# The nature of clusters of multimorbid patients in the UK: a retrospective cohort study

Yajing Zhu, PhD research associate[1], Duncan Edwards, MPH, MRCGP senior clinical research associate[2], Rupert A Payne, PhD MRCGP FRCPE consultant senior lecturer in primary health care[3], Steven Kiddle, PhD research fellow[1]

**Affiliations:** [1] MRC Biostatistics Unit, University of Cambridge, Cambridge, CB2 0SR, UK. [2] The Primary Care Unit, University of Cambridge, Department of Public Health and Primary Care, Strangeways Research Laboratory, Worts Causeway, Cambridge, CB1 8RN, UK.[3]Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2PS.

**Correspondence to:** Yajing Zhu, yajing.zhu@mrc-bsu.cam.ac.uk

**Competing interests:** All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: SJK reports grants from Medical Research Council, during the conduct of the study.

**Contributors:** SJK, YZ and DE conceived and designed the study. DE drafted the protocol, which all authors (YZ, DE, RAP, SJK) contributed to and revised critically. SJK and YZ were responsible for data management. YZ did the statistical analysis and drafted the manuscript, which all authors contributed, revised critically and approved. SJK is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting that criteria have been omitted.

**Transparency declaration:** The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained (see protocol and amendments in supplemental file 1).

**Ethical approval:** This study was approved by the CPRD ISAC, and so is covered by their ethics approval.
**Funding:** SJK and YZ are supported by SJK's MRC Career Development Award (MR/P021573/1).

## Abstract

**Objectives:** To identify, describe and validate clusters of patients based on their multimorbidity, to allow better design of health services and highlight groups that may require tailored interventions.

**Design:** Retrospective cohort study. Patients with multimorbidity were stratified by four age strata and clustered using latent class analysis. Associations between multimorbidity clusters, demographics and outcomes were quantified using generalised linear models.

**Setting:** 382 general practices in England contributed primary care health record data to the Clinical Practice Research Datalink (CPRD).

**Participants:** All multimorbid adults (18 years old or more, with two or more long-term conditions) whose diagnoses are defined in 2012 (N=113,211), from a random sample of CPRD-GOLD (N=391,669). Cluster identification used a random set of 80% of the multimorbid patients (N=90,571), with consistency of results checked in the remaining 20% of multimorbid patients (N=22,640).

**Main outcome measures:** NHS service utilisation was measured by three variables: primary care consultations, hospitalisations and repeat prescriptions in one year after January 2012. All-cause mortality was recorded within two and five years.

**Results:** Clinically distinct and meaningful clusters were identified using robust latent class analysis for 38 long-term conditions within each age strata. Associated patient profiles and outcomes were derived. "Physical-mental health co-morbidity" and "Respiratory disease and multimorbidity" clusters were common across all age strata. In under 65 year olds, "Substance misuse and mental illness co-morbidity" cluster (<7% prevalence, consisted mostly of male, current smokers and patients from derived areas) was found to have the highest demographic-adjusted 5-year mortality rate (in the 45-64 age strata, adjusted odds ratio =1.08 (95% CI 1.07 to 1.10), p<0.01) despite low health service utilisation. Cardiovascular-related clusters were prevalent in over 65 year olds where patients in the cluster "Chronic pain, cardiovascular disease and mental illness" (65-84 age strata) had the highest primary care consultations in one year (median=23, interquartile range [14-35]) and 5-year mortality (39%). In the 85+ age strata, patients in the "Low service use multimorbidity" cluster had the lowest number of morbidities (median=3 [2-4]), service use and mortality. Consistency of results across identification and validation data was confirmed.

**Conclusions:** Across age strata, a clear distinction between morbidity clusters was uncovered, both in the prevalence of long-term conditions within them, and in their associations with outcomes (service use and mortality). Interventions and policies to improve the care of multimorbid patients may be more effective when targeted on the distinct clusters of multimorbidity we have highlighted.

## Summary of contribution

**What is already known on this topic:**

- Multimorbidity is common and is associated with greater healthcare service utilisation and treatment burden.
- Multimorbidity is prevalent in the older population, and in areas with greater socioeconomic deprivation.
- Three main multimorbidity groups have been identified in the literature, involving: (1) cardiovascular-metabolic conditions, (2) mental health related conditions, and (3) musculoskeletal disorders.

**What this study adds:**

- Distinct, validated and clinically meaningful clusters of multimorbid patients have been identified using routine primary care data, with patients broadly representative of the English population.

- Among younger and middle-age individuals, those with substance misuse, alcohol problems, chronic pain and depression (low prevalence) had the highest mortality but relatively low service utilisation.

- Over half of the oldest old multimorbid patients can be grouped in a cluster with relatively low service use.

## Introduction

Owing to advances in medicine, improved life expectancy and ageing populations, a growing number of individuals are living with multimorbidity, i.e. more than one long-term condition (LTC) [1] [2]. Multimorbidity has been recognised as a global challenge for health care management [3] and it is estimated by the Health Foundation that 14 million individuals in England have multimorbidity, with over a third of these having more than four LTCs [4]. Patients with multimorbidity also account for the majority of primary care consultations and hospitalisations in the UK [5]. However, current clinical specialities, guidelines, quality improvement strategies and quality of care metrics are organised around single diseases [1] and treatments of multiple conditions are rarely coordinated, resulting in insufficient or even conflicting care [6].

Patients with multimorbidity have a diverse range of diseases, needs and outcomes [4] [7] [5]. Therefore, identifying and characterising clusters of multimorbid patients – in other words, groups of patients that share similar patterns of LTCs – is an essential step toward improving their healthcare. This would facilitate the development of effective strategies for early diagnosis and prevention of multimorbidity, and allow for a better design and delivery of targeted interventions [1] [8]. Several systematic reviews have found common multimorbidity clusters involving cardiovascular-metabolic conditions, mental health and musculoskeletal disorders [9] [10]. However, existing evidence has important limitations. First, most previous studies have focused on older populations (aged 60+), failing to provide evidence on how multimorbidity should be addressed across the adult population. A few studies have provided age-stratified clusters [9] [11] but they studied only a small list of LTCs, leaving scarce evidence for the younger multimorbid population. Second, systematic reviews have concluded that multimorbidity clusters composed of more than two conditions have not been well profiled mostly due to non-representative and smaller samples [1] [9] [10]. Third, there is substantial heterogeneity in the number of conditions considered (often less than 20) and in the statistical methods. Most studies focused on grouping diseases rather than patients, where each disease can only go into one cluster and so it is not straightforward to relate patients to outcomes in order to facilitate patient-centred policy-making [12]. Commonly used clustering methods were exploratory approaches such as factor analysis and hierarchical clustering [9] [13], where results were highly sensitive to the subjective choice of metrics [14] [15]. Finally, the validity and generalisability of cluster solutions in new samples is important for decision-making but is often ignored in the current literature [9] [10].

This study aims to address the limitation of previous research, and to identify age-stratified clusters of multimorbid adult patients in a large representative sample of UK patients using a

comprehensive list of 38 LTCs developed by Cassell et al. [5] and a robust model-based probabilistic approach, latent class analysis [14].

