOF), based on 'QOF38' definitions (9). Two of the conditions selected for inclusion within the definition of multimorbidity were not included within the QOF: chronic pain, defined on the basis of two or more repeat prescriptions for opioid analgesics (British National Formulary, chapter 4.7.2) or neuropathic pain medication (British National Formulary, chapter 4.7.3) (10); morbid obesity defined as a Body Mass Index ≥40 kg/m^2. For each LTC, the date of onset was obtained from the EHR and used in the longitudinal analysis.

Demographic data consisted of gender, age in years (on the date of data extraction) and self-ascribed ethnicity obtained from the EHR. Social deprivation data derived from residency data was based on the Index of Multiple Deprivation 2010 classification at Lower Super Output Area (LSOA), stratified into locally based quintiles (11). Local deprivation quintiles were used in place of national quintiles since mean deprivation levels are high in Lambeth, the 22nd most deprived local authority (out of 326) in England (12).

Cardiovascular risk factors included in the analysis were: hypertension (defined as patients on the QOF Hypertension register); moderate obesity (defined as a Body Mass Index of 30.0-39.9; note that patients with BMI ≥40 were included as one of the LTCs within the definition of multimorbidity and therefore not included as a cardiovascular risk factor); smoking (patients with any record of being a smoker).

Data variables

The data consisted of 'real world', routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Routinely collected electronic data was available from all included practices from 2004. Data were extracted from the EHR into a secure data warehouse and contained information on patient demographic characteristics, LTCs, clinical values and medication. The data used in this study were extracted in May 2018 and related to all patients registered at each of the included practices on that date.

Data analysis

We analysed socio-demographic (age, gender, ethnicity), social (area level deprivation) and cardiovascular risk factor (hypertension, moderate obesity, smoking status) data for the multimorbidity cohort and general population using univariable statistical methods applied at patient-level. Socio-demographic and cardiovascular risk factor determinants of multimorbidity were analysed using multilevel logistic regression models to model practice level variation. We also conducted a sensitivity analysis using logistic regression models adjusted for clustering at the practice level. The sensitivity analysis allowed pseudo-r^2 values and Receiver Operating Characteristic (ROC) curves to be derived. Analysis was conducted using the statistical software package STATA IC 15 (13).

The acquisition sequence for patients in the multimorbidity cohort was established by searching the EHR for date of onset of each LTC. For this analysis, we established the order of acquisition and

tabulated the frequency of first, second and third LTCs. Patients with identical dates of onset recorded for two or more LTCs were excluded from this analysis. We displayed findings using alluvial plots, an infographic allowing representation of multiple pathways. These were constructed using the software R, and the packages 'ggplot2' and 'ggalluvial' (14).

Patient and public involvement

The borough based and statutory organisation, Lambeth HealthWatch, represented the interests of patients and public in this work; they contributed to the original protocol design and shared in dissemination of the findings.

RESULTS

Multimorbidity cohort characteristics

The final study population consisted of 332,353 patients aged ≥18 years. Data from 13,369 (4.0%) patients had been excluded because a data sharing opt-out code was recorded in their EHR. Patients were included in the final sample even though some socio-demographic data were missing: 3289 (0.99%) patients could not be linked to a LSOA and therefore had missing IMD-2015 score data; ≤10 patients had missing coded gender data; ≤10 patients had missing coded age data. Patients with any category of missing data (n = 3301) were excluded from the multivariable analysis.

In all, 5597 (1.7%) patients had a record of three or more of the selected LTCs, the 'multimorbidity cohort'. Most (n = 3542) of this cohort had three LTCs (63.3%); 1333 (23.8%) had four LTCs; 492 (8.8%) had 5 LTCs; the remaining 230 (4.1%) had more than 5 LTCs. Of the remaining population, 45,241 (13.6%) had one LTC and 10,992 (3.3%) had two LTCs.

A summary of LTC frequencies within the multimorbidity cohort is displayed in Table 1. The most common LTCs within this cohort were: DM (63.0%) and chronic pain (42.8%). In contrast, the most common of the included LTCs in the adult general population were: depression (8.4%), DM (5.4%) and morbid obesity (3.2%).

The socio-demographic characteristics of the multimorbidity cohort are displayed in Table 2. Within the cohort, 33.9% were aged under 65 years (compared with 91.7% in the sample population); 27.7% were of a 'Black' ethnicity (18.0% in the sample population) and 46.0% were born in the UK (45.2% in the sample population). The mean age for the multimorbid cohort was 69.9 years (compared with a mean of 41.6 years in the sample population). The mean age for multimorbid patients in least and most deprived quintiles was 73.0 (standard deviation (SD), 13.5) and 69.3 (SD, 12.9) years, respectively; and for the white, black and south Asian populations was 71.2 (SD, 13.3), 69.4 (SD, 14.3) and 71.9 (SD, 11.8) years, respectively.

Multimorbidity cohort socio-demographic determinants

Socio-demographic determinants of the multimorbidity cohort were dominated by the effect of older age groups, but also included social deprivation and ethnicity (Supplementary file: Table 1). Based on the adjusted odds ratio (AOR) derived from the multilevel regression model, the strongest determinant for multimorbidity was related to age. After adjustment for age, both social deprivation (the more deprived quintiles) and ethnicity (Black and south Asian ethnicities) remained significant determinants, albeit with much smaller odds ratios. The AOR for black and South Asian compared with white ethnicity was 1.15 (95% CIs, 1.07, 1.23) and 1.19 (95% CIs, 1.07, 1.33), respectively; for most compared to least deprived quintile, the AOR was 1.83 (95% CIs, 1.66, 2.02).

Multimorbidity cohort cardiovascular risk factor determinants

Addition of the three cardiovascular risk factors included in the study attenuated the odds ratios for multimorbidity related to age (Table 3). Social deprivation remained a determinant of multimorbidity: for most compared to least deprived quintile, the AOR was 1.56 (95% CIs, 1.41, 1.72). South Asian ethnicity remained a significant determinant of multimorbidity: AOR 1.44 (95% CIs, 1.29, 1.61) but black ethnicity was no longer a positive determinant: AOR 0.86 (95% CIs, 0.80, 0.92).

The three cardiovascular risk factors were significant determinants of multimorbidity: hypertension, AOR 5.05 (95% CIs, 4.69, 5.44); moderate obesity, AOR 3.41 (95% CIs, 3.21, 3.63); smoking, AOR 2.30 (95%CIs, 2.16, 2.45).

Sensitivity analyses

Re-analysis of the determinants of multimorbidity using regression modelling adjusted for clustering at practice level resulted in similar adjusted odds ratios to those obtained in the primary analyses(Supplementary file: Table 2). We explored goodness-of-fit through derived pseudo-r^2 values for demographic determinants: pseudo-r^2 = 0.22, and for cardiovascular risk factor adjusted determinants (hypertension, moderate obesity, smoking): pseudo-r^2 = 0.32. The areas under the ROC curve for each model were 0.84 and 0.93, respectively (Supplementary file: Figures 1,2).

Multimorbidity acquisition sequence

Of the 5597 patients in the multimorbidity cohort, 5196 had three distinct dates of onset for each of their three or more component LTCs. The remaining 401 (7.2%) patients had identical dates of onset recorded for two or three of their first three LTCs and therefore could not be classified into a sequence. Alluvial plots were constructed displaying the acquisition sequence for LTCs and edited to display dominant flows of patients in each category. Figure 1 displays the three most common starting conditions and subsequent most commonly acquired second and third LTCs. Unedited alluvial plots alluvial plots displaying all patient flows are shown in the Supplementary file, Figure 3.

The alluvial plots illustrate that diabetes and depression were the most common starting conditions for patients with multimorbidity; diabetes was also relatively common as the second or third acquired LTC whereas depression was predominantly a first-onset LTC (Figure 1).

In the most deprived quintile, diabetes and depression were the most common starting conditions whereas in the least deprived quintile, diabetes and CHD were more common as starting conditions

than depression (Figures 2, 3). Unedited alluvial plots comparing most and least deprived quintiles are shown in the Supplementary file, Figures 4, 5.

Multimorbidity in the white ethnic group was dominated by depression as the starting condition, whereas in the black ethnic group diabetes was the most common starting condition, with depression and SMI also relatively common (Figures 4, 5). Relatively small numbers resulted in poor definition of the alluvial plot in the South Asian group and this figure is not presented. Unedited alluvial plots comparing black and white ethnic groups are shown in the Supplementary file Figures 6, 7.

Multimorbidity in the under 65 year old cohort was dominated by depression as the starting condition, with SMI also relatively common; in the ≥65 year old cohort, diabetes and CHD were the most common starting conditions (Figures 6, 7).

Chronic pain appeared to be more common as a second or third acquired LTC but less common as a first LTC. This sequence was apparent in the overall picture, in the pattern displayed by least and most deprived quintiles, for black and white ethnicities and for younger and older age cohorts.

Morbid obesity was among the more common starting conditions in the most deprived cohort and younger age cohort. However, in other socio-demographic samples, morbid obesity was more common as a second and third acquired LTC.

DISCUSSION

We report on multimorbidity using a definition originating from a health service commissioning perspective, consisting of three or more of 12 LTCs selected because of likely high impact upon health service and social care utilisation. In total, 1.7% of the adult population in our study sample had multimorbidity according to these narrowly defined criteria. Diabetes and chronic pain were the most prevalent LTCs within this cohort. Independent of age, both ethnicity and social deprivation were significant determinants of multimorbidity. However, the cardiovascular risk factors of

hypertension, obesity and smoking were more strongly associated with multimorbidity than social deprivation or ethnicity.

