REAL Centre

Quantifying health inequalities in England

Appendix

August 2022

Toby Watt,
Ann Raymond,
Laurie Rachet-Jacquet

The Health Foundation

Quantifying health inequalities in England: Appendix

Data description

Clinical Practice Research Datalink (CPRD) Aurum is a database containing pseudonymised primary care records from patients registered at 1,442 primary care practices in England. We have used CPRD Aurum records linked to secondary care through the Hospital Episode Statistics (HES), the Index of Multiple Deprivation (IMD) of the patients' local area and mortality data collected by the Office for National Statistics (ONS). Using this unique data set we can explore diagnosed illness across the English population by age, sex, ethnicity, region, and socioeconomic status through each patient's local area IMD.

Data permissions

The data is provided by patients and collected by the NHS as part of their care and support. Regulatory approvals to use CPRD data for the current project were granted by the CPRD Independent Scientific Advisory Committee (ISAC protocol number 20-000096). The Health Foundation's REAL Centre is working in collaboration with the University of Liverpool to develop a microsimulation model of health and care demand. Part of our analysis is to interrogate incidence, prevalence, case fatality, and mortality rates for common non-communicable diseases, stratified by year, age, sex, region, socioeconomics, and ethnicity. These results have important policy implications for the government's levelling up agenda.

Cleaning and analysis

For our randomly selected sample of 2 million patients who were registered at a CPRD primary care practice between April 2008 and March 2020, we can analyse complete patient histories. This allows us to learn about non-communicable long-term illness for patients that are diagnosed in primary and/or secondary care settings.

We included all patients who were registered at practices covered by the CPRD Aurum database during the period from 1 April 2008 to 31 March 2020. After excluding patients with incomplete or incorrect linkage and missing patient characteristics, our final sample included 1,756,084 patients.

Cambridge Multimorbidity Score

Long-term illness can affect multiple biological systems at the same time, and it is common for people to be diagnosed with more than one or even many conditions at once. Comparing two people with different conditions is therefore challenging because different conditions have different implications for patients' quality of life and their health and care needs. Improved measures of multimorbidity are needed for conducting research, planning services and allocating resources.

One measure of multimorbidity is the Cambridge Multimorbidity Score (CMS) developed by Payne et al.[1] This gives a 'severity score' for each of 20 non-communicable illnesses based on patient's needs within primary care, their likelihood of experiencing unplanned admissions and their mortality rates using linked primary and secondary data from 2015.

88% of the burden of disease in England is a result of non-communicable illness.[2] The 20 conditions contained within the Cambridge Multimorbidity Score are responsible for 66% of the total burden of non-communicable disease in England. These conditions and their weights (from highest to lowest) are presented in Table A1.

Table A1: Cambridge Multimorbidity Score weights by condition

Condition	**Weight**
Dementia	2.50
Cancer	1.53
COPD	1.46
Atrial fibrillation	1.34
Heart failure	1.18
Constipation	1.12
Epilepsy	0.92
Chronic pain	0.92
Stroke / transient ischaemic attack (TIA)	0.80
Diabetes (type I or II)	0.75
Alcohol problems	0.65
Psychosis/bipolar disorder	0.64
Chronic kidney disease	0.53
Anxiety/depression	0.50
Coronary heart disease	0.49
Connective tissue disorders	0.43
Irritable bowel syndrome	0.21
Asthma	0.19
Hearing loss	0.09
Hypertension	0.08

Using primary and secondary care patient records we define patients that have been diagnosed with these conditions and, using these weights, calculate their Cambridge Multimorbidity Score. Our definitions of diagnosis for these conditions are outlined in the next section.

Definition of long-term illness

We defined a diagnosis of one of these conditions by adapting the approach of Barnett et al. (2012)[3] and Head et al (2021).[4,5]

For most conditions, we use disease definitions that are commonly used in epidemiological research. We have provided a more detailed definition for the following conditions below:

- Alcohol problems include conditions associated with harmful levels of alcohol consumption including alcoholic liver cirrhosis, alcoholic hepatitis and mental and behavioural disorders associated with alcoholism.[6]
- Chronic kidney disease includes chronic kidney disease stages 3 and above and end stage renal disease.[7]
- Connective tissue disorders include conditions associated with rheumatoid arthritis,[8] lupus erythematosus,[9] polymyalgia rheumatica[10] and scleroderma.[11]
- Psychosis/bipolar disorder includes conditions associated with schizophrenia, paranoia and bipolar affective disorders and manias.[12,13]

A patient is diagnosed with a condition at the time of the first instance of a diagnosis across CPRD or HES inpatient data. This is the case for most conditions. For chronic pain, diagnosis is based on having at least four prescriptions of analgesics or epilepsy drugs (conditional on not having been diagnosed with epilepsy) in the span of a year in CPRD. We use a similar rule for constipation where diagnosis is based on having at least four prescriptions of laxatives within a year in CPRD.

Regarding the duration of these conditions, 15 of them are modelled as chronic, life-long conditions. These are alcohol problems, atrial fibrillation, connective tissue disorders, coronary heart disease, chronic kidney disease, chronic obstructive pulmonary disease (COPD), dementia, diabetes (type 1 or 2), epilepsy, hearing loss, heart failure, hypertension, irritable bowel syndrome, psychosis/bipolar disorder, and stroke/transient ischaemic attack.

In line with what has been done for the Cambridge Multimorbidity Score, we model potential remission for five conditions, ie asthma, anxiety/depression, cancer, chronic pain and constipation. We consider the condition to have been resolved if the patient has had no new diagnoses or related drug prescriptions for a year (for cancers, we assume that remission occurs 5 years after the first diagnosis).

We also treat any diagnosis that takes place within 12 months of a patient's registration with a new practice as an existing disease (prevalence) and not a new diagnosis (incidence).

Patient characteristics

We are able to define the patient's Index of Multiple Deprivation (IMD) of their local area, age, gender, ethnicity, and region for every patient.

Index of multiple deprivation

Although not disclosed to the researcher, CPRD have access to the patient's local area postcode. They have then linked the patient to the Index of Multiple Deprivation (IMD), a ranking of the Lower Super Output Areas in England by decile of deprivation. Because of data availability we use the 2015 IMD area classifications.

Ethnicity

Patient ethnicity is reported in both CPRD and HES. We use the 11-group classification that is commonly used by HES. The ethnic categories are Bangladeshi, Black African, Black Caribbean, Black Other, Chinese, Indian, Mixed, Other, Other Asian, Pakistani and white. The white ethnic group includes non-British white groups such as Gypsy, Roma or Irish travellers. When multiple ethnicities are reported for the same individual, we assign patient ethnicity based on the most commonly reported ethnicity across all datasets. In case of ties, we assign patient ethnicity based on the most recent ethnicity reported across CPRD and HES.

Region

Region pertains to the region in England of the GP practice with which the patient is registered.

---

[1] www.cmaj.ca/content/192/5/E107.long

[2] https://vizhub.healthdata.org/gbd-compare/

[3] www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)60240-2/fulltext

[4] www.thelancet.com/pdfs/journals/lanhl/PIIS2666-7568(21)00146-X.pdf

[5] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists

[6] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Alcohol%20Misuse.csv

[7] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Chronic%20Kidney%20Disease.csv

[8] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Rheumatoid%20Arthritis.csv

[9] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Lupus%20Erythematosus.csv

[10] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Polymyalgia%20Rheumatica.csv

[11] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Scleroderma.csv

[12] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Schizophrenia.csv

[13] https://github.com/annalhead/CPRD_multimorbidity_codelists/blob/main/codelists/Bipolar%20affective%20disorder%20and%20mania.csv