## Methods

### Data source

Our analysis used the Clinical Practice Research Datalink (CPRD)-GOLD database where anonymised primary care clinical data are contributed by UK general (family) practices (GP) [16]. Almost all UK residents are registered with a National Health Service (NHS) GP. CPRD has been validated to be representative of the UK population for age, sex, and ethnicity [16] [17]. Patients' GP records were linked with hospitalisation data (Hospital Episodes Statistics, HES), all-cause mortality data (Office for National Statistics, ONS) and area socioeconomic deprivation data (Index of Multiple Deprivation, IMD); these linked data were available for approximately 75% of CPRD practices, all of which are based in England. Data are available to approved researchers, for approved projects from CPRD (https://www.cprd.com/). The protocol for this study (16_057RA2) is provided in supplementary file 1, which is covered by the CPRD Independent Scientific Advisory Committee ethics approval.

### Study population

Data on a random selection of individuals were acquired from CPRD (the same individuals studied in Cassell et al., [5]). To be included in our study we required patients to be aged 18 years and above with valid registered-status in a practice with data classified by CPRD as "up-to-standard" in January 2012. We chose the year 2012 to allow complete ascertainment of five-year mortality. Additionally, we required that their practice allowed linkage to ONS, IMD and HES, resulting in the inclusion of only English practices. While the sample of individuals and inclusion criteria match Cassell et al., [5], final patient numbers between the two studies differ based on changes in the number of patients eligible for linkage to HES, ONS and IMD.

### Patient and public involvement

There was no patient or public involvement in this study.

### Statistical software

Data analysis was performed in R 3.4.4. R package names are given in the following sections where appropriate (*in brackets and italics*), for example memory efficient packages were used to extract data for analysis in R (*ff*, *CALIBERdatamanage*). For transparency and reproducibility, all analysis scripts and code lists are available from https://github.com/Kiddle-group.

### Definition of patient characteristics, morbidities and outcomes

Morbidities in this study were defined as binary variables (present or not) based on the classification of LTCs in primary care developed by Barnett et al., [7] This taxonomy attempted to include all conditions "likely to be chronic (defined as having significant impact

over at least the most recent year) and with significant impact on patients in terms of need for chronic treatment, reduced function, reduced quality of life, and risk of future morbidity and mortality", and was developed for use in UK primary care electronic health record research [7] and has been adapted for use in CPRD [5]. The specific definitions for each LTC is based on the UK Read code system and electronic prescription data coded using CPRD's prodcode; these have been updated from Cassell et al., [5], giving a total of 38 LTCs. (https://www.phpc.cam.ac.uk/pcu/cprd_cam/codelists/v11/ ). The LTCs used in this study largely match the only other large sample size UK multimorbidity cluster study [18].

Two sets of outcome variables related to service use and mortality were defined. NHS service utilisation or treatment burden was measured by three variables over the 12-month period after January 2012: primary care consultations (consultations with any clinician in the primary care team), the number of all-type hospitalisation spells (defined by discharge dates) and the amount of repeat prescriptions (counting the unique British National Formulary (BNF) codes). All-cause mortality at two and five years was extracted from ONS data.

Patient characteristics that were considered in this study include gender, age groups (stratified into 18-44, 45-64, 65-84 and 85+ years), last recorded pre-2012 body mass index (BMI) in the original continuous scale, last recorded pre-2012 smoking status (current, never and ex-smokers) and socioeconomic status measured by IMD in quintiles (1 for the least deprived and 5 for the most).

### Statistical analysis

Multimorbidity is heterogeneous and our aim was to identify distinct patient groups (i.e. multimorbidity clusters) such that within each cluster patients had similar patterns of morbidities. Many previous studies identify clusters by counting the prevalence of pairs of conditions, an approach which does not take into account the prevalence of each individual condition and focuses on conditions rather than patients [9]. Therefore, we used cluster analysis, an approach that assigns all patients to non-overlapping clusters (i.e. each patient is assigned to only one cluster) in a data-driven fashion [19] [20]. Specifically, we used latent class analysis (LCA) (*poLCA*). Compared to other exploratory clustering methods (e.g. factor analysis, hierarchical clustering method, [21] [22], LCA is a model-based probabilistic clustering approach that is not sensitive to rotation of factors and does not require any subjective choice of "distance measures" for multimorbidity patterns [14] [15] [23]. This greatly enhances the reproducibility and stability of the latent class solutions.

Guided by simulation studies performed by [24], the optimal number of latent classes was decided using a combination of statistics (Bayesian Information Criteria (BIC), sample-sized adjusted BIC, log-likelihood ratio test, entropy for classification quality) and clinical judgment. Within our datasets conditions are present (i.e. recorded) or not by definition, and so missing data methods were not needed for cluster analysis. More details on a technical review of commonly used clustering methods, the LCA methodology and application of selection statistics are provided in supplementary file 2, section 3.

To account for the different nature of multimorbidity clusters at different ages, four clinically meaningful age strata (18-44, 45-64, 65-84, 85+ years) were chosen. We derived the cluster solution and performed post-hoc statistical tests in a randomly selected subset of the multimorbid population that contained 80% of the patients, stratified by age group (i.e. training set). Separate LCAs were performed for each strata, and each patient allocated to a

single multimorbidity cluster. For ease of interpretation, clusters were labelled by their three most distinctive conditions, defined as the three conditions whose difference in prevalence between cluster and age strata were the highest. Short versions of condition names (Table 1) were used in tables and figures for ease of reference. To quantify the association between outcomes, multimorbidity clusters and patient demographics, generalised linear models were fitted. Individuals with missing data for last pre-2012 recording of smoking status and BMI represented a small percentage (<5%) of the population, and so were excluded from the generalised linear models (i.e. complete case analysis).

### *Assessment of the stability of morbidity clusters*

To assess the stability of age-stratified multimorbidity clusters, LCA was repeated in the remaining 20% of the population (i.e. test set), fixing the number of clusters to match that learned from the training set [25]. We employed three methods to indirectly validate our cluster solutions (a direct approach was not possible as clusters were unobserved). First, to check the consistency between cluster profiles for 38 LTCs in the training and test sets, each cluster in the test set was matched (using two criteria for robustness) with a corresponding cluster in the training set. Matched cluster pairs were selected such that Jensen–Shannon distance (JSD, a measure of the divergence between disease distributions [26]) is the smallest and the bivariate Pearson's correlation coefficient (the degree to which two disease profiles co-vary [27]) is the highest (supplementary tables 3a, 3b). Second, entropy measures [24] (for classification quality) computed in the training and test sets were expected to be similar. Finally, stability was further assessed by observing that the association between clusters, patient demographics and outcome variables were similar (in terms of size, direction and statistical significance). For more details see supplementary file 2 section 4.

## Results

### Characteristics of the study population

Out of the random sample (N=391,669), 49% and 22% of individuals had none or only one LTC respectively. The prevalence of each condition is provided in Table 1. Among the multimorbid patients (N=113,211, accounting for 29% of the sample), all unique combinations of conditions were less than 1% prevalent in the total population with the most prevalent 20 containing only pairs of conditions (supplementary table 1). This together with the large number of unique combinations of conditions (supplementary table 2) indicated that multimorbidity patterns were highly heterogeneous.