The acquisition sequence of multimorbidity differed substantially according to age, ethnicity and social deprivation. Diabetes and depression were the most common starting conditions overall. Diabetes as a starting condition was notably more common in the older and black ethnic group. Depression as a starting condition was notably more common in patients who were younger, more deprived and in the white ethnic group. Differences in acquisition sequence between most and least deprived areas and white and black ethnicities (Figures 2 - 5) are unlikely to have been strongly influenced by age differences since mean age was similar for each of these cohorts.

Strengths and Limitations

Data access was a limitation to this analysis. We were only able to obtain data from one of the two boroughs adopting this approach to multimorbidity, the other lacked a data extraction system preventing us from analysing large datasets of patient-level data. Had we gained access to the data, this would have approximately doubled our sample size and enabled further analysis of multimorbidity in deprived, multi-ethnic populations. This difficulty in accessing anonymised data hampers the analysis of patient-level data in many areas of the UK (15). Data coding constrains the analysis of primary care data and we were only able to study the association of multimorbidity with a limited range of cardiovascular risk factors while other known risk factors such as exercise and diet could not be captured. We aimed to display the acquisition sequence of LTCs using alluvial plots. Although these plots clearly display the sequence in which patients develop LTCs, they do not readily display time data and thus fail to distinguish between rapidly and slowly progressing multimorbidity. A time-to-event analysis is required for identifying those patients who progress rapidly from first LTC into multimorbidity, which is the subject of further study. Similarly, the acquisition sequence could not be determined for a small minority of patients with identical LTC date-of-onset recording by GPs whereas in reality, it is unlikely that the LTC onset dates were simultaneous. Furthermore, as with all

studies based on primary care data, there may be coding anomalies which introduce bias into the estimates of LTC prevalence. QOF coding criteria were used for 10 of the included LTCs, standardising the definition. However, the prevalence of conditions such as depression may be underestimated using QOF criteria (16). For the two LTCs not included in the QOF, the definition is dependent on GP coding. Thus 'morbid obesity' was only included in our study if there was a BMI recording which may have resulted in an under-estimate of prevalence. 'Chronic pain' was defined based on medication consumption whereas many patients with chronic pain may have sought alternatives to analgesic medication resulting in an underestimate of prevalence; conversely, our inclusion criteria of 'two or more prescriptions over the preceding year' may have resulted in an overestimate of prevalence, with some patients recovering from chronic pain during the course of the year. In common with other observational studies, significant associations between multimorbidity and socio-demographic or cardiovascular risk factor determinants may imply, but cannot prove, causality. Whilst interventional studies are required to obtain stronger evidence of causality, causal inference may be derived by time series analyses and further study of potential confounding and residual variance. Finally, the richness of locally based data covering a whole borough with unique socio-demographic characteristics has to be offset against possible loss of generalisability to other areas with very different social deprivation and ethnicity characteristics.

Comparison with the literature

Our cross-sectional data, although conducted in a deprived, multi-ethnic population, are similar to the findings of others reporting on increased multimorbidity prevalence associated with age, social deprivation and ethnic minority status (17). Comparison with other multimorbidity studies is difficult because of the highly restricted definition of multimorbidity used in the current study. Nevertheless, other studies have reported the high prevalence of both diabetes and depression in multimorbidity cohorts (3, 17). A higher prevalence of mental health and physical health LTC combinations has been noted in deprived areas, although in our own study we did not conduct an analysis to identify

specific LTC combinations (18). Certain conditions have been found to be more prevalent in deprived communities, also contributing to higher prevalence of multimorbidity in these areas, such as depression and addiction issues in younger deprived populations (19), or more generally, depression, drugs, anxiety, dyspepsia, chronic pain, CHD, DM (20). These findings are aligned to our own reporting of high proportions of multimorbid patients in deprived areas with depression and diabetes. The known influence of population age profiles on the demography of multimorbidity (19) is illustrated by our finding of markedly differing alluvial plot profiles describing multimorbidity acquisition in a younger cohort dominated by mental health related conditions (Figure 6) and an older cohort dominated by CHD and DM (Figure 7). The inner city population in our study is characterised by a much younger overall age profile compared with the national population: 8% of the population of Lambeth is aged 65 years or more compared to a mean of 18% in England (21). Added to this, the most socially deprived and black ethnic minority groups in our sample were somewhat younger. Both factors are likely to have reduced the overall study prevalence of conditions associated with ageing such as diabetes and CHD. Many multimorbidity studies do not include morbid obesity within their definition (3, 17); we found morbid obesity was particularly associated with social deprivation.

Several publications have reported on combinations and clusters of LTCs but few longitudinal analyses have been reported (22). One study from Australia reported the order of appearance for eight LTCs, reporting in detail for asthma and mood related disorders, the two LTCs most strongly associated with the risk of developing a second LTC. For those with baseline asthma, there was a higher subsequent risk of developing COPD and hypercholesterolaemia; for those with baseline mood disorders, their risk of subsequent asthma, diabetes and other mental disorders was increased (23). In a study of cardiometabolic conditions in Australia, nearly one-quarter of women initially diagnosed with stroke subsequently progressed to other conditions which was a much larger proportion progressing to other conditions than in those initially diagnosed with diabetes (9.9%) or heart disease (11.4%) (24). A further US study explored the acquisition sequence of 20 LTCs

describing dyads and triads of conditions and reporting, for example, that the most common triad sequence in 20-39 year olds was depression, asthma and substance misuse whereas in 50-59 year olds it was hyperlipidaemia, hypertension and diabetes (25). They concluded that combinations of LTCs vary extensively by age and sex. Our own findings confirm variation by age and sex, with ethnicity adding to the pattern of variation. Some authors have suggested that the study of acquisition sequence may imply potential interventions to prevent, minimise or delay progression toward multimorbidity (23). Our findings that three cardiovascular risk factors are more strongly associated with multimorbidity than deprivation and ethnicity, suggest that interventions to reduce the impact of these cardiovascular risk factors may contribute to a reduction in the prevalence of multimorbidity. The control of hypertension, smoking and obesity are often perceived in terms of primary cardiovascular disease prevention or of secondary prevention of single LTCs but may also be conceptualised in terms of multimorbidity prevention. However, a focus on cardiovascular risk factors should not detract from 'the causes of the causes', since social conditions themselves generate causal pathways leading from socio-economic determinants to risk behaviours (26) and interventions which address health behaviours while failing to engage in social determinants may paradoxically result in increased health inequalities (27).

CONCLUSION

We have confirmed the role of age, social deprivation and ethnicity as determinants of multimorbidity in an inner city multi-ethnic population and have extended previous findings demonstrating the way in which the acquisition sequence of multimorbidity is patterned by these determinants. Three cardiovascular risk factors, hypertension, obesity and smoking were stronger determinants of multimorbidity than either deprivation or ethnicity. The strength of these cardiovascular risk factors as determinants suggests interventions which may be effective in reducing the prevalence, delaying the onset or slowing the progression of multimorbidity.

Contributors: MA, HD and MW contributed to the idea and design of the study. SD, DW and JC led on data extraction and preparation. Statistical analysis was conducted by MA and HD. MA produced the first draft of the paper; all co-authors contributed and approved the final draft. MA is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This work was supported by the Guy's and St Thomas' Charity with a grant awarded November 2017. The views expressed are those of the authors and not necessarily those of the funders. The findings presented are independent from the funders who have had no role in data collection, data analysis or writing of the paper.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval

All data were extracted under the terms of a signed Data Sharing Agreement (DSA) with each practice and with project-specific approval following submission of a Data Privacy Impact Assessment (DPIA), approved by Lambeth Clinical Commissioning Group (CCG) on 2/11/17. Information Governance approval required 'low number suppression', ensuring that data could not be displayed if the patient number was 10 or less in any given category; in these circumstances, data reporting would state: '≤10 patients'. Separate Ethical Committee approval was not required (Health Research Authority, personal correspondence, 29/9/17) since all data were fully anonymised for the purposes of research access, and all Patient Identifiable Data (PID) had been removed.

Data sharing: No additional data available.

