

Use of primary care during the COVID-19 pandemic

Supporting information and data description

Toby Watt, Senior Economist, REAL Centre

September 2020

Introduction

Health care in England has undergone monumental changes in response to the COVID-19 pandemic. This has presented in: **a large number of reactive health policy changes, dramatic drops in A&E activity, dramatic increases in wait times, changes in public attitudes and willingness to visit NHS services and cancellation of treatments.**

Researchers and policy makers have been monitoring progress of mortality rates, COVID-19 case rates and secondary care activity made available by the Office for National Statistics (ONS), NHS Digital / X and other government sources.

Primary care accounted for less than 10% of Department of Health and Social Care resource spending in 2017, however it is responsible for around 80% of patient contact with the health system. It is the entry point for all non-emergency care. GPs act as the face of the NHS for their patients, helping with health care prevention, reducing unnecessary secondary care visits, and helping with mental and physical health issues.

However, the lack of available data means health and health policy researchers know very little about primary care and its outcomes. For example, we can't observe referral rates; how appointment frequency varies by age and sex; the extent of variation by practice; or how much more primary care activity occurs for older patients or patients with pre-existing conditions. Filling these gaps in knowledge is key to effective primary and wider care policy.

The Health Foundation's REAL Centre has used a patient-level primary care data set to explore the changes that lockdown and the wider pandemic have brought about. In this

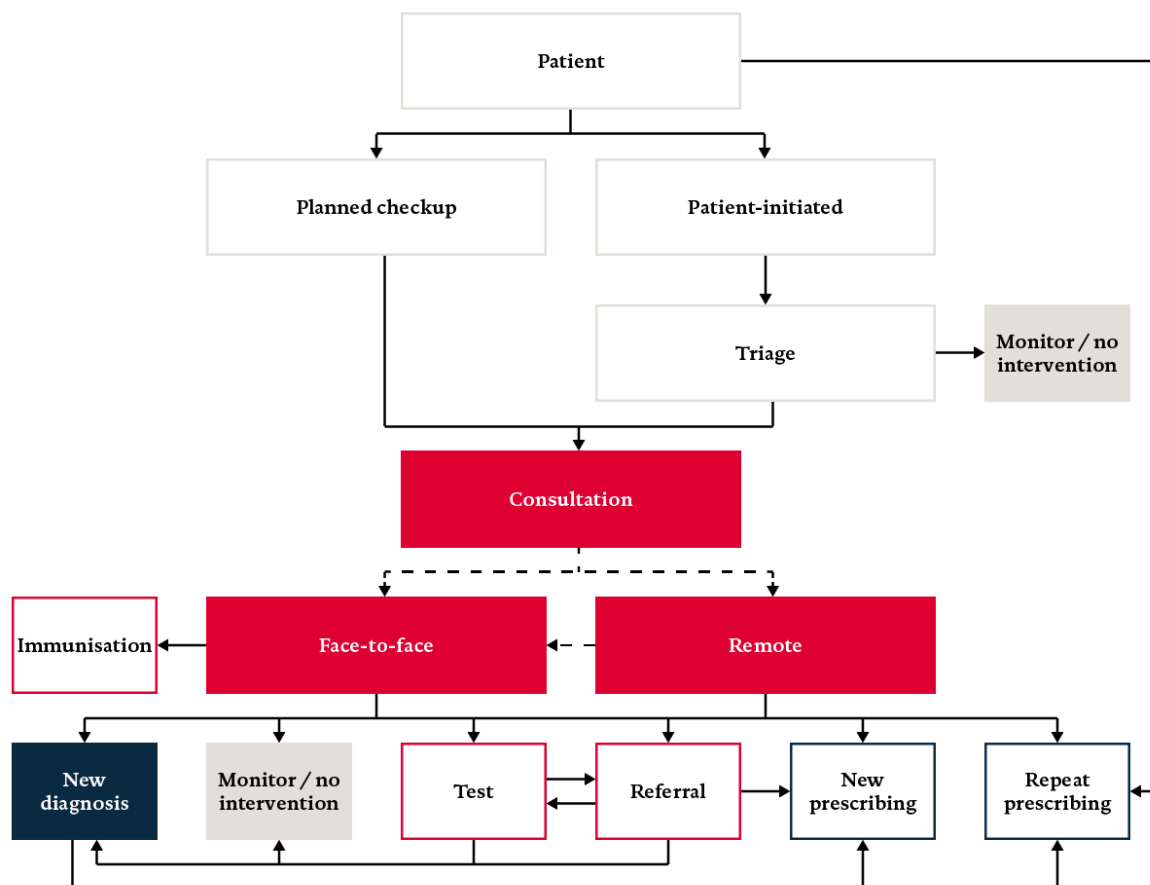
document we describe our data cleaning that generate the high-level trend analysis in primary care activity. We make our high-level data available for analysis in the hope it will be used to improve policy decisions in the UK recovery from COVID-19.