Key demographics of the whole sample (N=391,669) are summarised in Table 2. Females, older individuals and those from areas of greater socioeconomic deprivation had a higher prevalence of multimorbidity. Focusing on the multimorbid population (N=113,211, Table 3), for patients aged 18-44 years, multimorbidity was more common in areas with greater deprivation while for older groups, individuals from less deprived areas were more likely to have multiple LTCs.

### Multimorbidity clusters and outcomes

Latent class models were used to identify multimorbidity clusters using a random sample of 80% of the multimorbid patients in each age strata (N=90,571). These clusters differ across age strata, both in terms of number of clusters per strata and main components within each

cluster (Figures 1-4, Table 5). The association between multimorbidity clusters and outcomes (service use and mortality) remained significant ($p<0.01$ in almost all clusters) after stratifying by age strata, controlling for socioeconomic status and smoking behaviour (supplementary tables 7-10). Among the multimorbidity clusters, the cluster with the lowest impact on the outcomes of interest (in that age strata) was chosen as the reference cluster that all other clusters were compared to. Seven types of clusters were particular noteworthy, as summarised Table 4.

### *Main findings*

First, two sets of multimorbidity clusters were found to be common across all age strata: one that included clusters characterised by physical-mental health co-morbidity (characterised by depression, anxiety and painful conditions) and the other that was mostly composed of clusters of respiratory disease with multimorbidity (characterised by COPD and asthma). Second, clusters of patients with high rates of substance misuse and mental illness existed in both the 18-44 and 45-64 age strata and had the highest mortality rates (4% and 13% mortality within five years, respectively), but low service utilisation. Three different types of clusters with high rates of cardiovascular conditions occurred across patients aged 45 and over, and notably it was those clusters with the highest rates of established cardiovascular diseases (coronary heart disease, heart failure and atrial fibrillation), that had the highest rates of health service use and death (five-year mortality rates are 39% in the 65-84 age strata and 71% in the 85+ age strata). Clusters characterised by cardiovascular risk factors such as hypertension, diabetes and chronic kidney disease were common but had low mortality (13% mortality in five years in the 65-84 age strata) and hospitalisation spells (mean=0.6 per year). Among the oldest old (85+), the majority (58%) formed into the “Low service use multimorbidity” cluster, with the lowest rates of health care use, mortality, and total morbidities (approximately 40% lower than the rates of the highest risk cluster).

More results for the distribution of outcomes (e.g. median, interquartile range (IQR)), covariate-adjusted incidence rate ratios (IRR) for service utilisation and odds ratios (OR) for mortality derived from generalised regression models are available below and in supplementary tables 7-14 (covariates include gender, socioeconomic status, smoking status, BMI and age), but with a cautionary note that the relationship between clusters, patient demographics and outcomes should not be interpreted causally due to unmeasured confounding and simplified linear assumptions.

### *Under 65 year olds*

Five clusters were uncovered in the 18-44 age strata, with the most prevalent three being [Depression, Anxiety, Pain] (32%), [Pain, Hearing loss, Hypertension] (23%) and [Asthma, IBS (irritable bowel syndrome), Depression] (20%). Patients in these clusters also tended to have higher service utilisation. Compared to patients in the cluster with the lowest service use ([IBS, Depression, Hearing loss]), those in the physical-mental health co-morbidity cluster [Depression, Anxiety, Pain] were found to have the highest use of primary care consultations (median=12 per year, IQR [5-20], IRR=1.35 (95% CI 1.28 to 1.43), $p<0.01$) and those in cluster [Pain, Hearing loss, Hypertension] were found to have the highest admission rates (mean=0.6 spells per year, IRR=1.77 (95% CI 1.53 to 2.05), $p<0.01$). Prescribing rates were also particularly high in both these clusters. Highest mortality was found to be associated with the least prevalent (7%) multimorbidity cluster characterised by substance misuse and

mental illness co-morbidity [Drug-misuse, Alcohol, Depression] (4% mortality in five years, OR = 1.03 (95% CI 1.01 to 1.04), p<0.01).

In the 45-64 age strata, LCA revealed five clusters, with the most prevalent three being [Hypertension, Diabetes, Pain] (37%), [IBS, Hearing loss, Pain] (24%) and [Depression, Pain, Anxiety] (22%). Highest service use (per patient) was mostly associated with clusters with physical-mental health co-morbidities. For example, compared to patients in the cluster [IBS, Hearing loss, Pain] with the lowest service use and mortality, patients in the physical-mental health co-morbidity cluster [Depression, Pain, Anxiety] had the highest number of primary care consultations (median=14 per year, IQR [7-23], IRR=1.52 (95% CI 1.47 to 1.58), p<0.01), hospital spells (mean=0.6 per year, IRR=1.31 (95% CI 1.20 to 1.44), p<0.01) and prescribing rates (median = 4 unique drug classes over 12 months , IQR [2-7], IRR=2.37 (95% CI 2.29 to 2.46), p<0.01). As in the younger age strata, the least prevalent multimorbidity cluster characterised by substance misuse co-morbidity [Alcohol, Drug-misuse, Pain] (4%), was again associated with the highest deaths rate (13% mortality in 5 years, OR=1.08 (95% CI 1.07 to 1.10), p<0.01).

### *Above 65 year olds*

Six clusters were found in the 65-84 age strata where the commonest clusters were cardiovascular risk factors [Hypertension, Diabetes, Kidney disease] (41%) and physical-mental co-morbidities [Depression, Pain, Anxiety] (14%). Relative to the cluster with the lowest service use and mortality ([Hearing loss, Prostate disorder, IBS]), the least prevalent multimorbidity cluster, characterised by high rates of chronic pain, established cardiovascular disease and mental illness [Pain, CHD (coronary heart disease), Depression] (5%) was associated with the highest use of primary care consultations (median=23 per year, IQR[14-35], IRR=1.92 (95% CI 1.82 to 2.02), p<0.01), hospitalisations (mean=0.6 spells per year, IRR=2.15 (95% CI 1.94 to 2.40), p<0.01), prescribing rates (median = 11 unique drug classes over 12 months, IQR [8-14], IRR=2.88 (95% CI 2.77 to 3.00), p<0.01] and mortality (39% mortality in 5 years, OR =1.26 (95% CI 1.24 to 1.29), p<0.01).

The 85+ age strata was composed of four clusters with the majority of patients (58%) fitting within a “Low service use multimorbidity” cluster [Hypertension, Hearing loss, Diabetes]. Comparing against patients in the cluster characterised by established cardiovascular disease [Heart failure, CHD, Atrial fibrillation] with the highest mortality rate (71% mortality in five years), cluster [Hypertension, Hearing loss, Diabetes] had the lowest mortality (50% five-year mortality, OR = 1.23 (95% CI 1.19 to 1.27), p<0.01), fewer LTCs (median number of morbidities = 3, IQR [2-4], versus the highest 7, IQR [6-8]), and the least health care utilisation (roughly half the amount of the cluster with highest use) .