REFERENCES:

1. Palladino R, Tayu Lee J, Ashworth M, *et al*. Associations between multimorbidity, healthcare utilisation and health status: evidence from 16 European countries. *Age Ageing* 2016;45:431-5. https://dx.doi.org/10.11606%2FS1518-8787.2018052000639
2. Fortin M, Stewart M, Poitras ME, *et al*. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142-51. https://doi.org/10.1370/afm.1337
3. Cassell A, Edwards D, Harshfield A, *et al*. The epidemiology of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2018;68:e245-e51. https://doi.org/10.3399/bjgp18X695465
4. Hanlon P, Nicholl BI, Jani BD, *et al*. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. *Lancet Public Health* 2018;3:e323-e32. https://doi.org/10.1016/S2468-2667(18)30091-4
5. Billings J, Blunt I, Steventon A, *et al*.. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2. http://dx.doi.org/10.1136/bmjopen-2012-001667
6. Huntley AL, Johnson R, Purdy S, *et al*. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med*. 2012;10:134-41. https://doi.org/10.1370/afm.1363
7. Roland M, Abel G. Reducing emergency admissions: are we on the right track? *Br Med J* 2012;345:e6017. https://doi.org/10.1136/bmj.e6017
8. From one to many. Exploring people's progression to multiple long-term conditions in an urban environment. Guy's and St Thomas' Charity. London, 2018. https://www.gsttcharity.org.uk/sites/default/files/GSTTC_MLTC_Report_2018.pdf.
9. Quality and Outcomes Framework (QOF) business rules v 38 2017-2018 October code release.NHS Digital. Leeds, 2018. https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof/quality-and-outcome-framework-qof-business-rules/quality-and-outcomes-framework-qof-business-rules-v-38-2017-2018-october-code-release.
10. Joint Formulary Committee: British National Formulary (online). London: BMJ Group and Pharmaceutical Press,http://www.medicinescomplete.com
11. English Indices of Deprivation 2015 - LSOA Level. Ministry of Housing, Communities and Local Government. London, 2018. https://data.gov.uk/dataset/8f601edb-6974-417e-9c9d-85832dd2bbf2/english-indices-of-deprivation-2015-lsoa-level
12. Lambeth, State of the borough, 2016. Lambeth Council. London, 2017. https://www.lambeth.gov.uk/sites/default/files/State%20of%20Borough%202016%20-%20v3.pdf
13. Statacorp. Stata Statistical Software: Release 14 2015. Available from: http://www.stata.com/
14. Alluvial Diagrams in ggplot2. Brunson, JC, 2018. https://cran.r-project.org/web/packages/ggalluvial/vignettes/ggalluvial.html
15. Harwich E, Lasko-Skinner R. Making NHS data work for everyone. Reform. London, 2018. https://reform.uk/sites/default/files/2018-12/Making%20NHS%20data%20work%20for%20everyone%20WEB_1.pdf
16. Kendrick T, Stuart B, Newell C, *et al*. Changes in rates of recorded depression in English primary care 2003-2013: Time trend analyses of effects of the economic recession, and the GP contract quality outcomes framework (QOF). *J Affect Disord* 2015;180:68-78. https://doi.org/10.1016/j.jad.2015.03.040
17. Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37-43. https://doi.org/10.1016/S0140-6736(12)60240-2

18. McLean G, Guthrie B, Mercer SW, *et al*. General practice funding underpins the persistence of the inverse care law: cross-sectional study in Scotland. *Br J Gen Pract* 2015;65:e799-805. https://doi.org/10.3399/bjgp15X687829
19. Leiser S, Deruaz-Luyet A, N'Goran AA, *et al*. Determinants associated with deprivation in multimorbid patients in primary care-A cross-sectional study in Switzerland. *PLoS One* 2017;12:e0181534. https://doi.org/10.1371/journal.pone.0181534
20. McLean G, Gunn J, Wyke S, *et al*. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. *Br J Gen Prac*t 2014;64:e440-7. https://doi.org/10.3399/bjgp14X680545
21. Lambeth. Demographic factsheet, May 2015. Lambeth & Southwark's Public Health Intelligence Team, 2015. https://www.lambeth.gov.uk/sites/default/files/ssh-lambeth-demography-2015.pdf
22. Multimorbidity: a priority for global health research. Academy of Medical Sciences. London, 2018. https://acmedsci.ac.uk/file-download/82222577
23. Ruel G, Levesque JF, Stocks N, *et al*. Understanding the evolution of multimorbidity: evidences from the North West Adelaide Health Longitudinal Study (NWAHS). *PLoS One* 2014;9:e96291. https://doi.org/10.1371/journal.pone.0096291
24. Xu X, Mishra GD, Dobson AJ, *et al*. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. *PLoS Med* 2018;15:e1002516. https://doi.org/10.1371/journal.pmed.1002516
25. St Sauver JL, Boyd CM, Grossardt BR, *et al*. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open* 2015;5:e006413. https://doi.org/10.1136/bmjopen-2014-006413
26. Ferraro KF, Schaffer MH, Wilkinson LR. Childhood Disadvantage and Health Problems in Middle and Later Life: Early Imprints on Physical Health? *Am Sociol Rev* 2016;81:107-133. https://doi.org/10.1177/0003122415619617
27. Link BG, Phelan J. Social Conditions As Fundamental Causes of Disease. *J Health Soc Behavior* 1995;(Extra Issue): 80-94. DOI: 10.2307/2626958

Table 1. Frequencies of Long Term Conditions included in the multimorbidity cohort (n = 5597 patients) compared with the remainder of the registered population and with the general population aged ≥18 years.

Long Term Condition	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756	Total Population values: frequency (valid %) N = 332,353
DM	3525 (63.0%)	14,405 (4.4%)	17,930 (5.4%)
Chronic Pain	2397 (42.8%)	5813 (1.8%)	8210 (2.5%)
CKD	2101 (37.5%)	3470 (1.1%)	5571 (1.7%)
CHD	2099 (37.5%)	2668 (0.8%)	4767 (1.4%)
Depression	2086 (37.3%)	25,877 (7.9%)	27,963 (8.4%)
Morbid Obesity	1653 (29.5%)	8883 (2.7%)	10,486 (3.2%)
AF	1254 (22.4%)	1493 (0.5%)	2747 (0.8%)
COPD	1247 (22.3%)	2387 (0.7%)	3634 (1.1%)
Heart Failure	1186 (21.2%)	544 (0.2%)	1730 (0.5%)
Stroke	1095 (19.6%)	1571 (0.5%)	2666 (0.8%)
SMI	692 (12.4%)	4099 (1.3%)	4791 (1.4%)
Dementia	616 (11.0%)	738 (0.2%)	1354 (0.4%)

Table 2. Socio-demographic characteristics of multimorbidity cohort, compared with remainder of registered population aged ≥18 years.

Demographic characteristic	Multimorbidity cohort: frequency (valid %) N = 5597	Non-Multimorbidity cohort: frequency (valid %) N = 326,756
Female gender	3042 (54.4%)	161,445 (49.4%)
Age <65 years	1899 (33.9%)	299,742 (91.7%)
Age ≥65-74 years	1249 (22.3%)	15,992 (4.9%)
Age ≥75-84 years	1479 (26.4%)	8038 (2.5%)
Age ≥85 years	970 (17.3%)	2884 (0.9%)
White	3022 (54.0%)	179,859 (55.0%)
Black	1553 (27.7%)	58,939 (18.0%)
South Asian	469 (8.4%)	22,323 (6.8%)
Mixed	197 (3.5%)	15,177 (4.6%)
Other	100 (1.8%)	9804 (3.0%)
Unknown	256 (4.6%)	40,654 (12.4%)
Country of origin: UK*	1413 (46.0%)	69,675 (45.2%)
Language preference: English*	3755 (84.0%)	240,287 (73.8%)
Social deprivation: 1st quintile (most deprived)*	1500 (26.8%)	63,374 (19.4%)
Social deprivation: 2nd quintile*	1232 (22.0%)	63,073 (19.3%)
Social deprivation: 3rd quintile*	995 (17.8%)	66,409 (20.3%)
Social deprivation: 4th quintile*	1013 (18.1%)	66,334 (20.3%)
Social deprivation: 5th quintile (least deprived)*	828 (14.8%)	64,306 (19.7%)

*missing data with reduction in denominator number.

Table 3. Socio-demographic determinants of the multimorbidity cohort: odds ratios derived from multi-level logistic regression modelling with addition of three cardiovascular risk factors: Hypertension, Obesity (moderate), Smoking (ever).

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.10	1.01, 1.19
Age <65 years	(reference group)	
Age ≥65-74 years	4.01	3.69, 4.36
Age ≥75-84 years	8.12	7.46, 8.83
Age ≥85 years	15.71	14.20, 17.38
White	(reference group)	
Black	0.86	0.80, 0.92
South Asian	1.44	1.29, 1.61
Mixed	0.95	0.81, 1.11
Other	0.83	0.67, 1.03
Unknown	0.60	0.52, 0.69
Social deprivation: 1st quintile (most deprived)	1.56	1.41, 1.72
Social deprivation: 2nd quintile	1.35	1.22, 1.50
Social deprivation: 3rd quintile	1.18	1.06, 1.31
Social deprivation: 4th quintile	1.18	1.06, 1.30
Social deprivation: 5th quintile (least deprived)	(reference group)	
Hypertension register	5.05	4.69, 5.44
Moderate obesity	3.41	3.21, 3.63
Smoker (ever)	2.30	2.16, 2.45

FIGURE LEGENDS:

Figures 1-7 display edited alluvial plots showing dominant patient pathways from acquisition of first to second and third Long Term Conditions.

Figure 1: Acquisition sequence of Long Term Conditions; dominant pathways displayed with patient flows ≥ 35 (n = 769).

(FIGURE 1 ABOUT HERE)

Footnote for all Figures: Long Term Condition label abbreviations (n = 12): CHD = Coronary Heart Disease; STRK = Stroke; AF = Atrial Fibrillation; HF = Heart Failure; DM = Diabetes; CKD = Chronic Kidney Disease; MOBES = Morbid Obesity; Dep = Depression; SMI = Serious Mental Illness; DEM = Dementia; COPD = Chronic Obstructive Pulmonary Disease; CP = Chronic Pain.

Figure 2: Most deprived quintile: dominant pathways displayed with patient flows ≥ 13 (n = 145).

(FIGURE 2 ABOUT HERE)

Figure 3: Least deprived quintile: dominant pathways displayed with patient flows ≥ 8 (n = 89).

(FIGURE 3 ABOUT HERE)

Figure 4: 'White' ethnic group: dominant pathways displayed with patient flows ≥ 18 (n = 287).

(FIGURE 4 ABOUT HERE)

Figure 5: 'Black' ethnic group: dominant pathways displayed with patient flows ≥ 15 (n = 227).