How does primary care work and what can we observe?

GPs are the gatekeepers to the rest of the health system. The key roles of primary care are both to manage access to acute secondary care, and to help patients manage and monitor long-term conditions. At a check-up or an appointment (consultation) made by the patient, a GP will – depending on the needs of the patient – organise testing, therapy (including prescribing) or a referral to treatment / specialist examination in secondary care. Consultations can be conducted remotely or face-to-face, and the resulting activities occur in any combination.

This route through the various potential stages of primary care creates complex administrative data. This is visually represented in figure 1 (below).

Figure 1: Typical patient flows through primary care during COVID-19



This illustration shows the many ways in which the outcomes associated with a primary care appointment can vary: from the monitoring of a prospective new illness or management of an existing illness, through to prescribing or a referral to secondary care.

About the Clinical Practice Research Datalink (CPRD) dataset

The data generated by the primary care process are necessarily complex, driven by the multiple roles played by primary care practitioners. The systematic problem with data collection and analysis in primary care is that practices are, for the most part, independent businesses. There is therefore no administrative data set that reaches all practices.

The Health Foundation's REAL Centre has access to CPRD, a research service which collects these complicated, patient-level data from a subset of practices in the UK. There are two databases provided by CPRD – GOLD and Aurum. These data record consultations, referrals, tests, prescribing and immunisations.

Our analysis is of CPRD's GOLD database, the smaller of the two databases available in terms of practices that are currently active and reporting in England over the summer. We use these data because our team at the REAL Centre has prior experience of working with this dataset and the [published code lists](#) allow us to perform analysis of patients with pre-existing health conditions.*