## Demographic profiles of multimorbidity clusters

The associations between patient characteristics and different clusters are described in Table 6 (in descriptive statistics) and supplementary tables 11-14 (for model-based adjusted odds ratios and incidence rates; reference categories are male, non-smokers and the least deprived). Across all age strata, physical-mental co-morbidity clusters contained more female patients (over 66%, adjusted ORs are consistently around 1.1, p<0.01). Socioeconomic deprivation and smoking were strongly associated with “Substance misuse and mental illness co-morbidity” and “Physical-mental health co-morbidity” clusters (adjusted ORs for being in respective clusters are consistently around 1.1, p<0.01) in patients under 65 years old. In

older age strata, smoking was mainly associated with clusters involving respiratory diseases. For example, in the 65-84 age strata, cluster [COPD, Asthma, Pain] contained 24% current smokers (adjusted OR for being in this cluster =1.13 (95% CI 1.04 to 1.10), $p<0.01$). There was no evidence ($p>0.1$) for any association between deprivation and multimorbidity clusters in the 85+ age strata. The relationship between multimorbidity clusters and patient characteristics were highly consistent in the training and test sets (supplementary tables 5-6).

### Validation of cluster morbidity profiles

As well as validating the clusters by their association with patient characteristics and outcomes, the similarity of multimorbidity clusters were compared between the training set (80% of patients, N=90,571) and the test set (the other 20% of patients, N=22,640). Results are summarised below, and given in full in supplementary file 2 section 4. Measures of cluster quality (i.e. entropy) were found to be consistent between the training and test sets. As the training set contained more disease patterns, the derived clusters were more comprehensive. The test set (with fewer patients) contained fewer disease patterns and therefore we expect the derived clusters to be a subset of those in the training set. Validation of cluster profiles showed that every cluster in the test set found a match in the training set. Some clusters were particularly robust (had the smallest JSD and the highest Pearson's correlation coefficient), for instance, those in the largest age strata (65-84 age strata, N = 49,494), and clusters characterised by physical-mental co-morbidities or by alcohol and drug mis-use. Clusters with a less clear match were the respiratory disease and multimorbidity cluster in the 18-44 age strata and the clusters in the 85+ age strata, mostly due to patient heterogeneity in the smaller test set.

## Discussion

### Summary of results and comparison with other studies

This study identified and validated clinically meaningful clusters of patients with distinct patterns of multimorbidity across four age strata in the UK population. Our identification of clusters comprising cardio-metabolic disease and physical-mental health comorbidities is reassuringly consistent with other work [9] [10] [18] [28], but our results additionally benefits from a robust methodological approach and insights into important associated outcomes. There are also further specific findings of particular note. Firstly, a cluster of "substance misuse and mental illness comorbidity" was found in younger persons (<65), characterised by greater socioeconomic deprivation and a large male majority, and associated with low service use but high mortality. Within the cardiovascular-related clusters observed in the older population, we were also able to identify three distinct clusters: cardiovascular risk factors, established cardiovascular disease, and cardiovascular disease comorbid with chronic pain and mental illness. The first cluster was common but with lower rates of death and service use. In contrast, the third cluster was least prevalent but had substantially higher service use and mortality. Finally, we identified a large cluster within the oldest old (85+), which has not been previously described, characterised by relatively few morbidities, low service use and low mortality.

### Strengths and limitations

To the best of our knowledge, this is the first and the most comprehensive study to identify clusters of multimorbid patients using a large representative random sample from UK primary care records, and to relate multimorbidity clusters to mortality and service use. By including younger patients and stratifying by age we see how multimorbidity clusters differ over the life course. Our results are based on a robust model-based probabilistic approach (LCA, that does not rely on arbitrary choice of metrics), have been validated (evident through reasonable consistency between the main findings and those from a smaller held-out random sample), and are easily reproducible (diagnostic and programming codes are shared).

Limitations of this study include that patients may be misclassified due to undiagnosed, misdiagnosed or unrecorded conditions. While CPRD is representative of the UK population, recording rates for relevant primary care record codes may systematically differ between practices. Additionally, smaller or less-organised practices are less likely to participate in CPRD. This could affect the relevance of the results to certain GP practices. Methodologically, the LCA clustering approach used here is a robust probabilistic approach, but results may differ subtly if other approaches are used. Thorough methodological research is required to compare clustering approaches and account for datasets with varied patient characteristics and LTCs. Validation of latent clusters also requires further research where a larger sample size for the test set can strengthen the validation of patient assignments. Despite this, given the large and representative sample, the consistency of results both internally, across age strata and with existing literature, we are confident in our main results. Finally, multimorbidity evolves over time, but we only use longitudinal data to help define the presence in 2012 of conditions, and their relationship to patient outcomes. Future research can focus on identifying the development of multimorbidity clusters and patient outcomes over time, with results validated in multiple countries.

## Conclusion and policy implications

While patients with multimorbidity account for an ever increasing proportion of healthcare need and provision [1] [4] [5] [7], no existing interventions have shown convincing evidence of benefit in improving important outcomes [8] [29]. Our findings suggest that one reason for the failure of previous interventions is that multimorbidity is heterogeneous, with very different diseases, needs and outcomes in different clusters of patients. We propose that interventions to improve outcomes in multimorbidity may be more appropriately targeted on distinct types, and we have systematically highlighted clusters of patients and groups of clusters where tailored approaches should be attempted. For example, individuals within "physical-mental health co-morbidity" clusters may benefit from widespread development and adoption of integrated care interventions [30] [31].

Likewise, health service developments could take account of the clustering of multimorbidity to stratify care more appropriately. Clusters with relatively high mortality but low healthcare utilisation such as those characterised by substance misuse and mental illness may benefit from better access to support services and more proactive, targeted care provision. Conversely, a large proportion of patients over 85 years old with multimorbidity may not need intensive medical input, and patients within clusters characterised by high levels of cardiovascular risk factors but low levels of established cardiovascular disease or co-morbid disease could be targeted by initiatives to promote self-care and self-monitoring of LTCs such as hypertension [32] [33].

## Tables and Figures

## Supplementary files

Supplemental file 1 – ISAC protocol
Supplemental file 2 – Supplemetnary.docx
Supplemental file 3 – STROBE checklist

## Acknowledgements

We acknowledge CPRD @ Cambridge for developing and sharing disease definitions, and Dr Jennifer Quint (Imperial College London) for permission to use and share a codelist for smoking status. We also acknowledge the valuable statistical discussions with Dr. Robert Goudie and Dr. Paul Kirk at MRC Biostatistics Unit, University of Cambridge. This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare Products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. ONS is the provider of ONS mortality data used in this study. ONS and HES data copyright © (2018), re-used with the permission of The Health & Social Care Information Centre. All rights reserved. The interpretation and conclusions contained in this study are those of the author/s alone.

## References

[1] Academy of Medical Sciences, *Multimorbidity: a priority for global health research,* Academy of Medical Sciences, London, 2018.

[2] World Health Organization, *Multimorbidity,* World Health Organization, 2016.