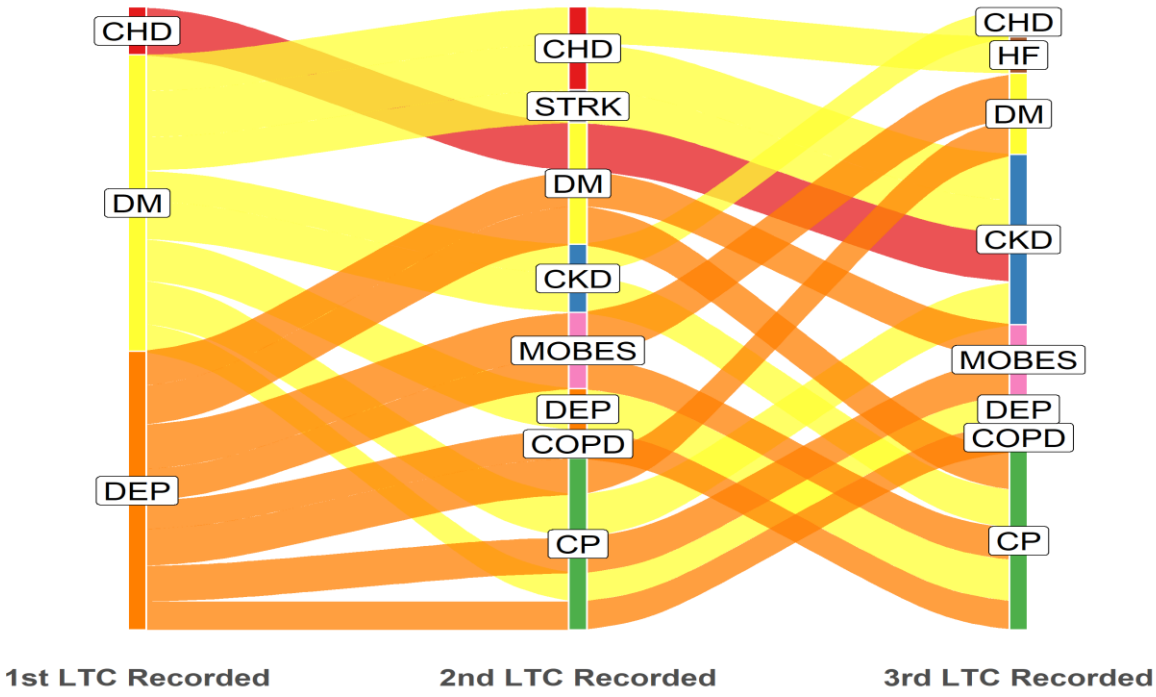
(FIGURE 5 ABOUT HERE)

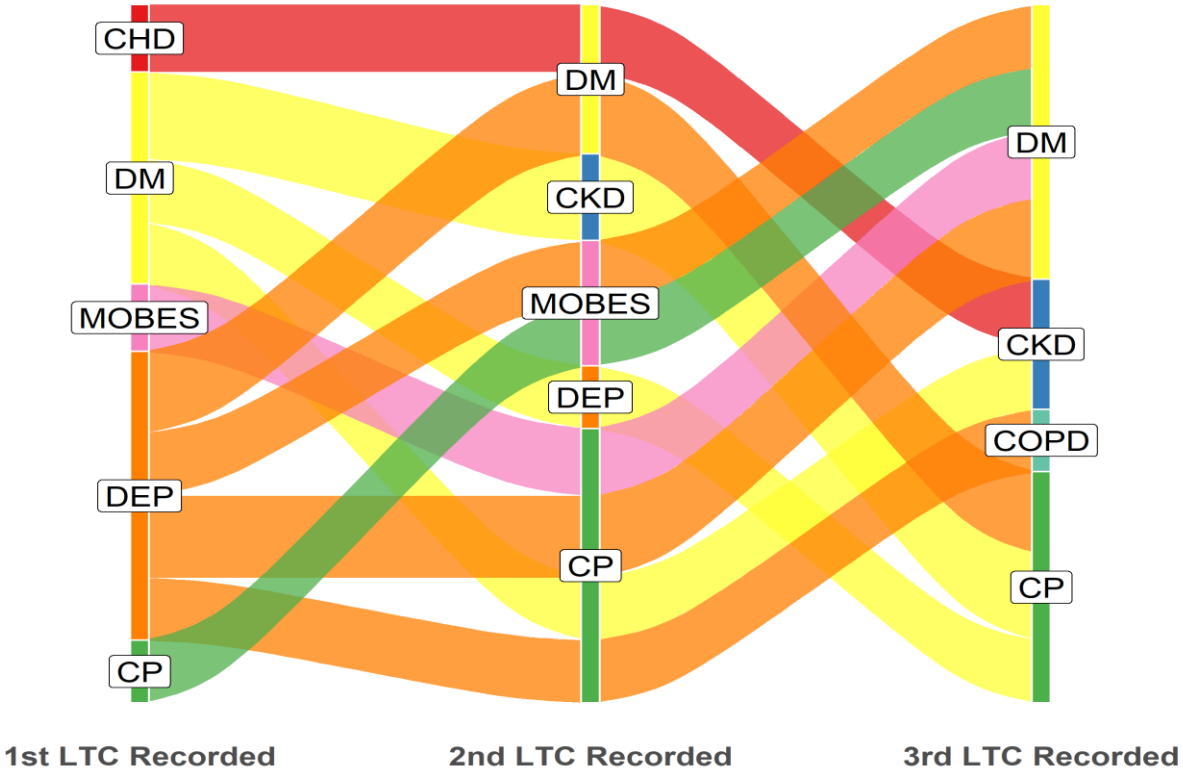
Figure 6: Age under 65 years: dominant pathways displayed with patient flows ≥ 25 (n = 343).

(FIGURE 6 ABOUT HERE)

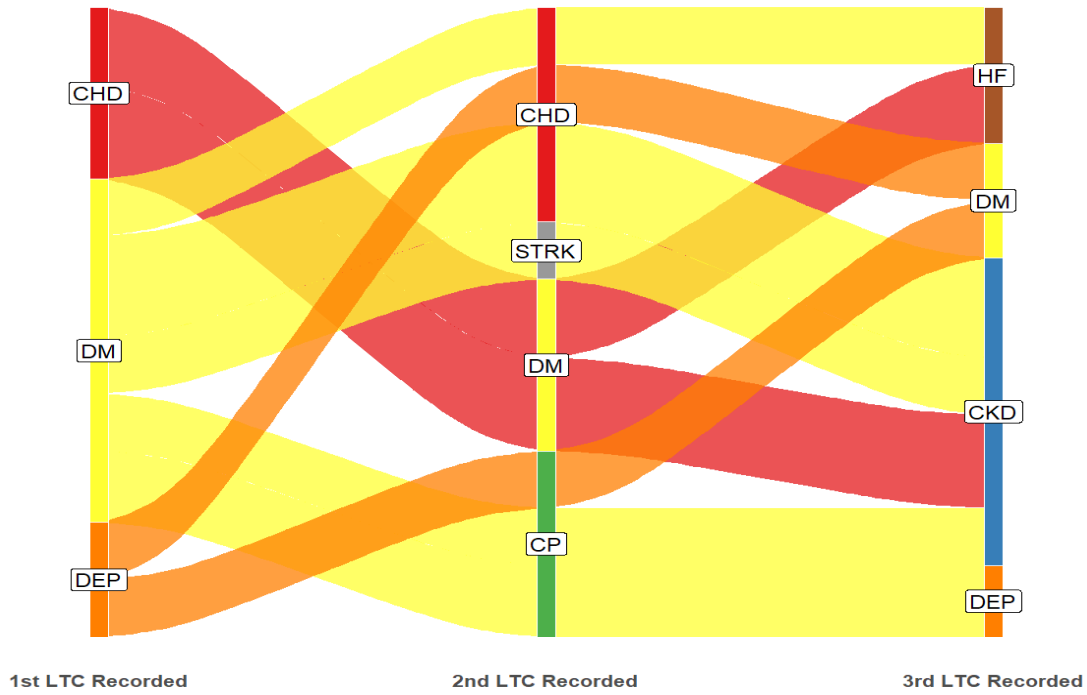
Figure 7: Age 65 years and over: dominant pathways displayed with patient flows ≥ 20 (n = 536).