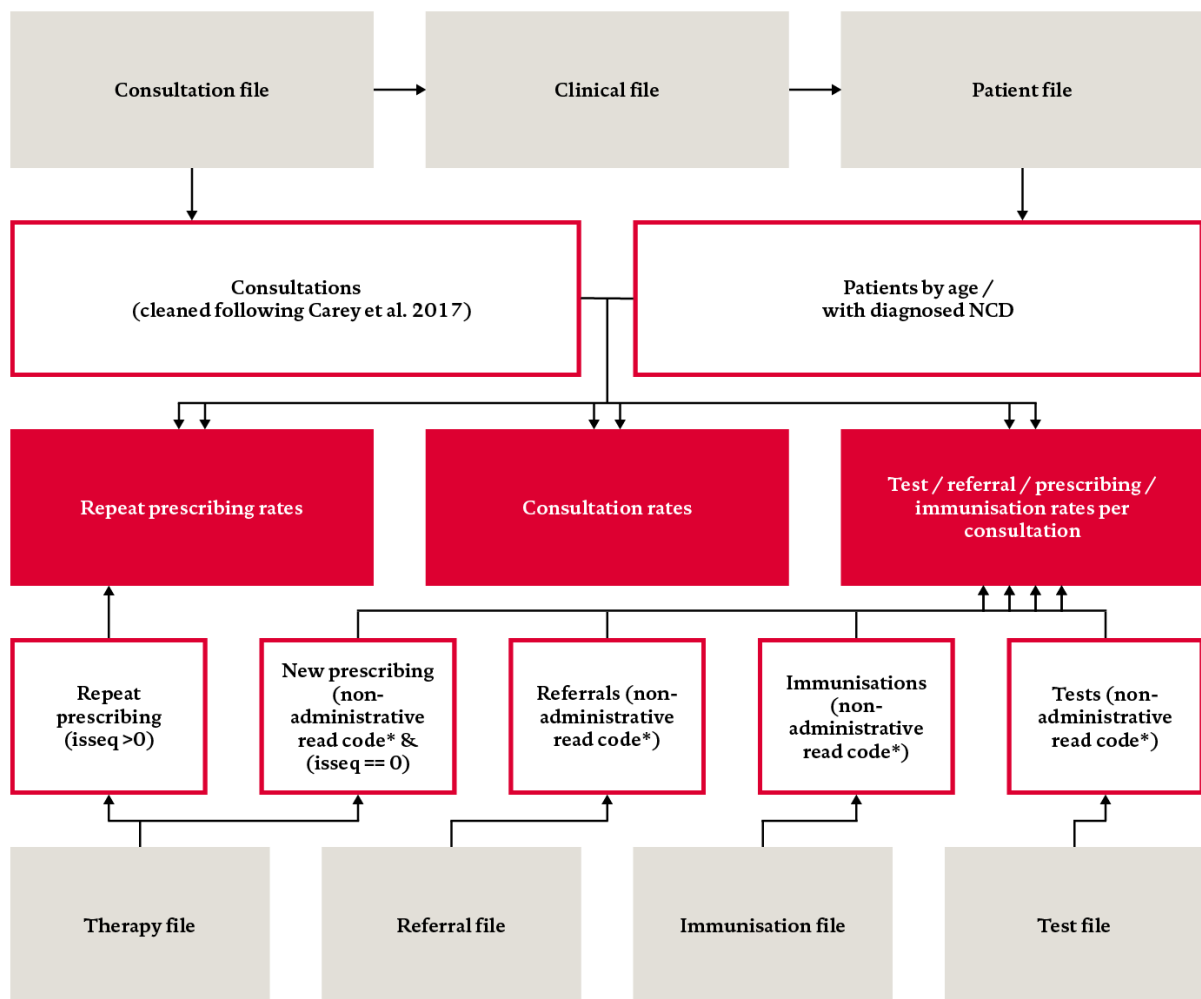
The practices included in the GOLD database have reported their data – from ~500,000 patients – up to the end of June 2020. This gives us around 10 weeks of data since England entered lockdown on 23 March 2020.

We present, for each week, consultation rates for the population that are registered at those practices at that time. The primary care data from CPRD is recorded first in consultations, with the associated clinical diagnosis, tests, therapy (prescribing) and referrals for each patient. These second stage data can occur several days after the original appointment. They are linked by a patient number and a consultation number that is unique to each appointment. We also calculate the *repeat* prescribing rates for patients, which do not require a full consultation to be fulfilled.

The figure below depicts the data setup used by CPRD, and the cleaning taken to arrive at our summary activity and rates of primary care outcomes. We can link outcomes such as tests and referrals to specific consultations using pseudonymised patient numbers and consultation numbers. This gives us a new insight into the typical outcomes associated with primary care and the treatment / management pathways that occur.

* The GOLD database is collected using Vision® software which is one of the last remaining sources that still makes use of [Read Codes](#) to identify primary care activity. Others make use of the [SNOMED](#) coding system.

Figure 2: CPRD Data format



Non-administrative read codes are defined in figure 39 of Carey et al (2017) presented in Annex 1.*

There are six data files in the CPRD database.† Each is cleaned in the following way:

- **Consultation file:** These data list all entries into the patient data base performed by a member of the practice staff. The staff role and the 'type' of consultation is recorded in these data, which we restrict to a definition of consultation defined by Carey et al (2017).†

* Carey IM, Hosking FJ, Harris T, DeWilde S, Beighton C, Cook DG. An evaluation of the effectiveness of annual health checks and quality of health care for adults with intellectual disability: an observational study using a primary care database. *Health Services and Delivery Research*. 2017; 5(25).

