# Medication data from Primary Care Electronic Extraction: August 2023

From the initial manual extraction of data for Exceed, medication data was restricted to prescriptions issued in the 5 years prior to the date of data extraction. Data was available for most British National Formulary chapters, apart from appliances, vaccines and food supplements.

The new electronic extraction has data on 8,274 patients with a study_ID starting with a 6, which are the main group of patients recruited through 36 local practices, with 3,363,514 prescriptions. In addition, there are 804 others with study_IDs starting with a different number, but typically with only one prescription record each, totalling 834 prescriptions.

Of participants registered with local practices, the range of prescription count per participant is 1 – 12,191, with an average of 406.

## Prescription Types

The full range of prescriptions typically issued in primary care appears to be represented (eyeballing list by experienced GP). There are 11,407 unique drugs in the data set. In addition to the original data set, we now have gluten free products, vitamins and food supplement, stoma, urinary appliances and dressings.

There are also 1,570 "TEXTUAL DRUGs" – which are probably uncoded but cover the full range of prescriptions. This data likely related to practices changing in clinical system and data in the old system being degraded to text data from the old system.

## Date Ranges

There is a very small amount of inaccurately dated data – the earliest record being from 1899! However, the new data now provides prescribing data going back to when electronic prescribing commenced in the NHS in the 1990s. It shows exponential growth as electronic prescribing took off,



Number of prescriptions issued per year
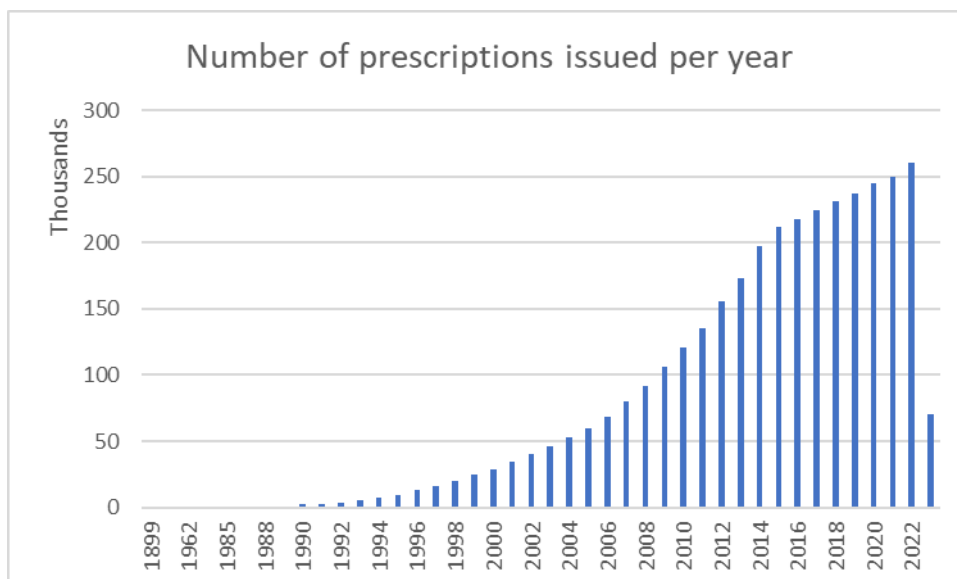
stabilising from about 2014, probably now reflecting the demographic effects of aging. It is likely therefore that we have complete prescribing data on our cohort from the early 2010s.

This data provides a number of new research opportunities. The data on drug dose, quantity and administration directions is better than in the original data set. This should enable better phenotype characterisation but this may require natural language processing as dose instructions are in free text and may not be standardised.

Drug prescribing can be used alongside diagnostic codes in phenotype definitions as many drug prescriptions are closely related to diagnosis. It can also be used to grade disease severity.

It may prove possible to link specific prescriptions to diagnostic and other codes directly through fields such as IDEvent.

DJS Aug 2023