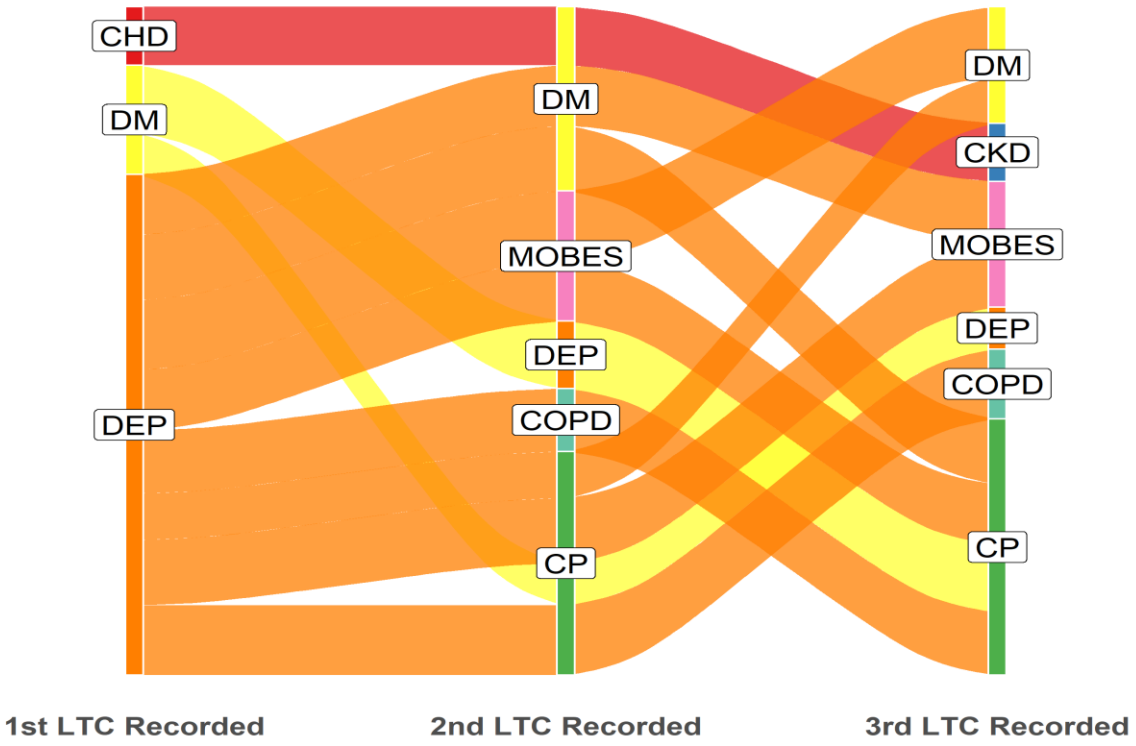
(FIGURE 7 ABOUT HERE)

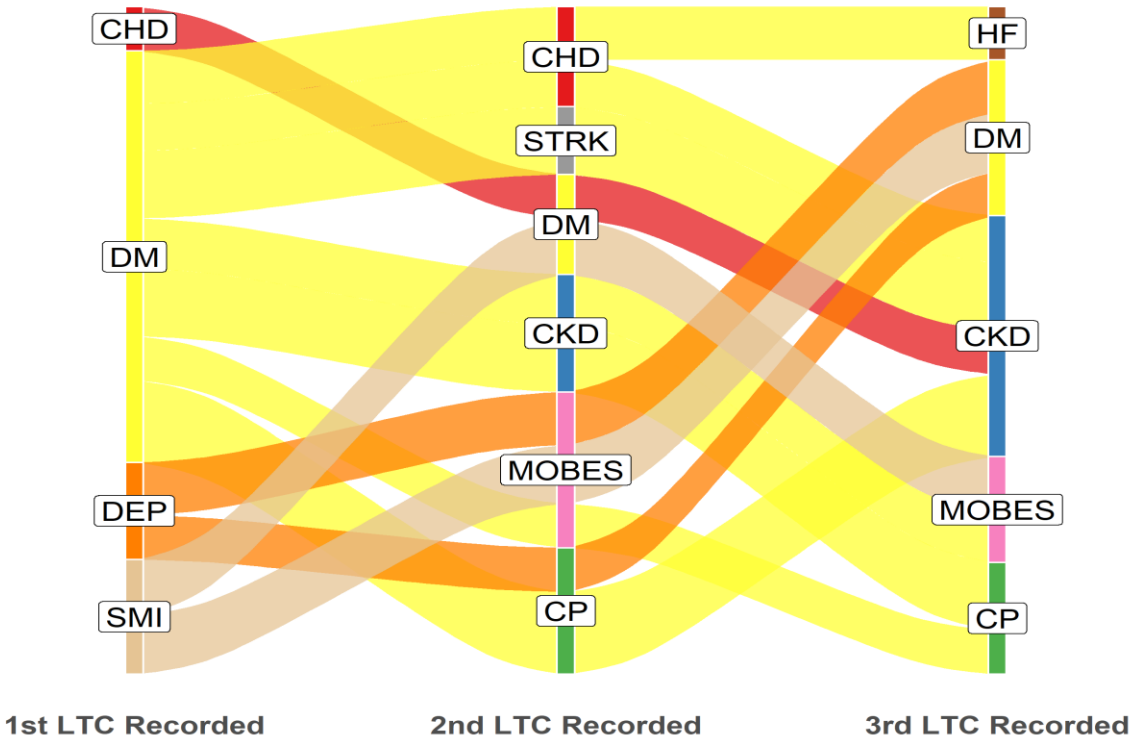


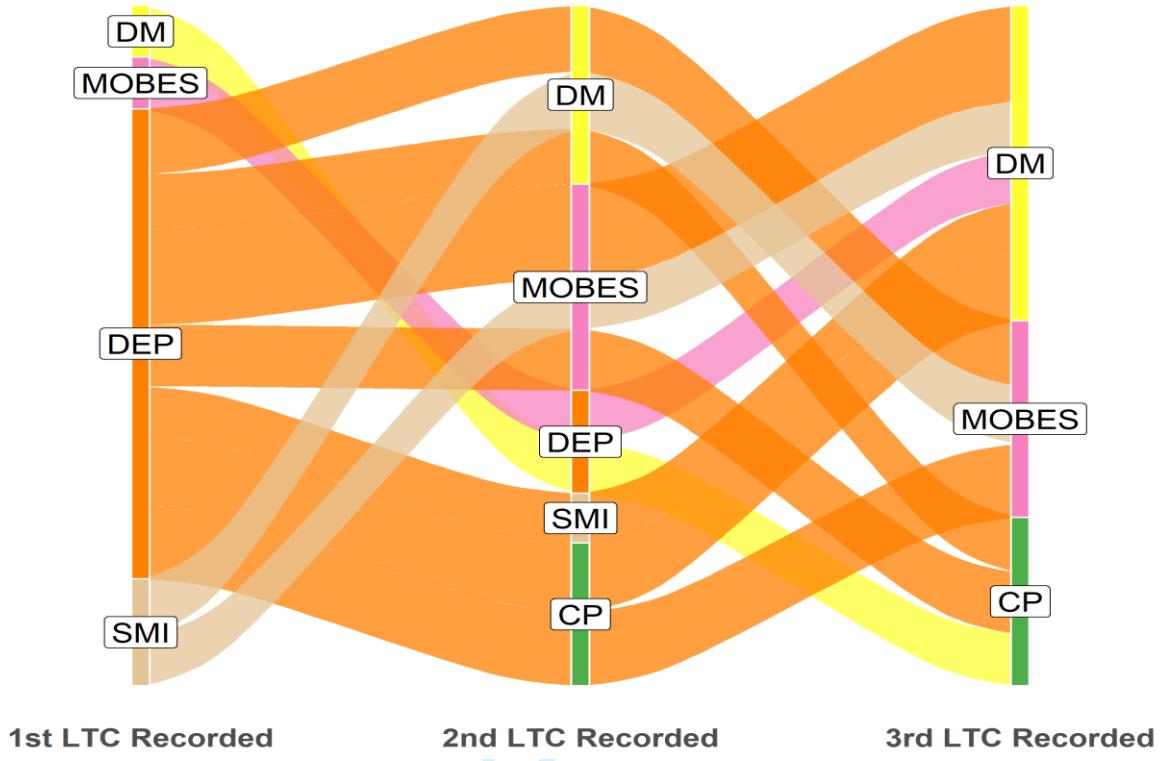


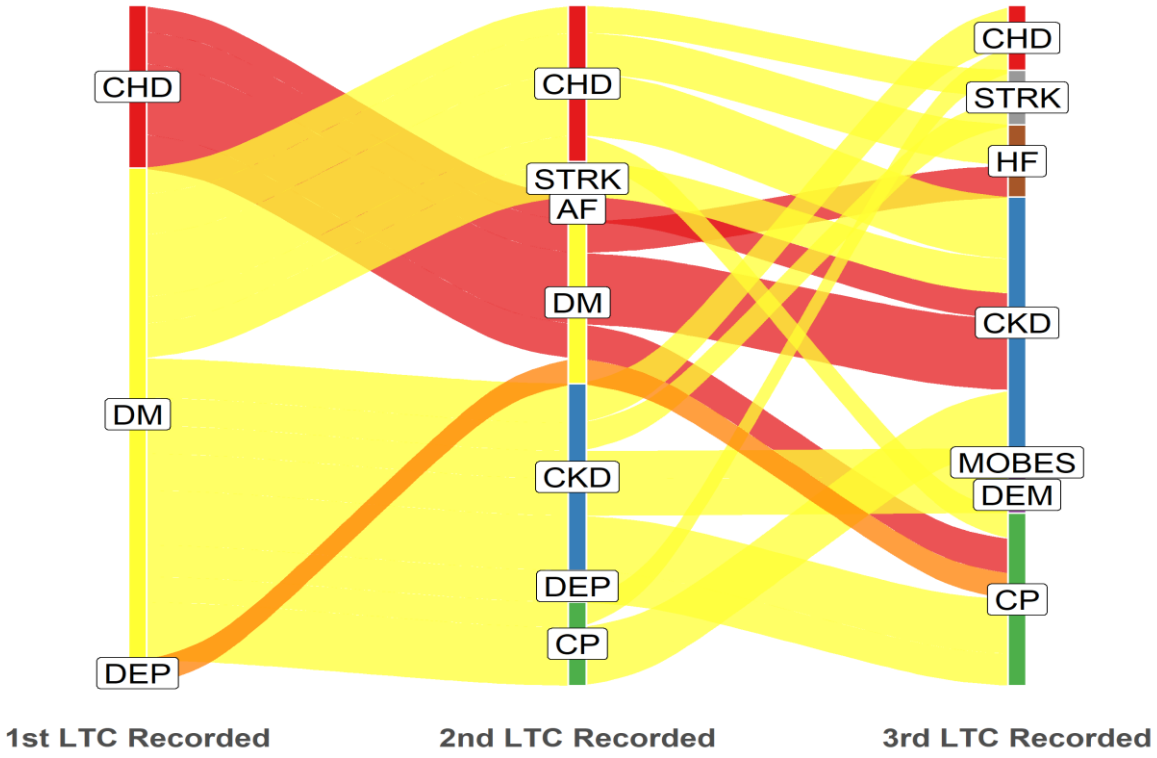
.