† https://cprdcw.cprd.com/_docs/CPRD_GOLD_Full_Data_Specification_v2.0.pdf

- **Clinical file:** These data correspond to a consultation. They contain diagnosis codes filled in by the practice staff. These data, using the definitions outlined by the Primary Care Unit at the School of Clinical Medicine, University of Cambridge,^{*} allow us to categorise patients by pre-existing condition. Our list of conditions used in this analysis are:
 - Anxiety, Asthma, Cancer, Coronary heart disease (CHD), Chronic obstructive pulmonary disease (COPD), Depression, Diabetes, Heart failure and Stroke
- **Patient file:** These data contain patient demographic information to split patients by age category. We can categorise patients by those with a long term or chronic illness using the diagnosis codes (Read Codes[†]) from the clinical file. This classification is done using patient clinical history dating back to 1980 (where available).
- **Test file:** These data contain records of test data on the GP system. These data are coded using a Read Code, chosen by the GP, which will generally identify the type of test used.
- **Referral file:** These data contain referral details recorded on the GP system. These files contain information involving patient referrals to external care centres (normally to secondary care locations such as hospitals for inpatient or outpatient care) and include specialty and referral type.
- **Therapy file:** These data contain details of all prescriptions (for drugs and appliances) issued by the GP system. Patients may have more than one row of data.
- **Immunisation file:** These data contain details of immunisation records on the GP system.

The purpose of this analysis is to help policy makers understand the high-level trends in primary care activity. In order to clean the data, we take the following steps:

1. Establish whether patients in our sample have been diagnosed with a pre-existing condition using the consultation file, in combination with the **code lists** published by the Primary Care Unit at the School of Clinical Medicine, University of Cambridge.
2. Categorise all registered patients from the sample by age and whether they have one of the long-term conditions listed. The analysis is presented for all patients, by age group and separately for people with each long-term condition (not by age). Patients can have multiple long-term conditions and can therefore be in multiple groups: the sum of the patient count from all of the long-term conditions will be greater than the total number of patients.

^{*} https://www.phpc.cam.ac.uk/pcu/research/research-groups/crmh/cprd_cam/codelist/
[†] [https://digital.nhs.uk/services/terminology-and-classifications/read-codes#:~:text=Read%20Codes%20are%20a%20coded,3%20\(CTV3%20or%20v3\).](https://digital.nhs.uk/services/terminology-and-classifications/read-codes#:~:text=Read%20Codes%20are%20a%20coded,3%20(CTV3%20or%20v3).)

3. Restrict the consultation file to reflect the Carey et al (2017)^{*} definitions of consultations. These are based on the staff code and the consultation type. We do include consultations that fit both the definitions of consultation and staff type, but do not have a clinical, referral, immunisation or test file that matches to it.
4. Restrict the test, immunisation and referral files to exclude those with administrative Read Codes and those without Read Codes. These data are then matched to a consultation number (unique when combined with patient number); we remove any data that does not match to a cleaned consultation. Referrals are then also split by urgency for all patients (not by age or non-communicable disease (NCD) due to small numbers). We limit the count to be equal to 1 if there is more than one test / referral / immunisation attached to a consultation. Our analysis therefore reflects the 'presence' of any referral / test / immunisation during a consultation, rather than the count.
5. We then use the consultation type and the Read Codes from the clinical and referral files to identify telephone or non-face-to-face consultations (see [Annex 1](#), Carey et al (2017)[†]).
6. Separate the therapy file into (a) repeat prescriptions (where the variable `issseq > 0`, which indicates a sequence of prescriptions) and (b) new consultation prescription.
 - a. We limit repeat prescriptions to a single indicator variable per consultation, but we do not limit the repeat prescription based on the consultation number – we match repeat prescriptions to the patient on a weekly basis. The 'consultation type' for these repeat prescriptions can therefore take any value, including 'other' or 'administrative'.
 - b. We attach the non-repeat prescriptions (`issseq == 0`) to the cleaned consultation file for new prescriptions from a consultation. We remove any prescriptions that do not match.