[3] E. Wallace, C. Salisbury, B. Guthrie, C. Lewis, T. Fahey and S. M. Smith, "Managing patients with multimorbidity in primary care," *BMJ,* vol. 350, pp. h176--h176, 1 2015.

[4] *Briefing: Understanding the health care needs of people with multiple health conditions.*

[5] A. Cassell, D. Edwards, A. Harshfield, K. Rhodes, J. Brimicombe, R. Payne and S. Griffin, "The epidemiology of multimorbidity in primary care: a retrospective cohort study," *British Journal of General Practice,* vol. 68, pp. e245--e251, 3 2018.

[6] K. Chaplin, P. Bower, M.-S. Man, S. T. Brookes, D. Gaunt, B. Guthrie, C. Mann, S. W. Mercer, I. Rafi, A. R. G. Shaw and others, "Understanding usual care for patients with multimorbidity: baseline data from a cluster-randomised trial of the 3D intervention in primary care," *BMJ open,* vol. 8, p. e019845, 2018.

[7] K. Barnett, S. W. Mercer, M. Norbury, G. Watt, S. Wyke and B. Guthrie, "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study," *The Lancet,* vol. 380, pp. 37-43, 7 2012.

[8] C. Salisbury, M.-S. Man, P. Bower, B. Guthrie, K. Chaplin, D. M. Gaunt, S. Brookes, B. Fitzpatrick, C. Gardner, S. Hollinghurst, V. Lee, J. McLeod, C. Mann, K. R. Moffat and S. W. Mercer, "Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach," *The Lancet,* vol. 392, pp. 41-50, 7 2018.

[9] A. Prados-Torres, A. Calderón-Larrañaga, J. Hancco-Saavedra, B. Poblador-Plou and M. Akker, “Multimorbidity patterns: a systematic review,” *Journal of Clinical Epidemiology,* vol. 67, pp. 254-266, 3 2014.

[10] C. Violan, Q. Foguet-Boreu, G. Flores-Mateo, C. Salisbury, J. Blom, M. Freitag, L. Glynn, C. Muth and J. M. Valderas, “Prevalence, Determinants and Patterns of Multimorbidity in Primary Care: A Systematic Review of Observational Studies,” *PLoS ONE,* vol. 9, p. e102149, 7 2014.

[11] B. Poblador-Plou, M. Akker, R. Vos, A. Calderón-Larrañaga, J. Metsemakers and A. Prados-Torres, “Similar Multimorbidity Patterns in Primary Care Patients from Two European Regions: Results of a Factor Analysis,” *PLoS ONE,* vol. 9, p. e100375, 6 2014.

[12] J. Collerton, C. Jagger, M. E. Yadegarfar, K. Davies, S. G. Parker, L. Robinson and T. B. L. Kirkwood, “Deconstructing Complex Multimorbidity in the Very Old: Findings from the Newcastle 85 Study,” *BioMed Research International,* vol. 2016, pp. 1-15, 2016.

[13] E. Ahlqvist, P. Storm, A. Käräjämäki, M. Martinell, M. Dorkhan, A. Carlsson, P. Vikman, R. B. Prasad, D. M. Aly, P. Almgren, Y. Wessman, N. Shaat, P. Spégel, H. Mulder, E. Lindholm, O. Melander, O. Hansson, U. Malmqvist, Å. Lernmark, K. Lahti, T. Forsén, T. Tuomi, A. H. Rosengren and L. Groop, “Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables,” *The Lancet Diabetes & Endocrinology,* vol. 6, pp. 361-369, 5 2018.

[14] D. J. Bartholomew, F. Steele, J. Galbraith and I. Moustaki, Analysis of multivariate social science data, Chapman and Hall/CRC, 2008.

[15] B. Muthen and L. K. Muthen, “Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes,” *Alcoholism: Clinical and Experimental Research,* vol. 24, pp. 882-891, 6 2000.

[16] E. Herrett, A. M. Gallagher, K. Bhaskaran, H. Forbes, R. Mathur, T. Staa and L. Smeeth, “Data Resource Profile: Clinical Practice Research Datalink (CPRD),” *International Journal of Epidemiology,* vol. 44, pp. 827-836, 6 2015.

[17] R. Mathur, K. Bhaskaran, N. Chaturvedi, D. A. Leon, T. E. Grundy and L. Smeeth, “Completeness and usability of ethnicity data in UK-based primary care and hospital databases,” *Journal of Public Health,* vol. 36, pp. 684-692, 12 2013.

[18] D. T. Zemedikun, L. J. Gray, K. Khunti, M. J. Davies and N. N. Dhalwani, “Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data,” *Mayo Clinic Proceedings,* vol. 93, pp. 857-866, 7 2018.

[19] F. B. Larsen, M. H. Pedersen, K. Friis, C. Glümer and M. Lasgaard, “A Latent Class Analysis of Multimorbidity and the Relationship to Socio-Demographic Factors and Health-Related Quality of Life. A National Population-Based Study of 162,283 Danish Adults,” *PLOS ONE,* vol. 12, p. e0169426, 1 2017.

[20] M. Hall, T. B. Dondo, A. T. Yan, M. A. Mamas, A. D. Timmis, J. E. Deanfield, T. Jernberg, H. Hemingway, K. A. A. Fox and C. P. Gale, “Multimorbidity and survival for patients with acute myocardial infarction in England and Wales: Latent class analysis of a nationwide population-based cohort,” *PLOS Medicine,* vol. 15, p. e1002501, 3 2018.

[21] A. Marengoni, D. Rizzuto, H.-X. Wang, B. Winblad and L. Fratiglioni, “Patterns of Chronic Multimorbidity in the Elderly Population,” *Journal of the American Geriatrics Society,* vol. 57, pp. 225-230, 2 2009.

[22] A. Prados-Torres, B. Poblador-Plou, A. Calderón-Larrañaga, L. A. Gimeno-Feliu, F.

González-Rubio, A. Poncel-Falcó, A. Sicras-Mainar and J. T. Alcalá-Nalvaiz, "Multimorbidity Patterns in Primary Care: Interactions among Chronic Diseases Using Factor Analysis," *PLoS ONE*, vol. 7, p. e32190, 2 2012.

[23] N. D. a. I. S. a. J. d. G. a. M. L. a. K. M. a. J. R. a. M. v. Smeden, "Concerns about composite reference standards in diagnostic research," *BMJ*, p. j5779, 2018.

[24] K. L. Nylund, T. Asparouhov and B. O. Muthén, "Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 14, pp. 535-569, 10 2007.

[25] W. H. F. a. K. C. Bronk, "Conducting Confirmatory Latent Class Analysis Using Mplus," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 18, no. 1, pp. 132--151, 2011.

[26] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information theory*, vol. 37, no. 1, pp. 145--151, 1991.

[27] D. G. Altman, Practical statistics for medical research, CRC press, 1990.