CHD
DM
DEP
SMI
CHD
STRK
DM
CKD
MOBES
CP
HF
DM
CKD
MOBES
CP
1st LTC Recorded
2nd LTC Recorded
3rd LTC Recorded




CHD
DM
DEP
CHD
STRK
AF
DM
CKD
DEP
CP
CHD
STRK
HF
CKD
MOBES
DEM
CP
1st LTC Recorded
2nd LTC Recorded
3rd LTC Recorded

Supplementary File

Table 1. Socio-demographic determinants of the multimorbidity cohort: adjusted odds ratios derived from multi-level logistic regression modelling.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.99, 1.11
Age <65 years	(reference group)	
Age ≥65-74 years	11.81	10.96, 12.72
Age ≥75-84 years	27.14	25.22, 29.20
Age ≥85 years	48.70	44.60, 53.18
White	(reference group)	
Black	1.15	1.07, 1.23
South Asian	1.19	1.07, 1.33
Mixed	0.96	0.83, 1.12
Other	0.76	0.62, 0.93
Unknown	0.44	0.38, 0.50
Social deprivation: 1st quintile (most deprived)	1.83	1.66, 2.02
Social deprivation: 2nd quintile	1.58	1.43, 1.75
Social deprivation: 3rd quintile	1.27	1.15, 1.40
Social deprivation: 4th quintile	1.21	1.10, 1.33
Social deprivation: 5th quintile (least deprived)	(reference group)	

Table 2. Socio-demographic determinants of the multimorbidity cohort: odds ratios derived from logistic regression modelling adjusted for clustering at practice level.

Demographic characteristic	Adjusted odds ratio (AOR)	95% confidence intervals for AOR
Female gender	1.05	0.97, 1.13
Age <65 years	(reference group)	
Age ≥65-74 years	12.08	10.53, 13.86
Age ≥75-84 years	27.74	24.28, 31.71
Age ≥85 years	49.84	43.00, 57.78
White	(reference group)	
Black	1.19	1.06, 1.33
South Asian	1.16	1.00, 1.34
Mixed	0.98	0.81, 1.20
Other	0.77	0.62, 0.96
Unknown	0.44	0.36, 0.55
Social deprivation: 1st quintile (most deprived)	1.96	1.69, 2.26
Social deprivation: 2nd quintile	1.66	1.43, 1.92
Social deprivation: 3rd quintile	1.31	1.13, 1.51
Social deprivation: 4th quintile	1.24	1.09, 1.42
Social deprivation: 5th quintile (least deprived)	(reference group)	

Figure 1: Area under Receiver Operating Characteristic (ROC) curve = 0.84; based on regression model presented in Table 5.

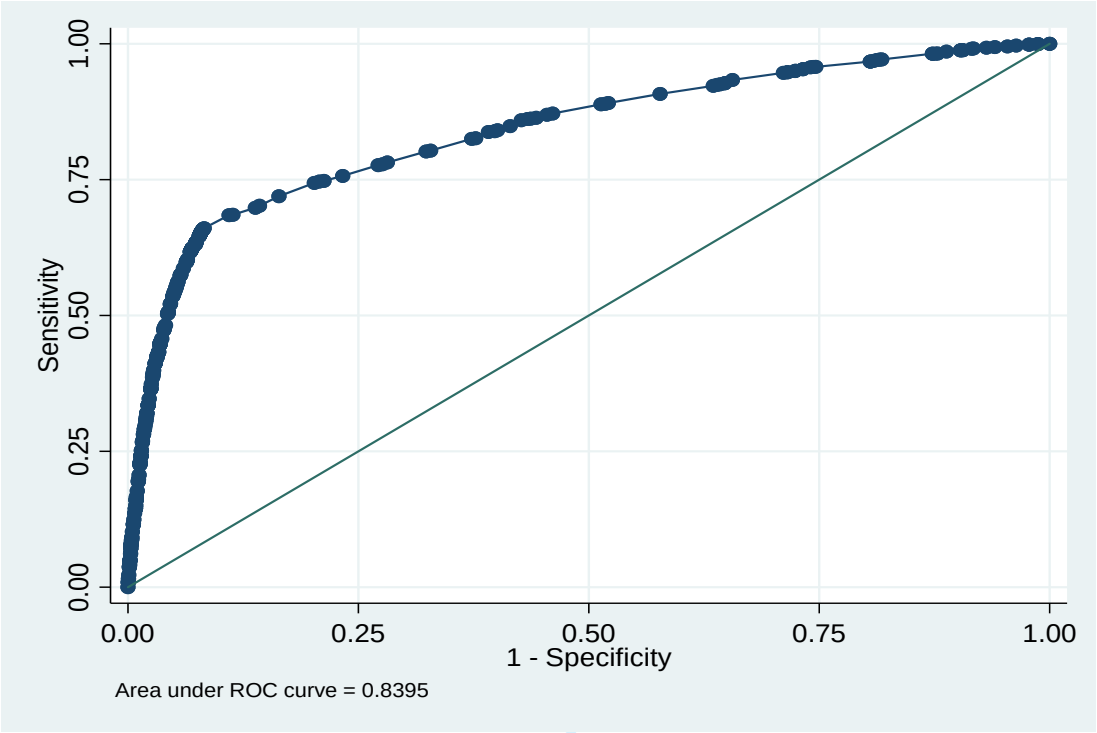
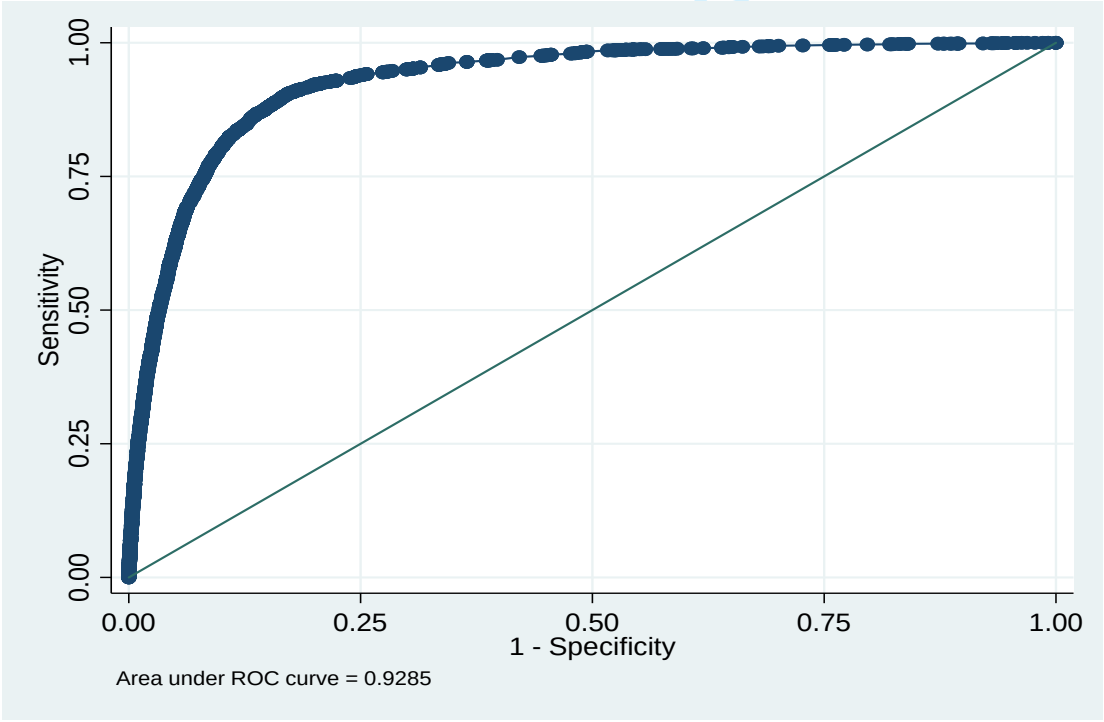
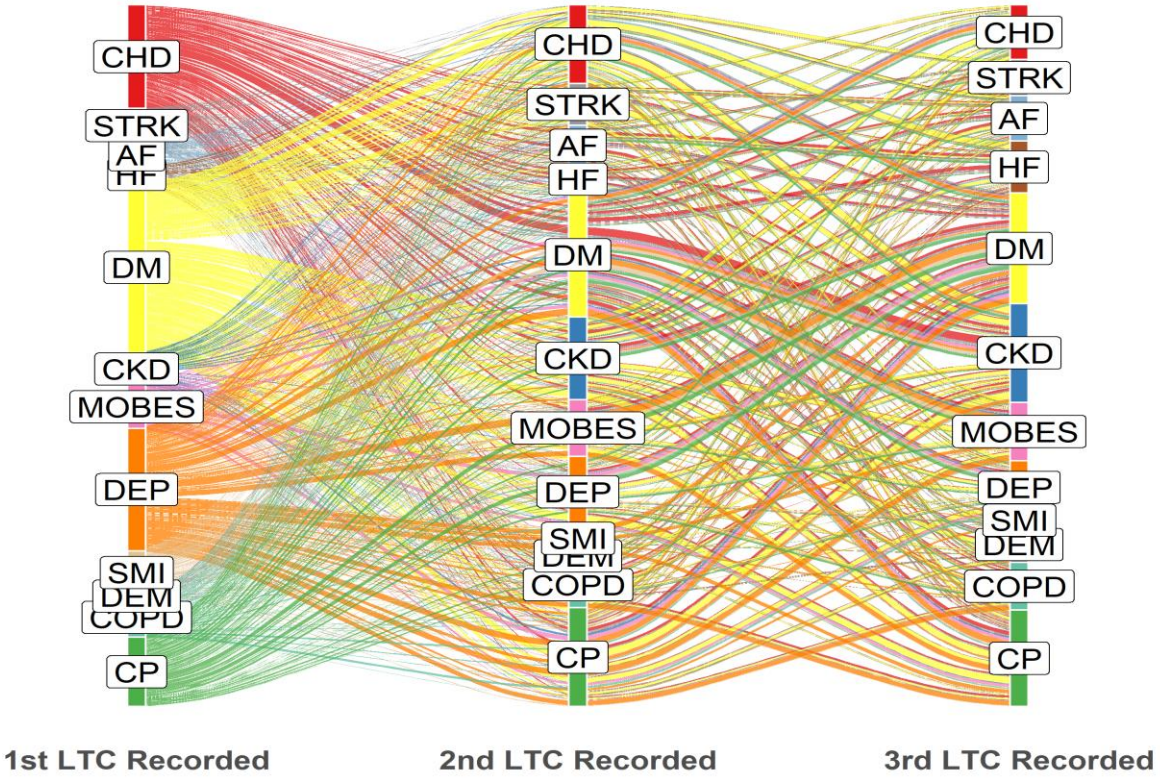