We then limit the prescriptions to be a binary indicator variable, showing the 'presence' of a repeat prescription per consultation, rather than the count of different drugs prescribed. Prescribing data is extremely variable from drug to drug and from condition to condition. Volumes and potency are difficult to manage, and multiple drugs could be prescribed to treat the same condition. This is a limiting factor of our analysis, in an area that is deserving of further work, especially in relation to the changes in patient and practice behaviour brought about by lockdown.

^{*} Carey IM, Hosking FJ, Harris T, DeWilde S, Beighton C, Cook DG. An evaluation of the effectiveness of annual health checks and quality of health care for adults with intellectual disability: an observational study using a primary care database. *Health Services and Delivery Research*. 2017; 5(25).

[†] Carey IM, Hosking FJ, Harris T, DeWilde S, Beighton C, Cook DG. An evaluation of the effectiveness of annual health checks and quality of health care for adults with intellectual disability: an observational study using a primary care database. *Health Services and Delivery Research*. 2017; 5(25).

7. We then combine the cleaned consultation and repeat prescribing data with the cleaned patient file, and aggregate weekly. We calculate the crude rates of consultation and repeat prescribing per person per year (weekly rate x 52). We can then calculate these rates by age of the patient and categorise the rates for patients with pre-existing conditions.
8. We match the referral, test, immunisation and new prescribing data with the cleaned consultation and patient data to calculate the rates per person per year, as well as the rates per consultation.
9. We also offer the percentage changes in 2020 when compared to the average of previous years in our presentation. We use data from 2016–19 to calculate the average weekly activity rates, and then calculate the percentage difference between the respective week in 2020. These do not take into account time trends in activity rates from 2016–20.
10. We also present a breakdown of primary care outcomes (immunisations, referrals, new prescribing and tests) in terms of the fall in consultations. For example, we present the actual levels of referrals per person, and we estimate the levels that we would have expected conditional on the rate of referrals per consultation remaining at pre-lockdown levels. We do this by applying the average rate of referral per consultation in 2016–19, to the rate of consultation in 2020. (Figures 4.4, 6.1, 6.2 and 8.2 in the [analysis](#)).

These data are crude rates, they are not adjusted for age and they represent the population sample that is available to us through CPRD GOLD. We do not attempt to adjust these figures to be nationally representative, because they are based on a small number of practices. The figures could be adjusted to match the demographics of the English population, however there may be significant differences in the underlying health of the population in the sample when compared with the wider population. The data are available for download as part of the [analysis](#).

The data are presented to the end of June 2020. However, the final weeks show a drop off in patients due to the number of practices that report data all through May. We therefore produce charts up to and including week commencing 21 June 2020 or week 25 2020, excluding the last week of June (week 26).

In the tab 'Raw Data', we present four tables:

- **Table 1** shows the activity levels per person per year and per consultation (where appropriate) for the groups and the primary care information discussed.