[28] A. Déruaz-Luyet, A. A. NGoran, N. Senn, P. Bodenmann, J. Pasquier, D. Widmer, R. Tandjung, T. Rosemann, P. Frey, S. Streit, A. Zeller, D. M. Haller, S. Excoffier, B. Burnand and L. Herzig, "Multimorbidity and patterns of chronic conditions in a primary care population in Switzerland: a cross-sectional study," *BMJ Open*, vol. 7, p. e013664, 6 2017.

[29] S. M. Smith, E. Wallace, T. ODowd and M. Fortin, "Interventions for improving outcomes in patients with multimorbidity in primary care and community settings," *Cochrane Database of Systematic Reviews*, 3 2016.

[30] M. Sharpe, J. Walker, C. H. Hansen, P. Martin, S. Symeonides, C. Gourley, L. Wall, D. Weller and G. Murray, "Integrated collaborative care for comorbid major depression in patients with cancer (SMaRT Oncology-2): a multicentre randomised controlled effectiveness trial," *The Lancet*, vol. 384, pp. 1099-1108, 9 2014.

[31] P. Coventry, K. Lovell, C. Dickens, P. Bower, C. Chew-Graham, D. McElvenny, M. Hann, A. Cherrington, C. Garrett, C. J. Gibbons, C. Baguley, K. Roughley, I. Adeyemi, D. Reeves, W. Waheed and L. Gask, "Integrated primary care for patients with mental and physical multimorbidity: cluster randomised controlled trial of collaborative care for patients with depression comorbid with diabetes or cardiovascular disease," *BMJ*, vol. 350, pp. h638--h638, 2 2015.

[32] R. J. McManus, J. Mant, M. Franssen, A. Nickless, C. Schwartz, J. Hodgkinson, P. Bradburn, A. Farmer, S. Grant, S. M. Greenfield, C. Heneghan, S. Jowett, U. Martin, S. Milner, M. Monahan, S. Mort, E. Ogburn, R. Perera-Salazar, S. A. Shah, L.-M. Yu, L. Tarassenko, F. D. R. Hobbs, B. Bradley, C. Lovekin, D. Judge, L. Castello, M. Dawson, R. Brice, B. Dunbabin, S. Maslen, H. Rutter, M. Norris, L. French, M. Loynd, P. Whitbread, L. S. Ortaga, I. Noel, K. Madronal, J. Timmins, P. Bradburn, L. Hughes, B. Hinks, S. Bailey, S. Read, A. Weston, S. Spannuth, S. Maiden, M. Chermahini, A. McDonald, S. Rajan, S. Allen, B. Deboys, K. Fell, J. Johnson, H. Jung, R. Lister, R. Osborne, A. Secker, I. Qasim, K. William, A. Harris, S. Zhao, E. Butcher, P. Darbyshire, S. Joshi, J. Davies, C. Talbot, E. Hoverd, L. Field, T. Adcock, J. Rooney, N. Cooter, A. Butler, N. Allen, M. Abdul-Wahab, K. McNicholas, L. Peniket, K. Dodd, J. Mugurza, R. Baskerville, R. Syed, C. Bailey, J. Adams, P. Uglow, N. Townsend, A. Macleod, C. Hawkins, S. Behura, J. Crawshaw, R. Fox, W. Doski, M. Aylward, C. ACourt, D. Rapley, J. Walsh, P. Batra, A. Seoane, S. Mukherjee, J. Dixon, P. Arthur, K. Sutcliffe, C. Paschallides, R. Woof, P. Winfrey, M. Clark, R. Kamali, P. Thomas, D. Ebbs, L.

Mather, A. Beattie, K. Ladha, L. Smondulak, S. Jemahl, P. Hickson, L. Stevens, T. Crockett, D. Shukla, I. Binnian, P. Vinson, N. DeKare-Silver, R. Patel, I. Singh, L. Lumley, G. Williams, M. Webb, J. Bambrough, N. Shah, H. Dosanjh, F. Spannuth, C. Paul, J. Ganesegaram, L. Pike, V. Maheswaran, F. Paruk, S. Ford, V. Verma, K. Milne, F. Lockhat, J. Ferguson, A.-M. Quirk, H. Wilson, D. Copping, S. Bajallan, S. Tanvir, F. Khan, T. Alderson, A. Ali, R. Young, U. Chauhan, L. Crockett, L. McGovern, C. Cubitt, S. Weatherill, A. Tabassum, P. Saunders, N. Chauhan, S. Johnson, J. Walsh, I. Marok, R. Sharma, W. Lumb, J. Tweedale, I. Smith, L. Miller, T. Ahmed, M. Sanderson, C. Jones, P. Stokell, M. J. Edwards, A. Askey, J. Spencer, K. Morgan, K. Knox, R. Baker, C. Fisher, R. Halstead, N. Modha, D. Buckley, C. Stokell, J. G. McCabe, J. Taylor, H. Nutbeam, R. Smith, C. MacGregor, S. Davies, M. Lindsey, S. Cartwright, J. Whittle, J. Colclough, A. Crumbie, N. Thomas, V. Premchand, R. Hamid, Z. Ali, J. Ward, P. Pinney, S. Thurston and T. Banerjee, "Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial," *The Lancet*, vol. 391, pp. 949-959, 3 2018.

[33] R. J. McManus, J. Mant, M. S. Haque, E. P. Bray, S. Bryan, S. M. Greenfield, M. I. Jones, S. Jowett, P. Little, C. Penaloza, C. Schwartz, H. Shackleford, C. Shovelton, J. Varghese, B. Williams and F. D. R. Hobbs, "Effect of Self-monitoring and Medication Self-titration on Systolic Blood Pressure in Hypertensive Patients at High Risk of Cardiovascular Disease," *JAMA*, vol. 312, p. 799, 8 2014.