Figure 2: Area under Receiver Operating Characteristic (ROC) curve = 0.93; based on regression model presented in Table 5 with the addition of three risk factors: Hypertension, Smoking (ever), Obesity (moderate).



Figures 3-7 display un-edited alluvial plots for all available data. The figures in the main paper display alluvial plots edited to show dominant pathways.

Figure 3: Acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 5196); data excluded if simultaneous onset dates.



Long Term Condition label abbreviations (n = 12): CHD = Coronary Heart Disease; STRK = Stroke; AF = Atrial Fibrillation; HF = Heart Failure; DM = Diabetes; CKD = Chronic Kidney Disease; MOBES = Morbid Obesity; Dep = Depression; SMI = Serious Mental Illness; DEM = Dementia; COPD = Chronic Obstructive Pulmonary Disease; CP = Chronic Pain.

Figure 4: Most deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1394).

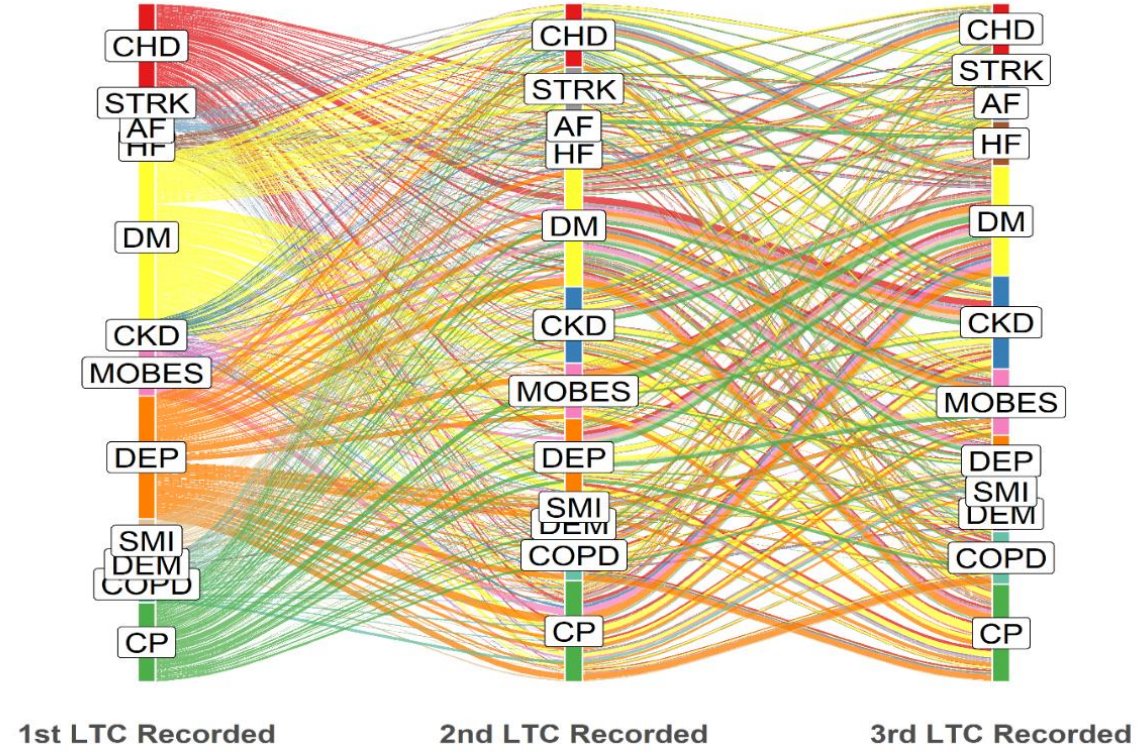


Figure 5: Least deprived quintile: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 763).

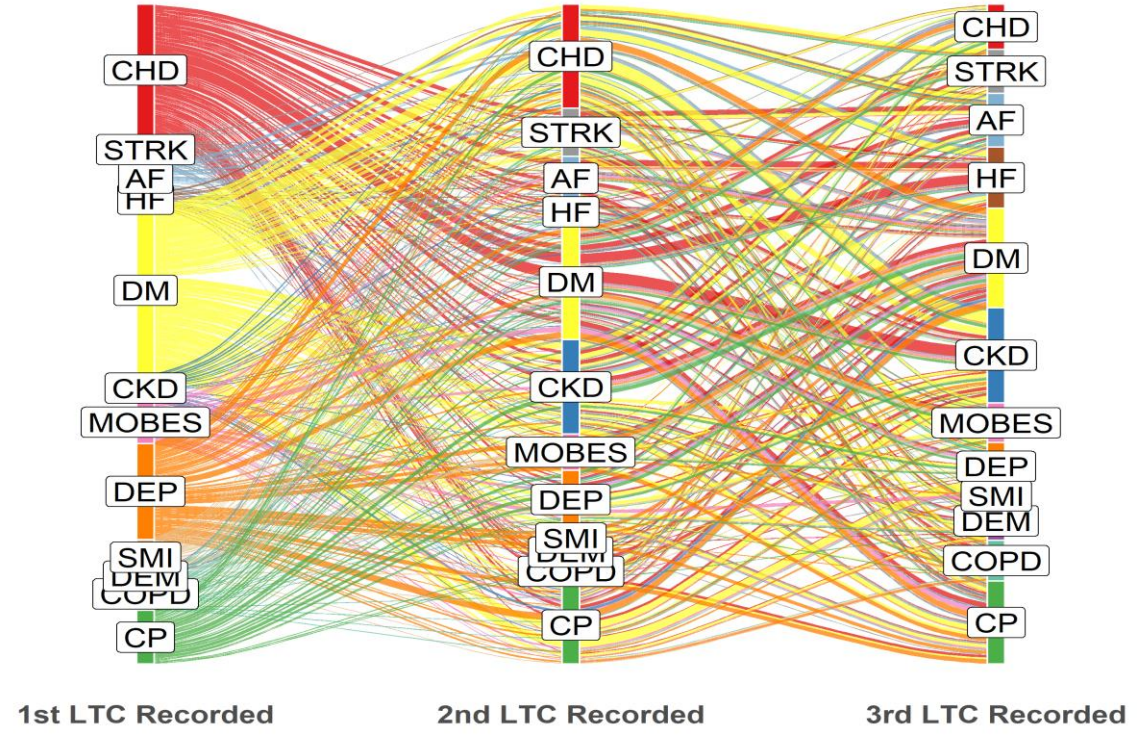


Figure 6: 'White' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 2788).