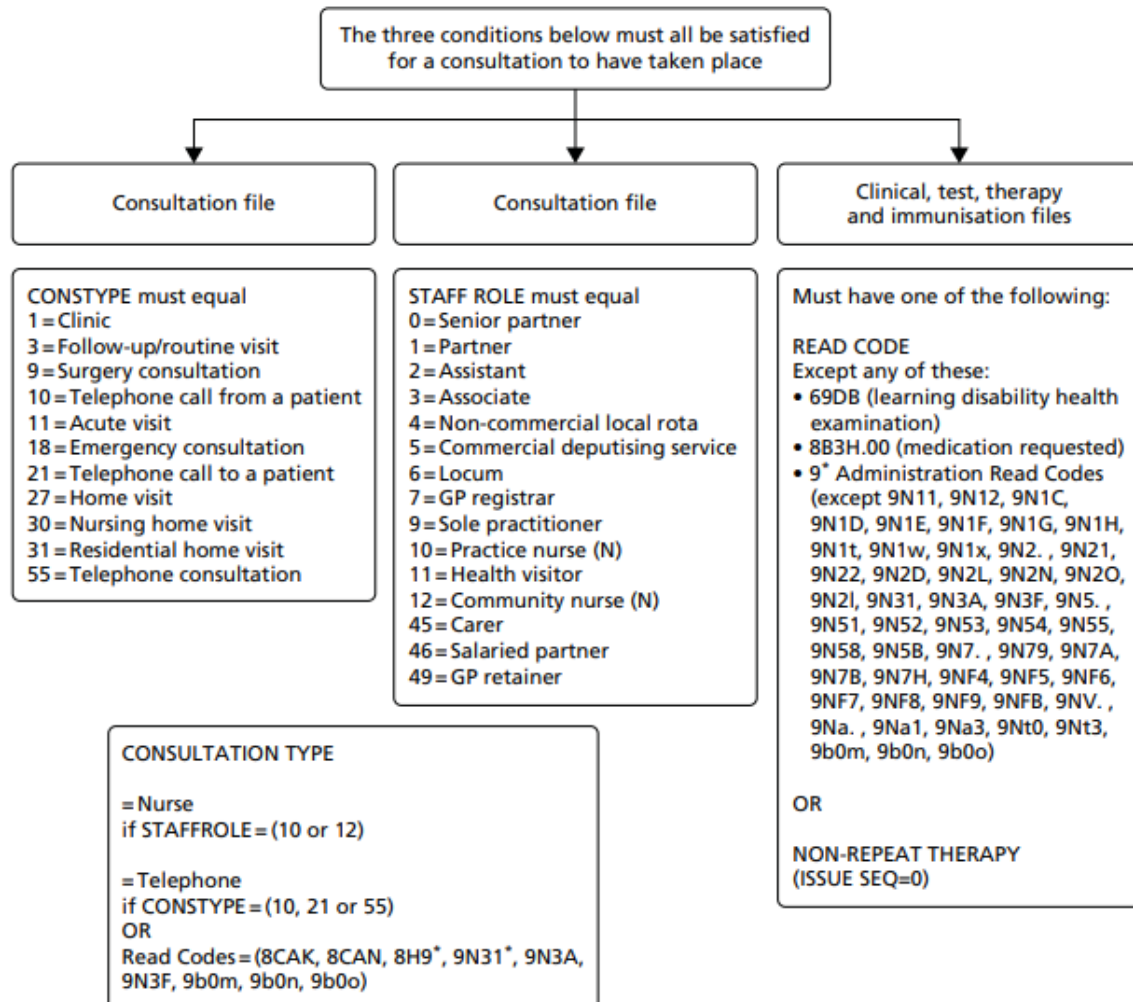
The data are shaped long, the variable 'group_type' tells us whether the values are for the whole sample population, split by age or split by NCD. The variable 'group' then tells us which age or NCD is represented by 'value'. 'activity_label' tells us what the data represent, and 'measure' tells us the unit. 'patient_n' gives the count of patients registered in the category, the week gives the week of the 'year' and 'sunday' gives us the corresponding Sunday at the start of the week. The analysis was performed in R which categorises weeks by the preceding Sunday. For example, lockdown in England started on Monday 23 March 2020; this was week 12 2020.

- **Table 2** is derived from Table 1 and shows the percentage difference each week during 2020 compared with the average in the same week from 2016–19.

- **Table 3** shows the referrals for all patients split by level of reported urgency.
- **Table 4** shows the national appointment data reported monthly by NHS Digital.

Annex 1: Definition of a consultation, and therapies, immunisations, tests and referrals associated with a consultation

From Carey et al (2017).



We also include constype == 18 as non-face-to-face because its setting is unknown – Kasteridis et al (2019)

(https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP164_care-home_placements_dementia.pdf).