| | | | | | |
|---|---|---|---|---|---|
| 31.9 | 22.8 | 20.3 | 18.3 | 6.7 | **% patients** |
| 49.6 | 45.5 | 40.7 | 37.1 | 63.3 | **% greater deprivation** |
| 45.8 | 27.3 | 29.1 | 28.4 | 76.2 | **% current smokers** |
| 14.5 | 11.9 | 11.3 | 10.6 | 10.7 | **Avg. # GP consultations** |
| 0.4 | 0.6 | 0.4 | 0.4 | 0.4 | **Avg. # hospital spells** |
| 2.4 | 2.5 | 1.9 | 1.2 | 1.3 | **Avg. # repeat prescriptions** |
| 0.9 | 1.0 | 0.2 | 0.2 | 1.8 | **% 2-year mortality** |
| 1.8 | 2.7 | 0.6 | 0.4 | 3.9 | **% 5-year mortality** |
| **DEP,ANX,PNC** | **PNC,HL,HYP** | **AST,IBS,DEP** | **IBS,DEP,HL** | **PSM,ALC,DEP** | |
| 4% | 23% | 7% | 6% | 2% | Hypertension |
| 31% | 36% | 17% | 15% | 19% | Painful condition |
| 11% | 30% | 18% | 21% | 14% | Hearing loss |
| 100% | 18% | 20% | 29% | 24% | Depression |
| 15% | 2% | 26% | 100% | 7% | Irritable bowel syndrome |
| 15% | 5% | 100% | 0% | 11% | Asthma |
| 2% | 19% | 3% | 3% | 2% | Diabetes |
| 0% | 3% | 0% | 0% | 1% | Coronary heart disease |
| 5% | 16% | 6% | 8% | 2% | Thyroid disorders |
| 41% | 9% | 5% | 8% | 15% | Anxiety |
| 0% | 1% | 0% | 1% | 1% | Diverticular disease of intestine |
| 0% | 1% | 1% | 1% | 1% | Prostate disorder |
| 0% | 1% | 0% | 0% | 0% | Chronic kidney disease |
| 4% | 11% | 8% | 10% | 1% | Chronic sinusitis |
| 0% | 2% | 0% | 0% | 0% | Stroke & transient ischaemic attack |
| 0% | 1% | 1% | 0% | 1% | COPD |
| 0% | 1% | 0% | 0% | 0% | Atrial fibrillation |
| 2% | 7% | 2% | 3% | 2% | Rheumatoid arthritis |
| 3% | 5% | 2% | 4% | 1% | Constipation |
| 1% | 3% | 1% | 1% | 1% | Cancer |
| 1% | 2% | 1% | 1% | 3% | Peptic ulcer disease |
| 6% | 4% | 2% | 2% | 42% | Alcohol problems |
| 8% | 1% | 3% | 2% | 75% | Psychoactive substance misuse |
| 1% | 6% | 1% | 2% | 1% | Blindness and low vision |
| 2% | 5% | 9% | 2% | 2% | Psoriasis or eczema |
| 0% | 1% | 0% | 0% | 0% | Heart failure |
| 6% | 5% | 1% | 2% | 17% | Schizophrenia |
| 1% | 4% | 2% | 6% | 1% | Inflammatory bowel disease |
| 0% | 0% | 0% | 0% | 0% | Peripheral vascular disease |
| 0% | 0% | 0% | 0% | 0% | Dementia |
| 5% | 4% | 2% | 5% | 5% | Anorexia or bulimia |
| 1% | 2% | 0% | 1% | 6% | Chronic liver disease |
| 12% | 0% | 0% | 0% | 0% | Epilepsy |
| 2% | 2% | 1% | 2% | 0% | Migraine |
| 1% | 10% | 1% | 1% | 0% | Learning disability |
| 0% | 1% | 1% | 0% | 0% | Bronchiectasis |
| 1% | 2% | 0% | 1% | 0% | Multiple sclerosis |
| 0% | 0% | 0% | 0% | 0% | Parkinson's diseases |

| HYP,DIA,PNC | IBS,HL,PNC | DEP,PNC,ANX | AST,COPD,PNC | ALC,PSM,PNC | |
|---|---|---|---|---|---|
| 37.2 | 24.3 | 21.7 | 12.3 | 4.4 | **% patients** |
| 37.7 | 29.5 | 45.9 | 35.1 | 57.2 | **% greater deprivation** |
| 20.5 | 20.5 | 35.1 | 20.2 | 63.3 | **% current smokers** |
| 11.5 | 10.5 | 16.7 | 12.6 | 10.7 | **Avg. # GP consultations** |
| 0.5 | 0.5 | 0.6 | 0.4 | 0.5 | **Avg. # hospital spells** |
| 4.1 | 2.0 | 5.1 | 3.4 | 2.4 | **Avg. # repeat prescriptions** |
| 1.6 | 1.3 | 2.4 | 1.0 | 4.5 | **% 2-year mortality** |
| 4.4 | 3.0 | 5.8 | 2.7 | 12.5 | **% 5-year mortality** |
| 76% | 18% | 22% | 26% | 19% | Hypertension |
| 25% | 28% | 53% | 24% | 28% | Painful condition |
| 15% | 29% | 11% | 15% | 13% | Hearing loss |
| 13% | 20% | 93% | 15% | 26% | Depression |
| 8% | 40% | 19% | 18% | 7% | Irritable bowel syndrome |
| 6% | 3% | 16% | 100% | 12% | Asthma |
| 37% | 3% | 11% | 8% | 6% | Diabetes |
| 17% | 4% | 5% | 3% | 5% | Coronary heart disease |
| 11% | 20% | 11% | 10% | 3% | Thyroid disorders |
| 2% | 5% | 31% | 3% | 11% | Anxiety |
| 4% | 9% | 4% | 4% | 2% | Diverticular disease of intestine |
| 5% | 7% | 2% | 4% | 3% | Prostate disorder |
| 5% | 1% | 2% | 1% | 0% | Chronic kidney disease |
| 5% | 16% | 4% | 12% | 2% | Chronic sinusitis |
| 6% | 1% | 3% | 1% | 3% | Stroke & transient ischaemic attack |
| 3% | 3% | 5% | 16% | 11% | COPD |
| 5% | 1% | 1% | 1% | 2% | Atrial fibrillation |
| 4% | 9% | 5% | 3% | 2% | Rheumatoid arthritis |
| 2% | 5% | 10% | 2% | 2% | Constipation |
| 4% | 6% | 3% | 3% | 2% | Cancer |
| 4% | 6% | 3% | 2% | 8% | Peptic ulcer disease |
| 3% | 1% | 6% | 1% | 62% | Alcohol problems |
| 0% | 0% | 4% | 0% | 42% | Psychoactive substance misuse |
| 2% | 3% | 1% | 1% | 2% | Blindness and low vision |
| 3% | 3% | 3% | 5% | 3% | Psoriasis or eczema |
| 3% | 0% | 0% | 0% | 0% | Heart failure |
| 2% | 3% | 8% | 1% | 9% | Schizophrenia |
| 1% | 5% | 1% | 2% | 0% | Inflammatory bowel disease |
| 2% | 0% | 1% | 0% | 1% | Peripheral vascular disease |
| 0% | 0% | 0% | 0% | 0% | Dementia |
| 0% | 2% | 2% | 0% | 3% | Anorexia or bulimia |
| 1% | 1% | 1% | 0% | 14% | Chronic liver disease |
| 0% | 0% | 12% | 0% | 0% | Epilepsy |
| 0% | 4% | 3% | 1% | 0% | Migraine |
| 1% | 2% | 1% | 1% | 1% | Learning disability |
| 0% | 1% | 1% | 3% | 1% | Bronchiectasis |
| 0% | 2% | 2% | 0% | 0% | Multiple sclerosis |
| 0% | 0% | 0% | 0% | 0% | Parkinson's diseases |