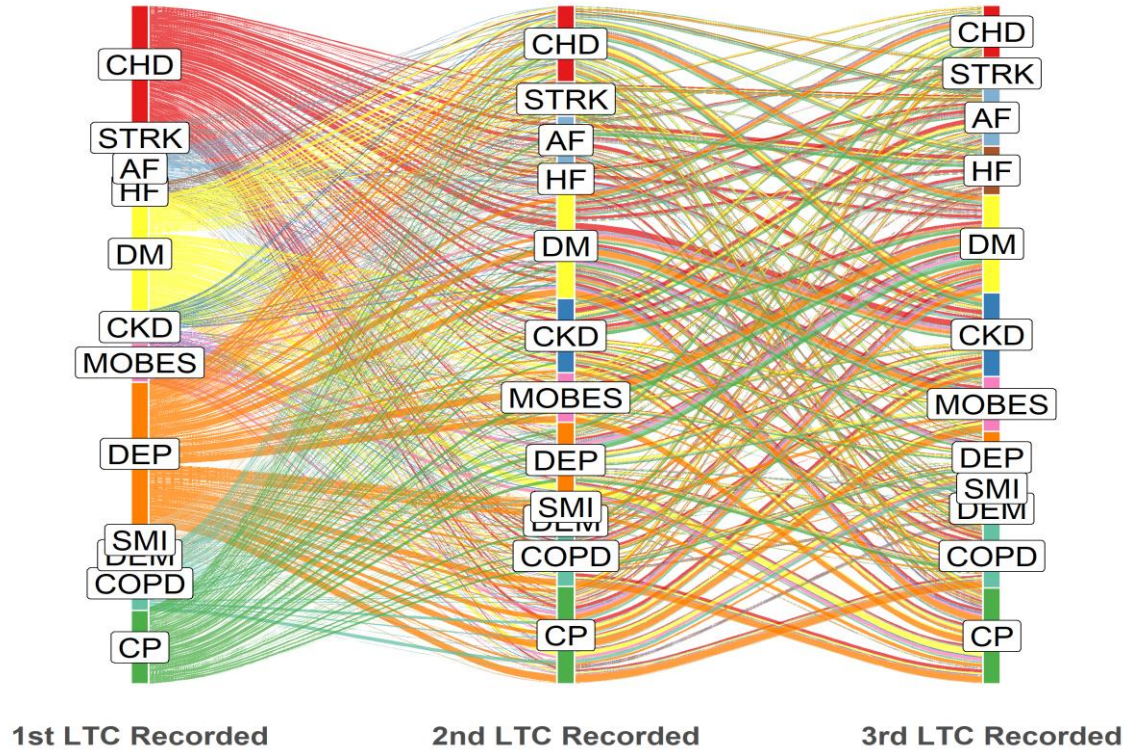
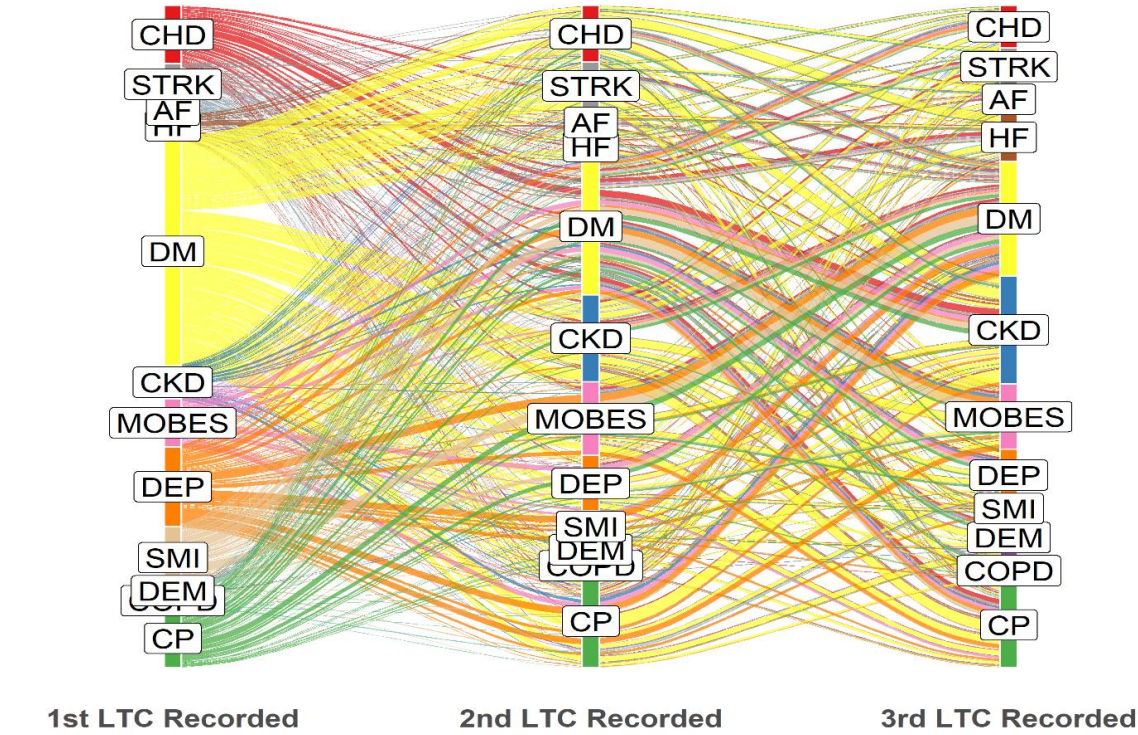


Figure 7: 'Black' ethnic group: alluvial plot displaying acquisition sequence of Long Term Conditions in the multimorbidity cohort (n = 1448).



STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	"Cross-sectional analysis and longitudinal study"
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1	Summarised in Abstract: Design, Setting, Participants, Main outcome measures, Results
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4	The association between multimorbidity, ~~and~~ social determinants and risk factors, and acquisition sequence of Long Term Conditions has not previously been determined in a deprived multi-ethnic community; identifying both the determinants of multimorbidity and the acquisition sequence may suggest interventions to prevent multimorbidity or slow its progression. For the purposes of this study, a locally derived definition of 'multimorbidity' has been used based on predicted high healthcare and social care demand. In contrast, most previously reported studies of multimorbidity have more

				inclusive definitions of multimorbidity.
Objectives	3	State specific objectives, including any prespecified hypotheses	1	"To study the social determinants and risk factors for multimorbidity and the acquisition sequence of multimorbidity"
Methods				
Study design	4	Present key elements of study design early in the paper	5	Presented under 'Data Variables' and 'Data Analysis'.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3	Presented under 'Study Setting', 'Study Design', 'Study population'.
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	4-5	Cross-sectional study: all patients currently (on the data of data extraction) registered at general practices in one south London borough (Lambeth) with the exception of those with an 'informed dissent' code in their case-notes. Longitudinal component: the same cohort of patients studied retrospectively from onset of first Long Term Condition (LTC) to acquisition of 3 or more LTCs.
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed		

		Case-control study—For matched studies, give matching criteria and the number of controls per case		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5	Summarised under the heading, 'Data Variables': "….routinely collected, anonymised, patient-level Read, EMIS and SNOMED coded information. Data were extracted from the EHR into a secure data warehouse …..and contained information on patient demographic characteristics, LTCs, clinical values and medication."
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5	All data was derived from the Electronic Health Record of patients registered at GP practices in the sample area.
Bias	9	Describe any efforts to address potential sources of bias	5-6	Patient records were all included in the analysis, reducing sampling bias. However, 4.0% of patients had an 'informed dissent' code in their case-notes, prohibiting any access to data. We were therefore unable to determine if the omission of these patients may have introduced bias
Study size	10	Explain how the study size was arrived at	5-6	As above – the sample of 332,353 patients represented all patients registered at GP

	practices in the study area, with the exception of those with informed dissent codes. All Odds Ratios were presented with 95% CI's so that the effect of small numbers on the CI could be seen

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5/6	Quantitative variables analysed according to description in section headed, 'Data Analysis'.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	5/6	Regression modelling – see 'Data Analysis'
		(*b*) Describe any methods used to examine subgroups and interactions	6	Sub-groups of the key predictor variables were stratified (into age bands, deprivation quintiles, major ethnic groups)
		(*c*) Explain how missing data were addressed	5/6	See section on missing demographic codes, missing clinical codes, 'informed dissent codes'.
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a	The study was not a sample
		(*e*) Describe any sensitivity analyses	8	See heading 'Sensitivity analyses'
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6/7	Summarised under 'Multimorbidity cohort characteristics'.
		(b) Give reasons for non-participation at each stage	n/a	n/a (exclusion criteria stated above)
		(c) Consider use of a flow diagram	n/a	Exclusion criteria summarised on pg4-6 without use of a Flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Tables 1,2	Table 1 summarises, 'Frequencies of Long Term Conditions included in the multimorbidity cohort'; Table 2 summarises, 'Demographic characteristics of multimorbidity cohort'.

		(b) Indicate number of participants with missing data for each variable of interest	5/6	Missing data numbers summarised under, 'Multimorbidity cohort characteristics'
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a	
		Cross-sectional study—Report numbers of outcome events or summary measures	13	Summary measures in Table 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Tables 1-5	Univariable analysis presented in Tables 1, 2. Multivariable analysis presented in Tables 3,4,5. Table 3 includes different confounder variables to Tables 4,5. All confounder variables are presented in the Tables.
		(*b*) Report category boundaries when continuous variables were categorized	Tables 2-5	Continuous variables were categorised: 'age' into 3 age bands; social deprivation into quintiles
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a	The study was about the high risk of multimorbidity in specified groups, comparing the power of risk factors.

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Table 5; Supplementary file 1	Sensitivity analyses summarised: logistic regression adjusted for clustering; Receiver Operating Characteristic curve
Discussion				
Key results	18	Summarise key results with reference to study objectives	13/14	Key finding summarised at opening of Discussion. The demographic and risk factors for multimorbidity are defined. The acquisition sequence of multimorbidity is described.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14/15	Discussed under heading, 'Strengths and Limitations'
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-17	Cautious interpretation summarised with reference to the literature. Cautious conclusion reached about possible interventions to prevent or delay multimorbidity onset.
Generalisability	21	Discuss the generalisability (external validity) of the study results	14/15	Summarised under, 'Strengths and Limitations'. Findings were derived from one deprived, multi-ethnic community and may not generalise to less deprived, less diverse populations.
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18/19	The study was funded by the Guy's and St Thomas' Charity

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.