| HYP,DIA,CKD | HL,PSD,IBS | DEP,PNC,ANX | CHD,AF,DIA | COPD,AST,PNC | PNC,CHD,DEP | |
|---|---|---|---|---|---|---|
| 41.0 | 21.6 | 13.8 | 10.8 | 8.2 | 4.7 | **% patients** |
| 30.4 | 25.3 | 32.5 | 33.3 | 40.0 | 43.1 | **% greater deprivation** |
| 10.1 | 9.5 | 15.5 | 12.0 | 24.5 | 16.5 | **% current smokers** |
| 12.4 | 13.2 | 17.3 | 16.8 | 15.7 | 26.3 | **Avg. # GP consultations** |
| 0.6 | 0.7 | 0.8 | 1.1 | 0.8 | 1.6 | **Avg. # hospital spells** |
| 4.5 | 3.3 | 5.9 | 5.9 | 5.5 | 10.7 | **Avg. # repeat prescriptions** |
| 4.7 | 4.4 | 8.4 | 11.3 | 9.2 | 16.2 | **% 2-year mortality** |
| 13.2 | 11.1 | 20.9 | 28.8 | 25.5 | 39.2 | **% 5-year mortality** |
| 100% | 34% | 43% | 61% | 44% | 75% | Hypertension |
| 27% | 30% | 56% | 22% | 33% | 81% | Painful condition |
| 19% | 40% | 17% | 22% | 21% | 29% | Hearing loss |
| 6% | 4% | 56% | 6% | 12% | 45% | Depression |
| 6% | 20% | 18% | 2% | 8% | 18% | Irritable bowel syndrome |
| 7% | 8% | 8% | 4% | 49% | 24% | Asthma |
| 31% | 10% | 12% | 32% | 11% | 36% | Diabetes |
| 15% | 14% | 8% | 54% | 12% | 53% | Coronary heart disease |
| 13% | 16% | 17% | 11% | 9% | 16% | Thyroid disorders |
| 1% | 2% | 23% | 1% | 4% | 20% | Anxiety |
| 8% | 19% | 14% | 7% | 10% | 28% | Diverticular disease of intestine |
| 8% | 21% | 5% | 16% | 10% | 14% | Prostate disorder |
| 16% | 6% | 9% | 20% | 5% | 27% | Chronic kidney disease |
| 3% | 10% | 4% | 1% | 7% | 7% | Chronic sinusitis |
| 9% | 5% | 9% | 18% | 6% | 27% | Stroke & transient ischaemic attack |
| 3% | 2% | 6% | 9% | 57% | 25% | COPD |
| 6% | 6% | 4% | 29% | 6% | 21% | Atrial fibrillation |
| 6% | 10% | 8% | 5% | 8% | 14% | Rheumatoid arthritis |
| 3% | 7% | 19% | 3% | 3% | 34% | Constipation |
| 6% | 11% | 6% | 5% | 5% | 8% | Cancer |
| 3% | 8% | 5% | 7% | 7% | 10% | Peptic ulcer disease |
| 1% | 0% | 3% | 1% | 5% | 3% | Alcohol problems |
| 0% | 0% | 1% | 0% | 3% | 2% | Psychoactive substance misuse |
| 2% | 3% | 2% | 3% | 2% | 7% | Blindness and low vision |
| 2% | 2% | 2% | 2% | 4% | 7% | Psoriasis or eczema |
| 0% | 0% | 0% | 19% | 2% | 19% | Heart failure |
| 1% | 1% | 7% | 1% | 1% | 3% | Schizophrenia |
| 1% | 3% | 2% | 1% | 2% | 2% | Inflammatory bowel disease |
| 3% | 1% | 1% | 12% | 4% | 12% | Peripheral vascular disease |
| 1% | 1% | 8% | 3% | 1% | 5% | Dementia |
| 0% | 0% | 1% | 0% | 1% | 1% | Anorexia or bulimia |
| 1% | 1% | 1% | 1% | 1% | 1% | Chronic liver disease |
| 0% | 0% | 2% | 0% | 0% | 2% | Epilepsy |
| 0% | 1% | 2% | 0% | 0% | 0% | Migraine |
| 0% | 0% | 1% | 0% | 0% | 0% | Learning disability |
| 0% | 1% | 1% | 1% | 9% | 3% | Bronchiectasis |
| 0% | 1% | 1% | 0% | 0% | 0% | Multiple sclerosis |
| 0% | 1% | 3% | 1% | 0% | 2% | Parkinson's diseases |

| | HYP,HL,DIA | PNC,DEP,CSP | HF,CHD,AF | AST,COPD,PNC |
|---|---|---|---|---|
| % patients | 57.9 | 23.3 | 10.6 | 8.3 |
| % greater deprivation | 29.6 | 30.0 | 30.5 | 30.2 |
| % current smokers | 5.0 | 5.4 | 3.8 | 8.2 |
| Avg. # GP consultations | 13.0 | 17.3 | 21.9 | 19.6 |
| Avg. # hospital spells | 0.8 | 0.8 | 1.5 | 1.1 |
| Avg. # repeat prescriptions | 4.1 | 6.7 | 8.0 | 6.9 |
| % 2-year mortality | 20.9 | 31.1 | 37.7 | 28.0 |
| % 5-year mortality | 49.5 | 62.9 | 70.8 | 56.5 |
| Hypertension | 72% | 57% | 70% | 57% |
| Painful condition | 23% | 64% | 54% | 44% |
| Hearing loss | 39% | 36% | 41% | 42% |
| Depression | 3% | 41% | 26% | 12% |
| Irritable bowel syndrome | 4% | 9% | 9% | 7% |
| Asthma | 2% | 4% | 9% | 48% |
| Diabetes | 18% | 11% | 22% | 11% |
| Coronary heart disease | 25% | 17% | 61% | 22% |
| Thyroid disorders | 14% | 15% | 20% | 12% |
| Anxiety | 2% | 20% | 12% | 8% |
| Diverticular disease of intestine | 13% | 20% | 22% | 21% |
| Prostate disorder | 13% | 4% | 16% | 17% |
| Chronic kidney disease | 24% | 20% | 39% | 23% |
| Chronic sinusitis | 2% | 2% | 3% | 6% |
| Stroke & transient ischaemic attack | 18% | 20% | 31% | 12% |
| COPD | 4% | 3% | 11% | 48% |
| Atrial fibrillation | 18% | 8% | 53% | 12% |
| Rheumatoid arthritis | 9% | 10% | 11% | 13% |
| Constipation | 5% | 42% | 28% | 20% |
| Cancer | 7% | 3% | 6% | 9% |
| Peptic ulcer disease | 6% | 6% | 10% | 7% |
| Alcohol problems | 0% | 1% | 0% | 0% |
| Psychoactive substance misuse | 0% | 0% | 0% | 0% |
| Blindness and low vision | 8% | 9% | 10% | 7% |
| Psoriasis or eczema | 2% | 3% | 4% | 3% |
| Heart failure | 5% | 2% | 49% | 8% |
| Schizophrenia | 0% | 3% | 1% | 0% |
| Inflammatory bowel disease | 1% | 1% | 1% | 2% |
| Peripheral vascular disease | 5% | 2% | 13% | 6% |
| Dementia | 7% | 24% | 11% | 4% |
| Anorexia or bulimia | 0% | 1% | 1% | 0% |
| Chronic liver disease | 0% | 0% | 0% | 1% |
| Epilepsy | 0% | 0% | 0% | 0% |
| Migraine | 0% | 0% | 0% | 0% |
| Learning disability | 0% | 0% | 0% | 0% |
| Bronchiectasis | 0% | 1% | 1% | 8% |
| Multiple sclerosis | 0% | 0% | 0% | 0% |
| Parkinson's diseases | 1% | 3% | 2% | 1% |