

# Assessment

Stella Doherty

2024-11-01

## Introduction

**Research question:** Did the COVID-19 pandemic influence prescriptions of SSRI antidepressants around Scotland?

This research will focus on prescriptions of SSRI antidepressants in different Health Boards in Scotland during the colder months (October, November, December) pre-COVID (2019) versus peak-COVID (2020). Focus will be on the trends of prescriptions and whether different regions in Scotland have higher dependencies on SSRI antidepressants. I am interested in this research because since the pandemic was very sudden and an unusual occurrence I am curious how it affected peoples well-being and also being from Scotland myself, the difference in affected areas also interests me.

##Data Wrangling

```
library(tidyverse)
library(janitor)
library(gt)
library(here)
library(sf)
library(lubridate)
```

Data sets were obtained from opendata.nhs.scot and downloaded in csv format.

```
data_dec2019 <- read_csv(here("data", "opendata_prescriptions_dec2019.csv")) %>%
  clean_names()

data_dec2020 <- read_csv(here("data", "opendata_prescriptions_dec2020.csv")) %>%
  clean_names()

all_prescriptions <- data_dec2019 %>%
  rename(hbt = "hbt2014") %>%
  full_join(data_dec2020)

ssri_all_prescriptions <- all_prescriptions %>%
  filter(!is.na(bnf_item_description), str_detect(bnf_item_description, "FLUOXETINE|CITALOPRAM|ESCAITAL"))
  select(hbt, bnf_item_description, number_of_paid_items, paid_date_month)

ssri_all_prescriptions_table <- ssri_all_prescriptions %>%
  mutate(paid_date_month = as.Date(paste0(paid_date_month, "01"), format = "%Y%m%d")) %>%
  group_by(bnf_item_description, paid_date_month) %>%
```

```

summarise(quantity_sum = sum(number_of_paid_items)) %>%
arrange(-quantity_sum) %>%
head(n = 10)

ssri_all_prescriptions_table%>%
  gt(groupname_col = "bnf_item_description",
     row_group_as_column = TRUE) %>%
  cols_label(quantity_sum = "Total prescriptions",
             paid_date_month = "Year") %>%
  tab_stubhead(label = "SSRI prescription") %>%
  tab_header(title = md("Top 5 SSRI prescriptions"),
             subtitle = md("December 2019 vs December 2020 across the whole of Scotland")) %>%
  tab_source_note(source_note = "Source:...") %>%
  tab_footnote(footnote = "Fluoxetine Hydrochloride (HCL) 20mg capsules",
               location = cells_body(columns = bnf_item_description, rows = starts_with("FLU")))) %>%
  fmt_date(columns = paid_date_month,
            date_style = "yMMM") %>%
  cols_align(align = "center",
             columns = paid_date_month)

```

## Top 5 SSRI prescriptions

December 2019 vs December 2020 across the whole of Scotland

SSRI prescription	Year
FLUOXETINE HCL_CAP 20MG	Dec 2020 Dec 2019
SERTRALINE HCL_TAB 50MG	Dec 2020 Dec 2019
SERTRALINE HCL_TAB 100MG	Dec 2020 Dec 2019
CITALOPRAM HYDROB_TAB 20MG	Dec 2020 Dec 2019
CITALOPRAM HYDROB_TAB 10MG	Dec 2020 Dec 2019

Source:...

\*Description of results

```

health_boards <- read_csv("https://www.opendata.nhs.scot/dataset/9f942fdb-e59e-44f5-b534-d6e17229cc7b/r
  clean_names()

hb_general_health <- read_csv(here("data", "UV302_general_health.csv"), skip = 10) %>%
  clean_names() %>% # remove the first row (with extraneous information)
  select(-x5) %>% # remove the final (unnecessary) column
  rename(hb_name = "health_board_area_2019",
         hb_population = count) %>%
  # filter the data so that we get the population of the entire health board

```

```

filter(general_health == "All people") %>%
# select only the relevant columns
select(hb_name, hb_population) %>%
# change health board names so they match the prescription data
mutate(hb_name = paste("NHS", hb_name))

## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(....)
##   problems(dat)

joined_hb_data <- health_boards %>%
  left_join(hb_general_health, by = "hb_name") %>%
  select(hb, hb_name, hb_population) %>%
  rename(hbt = "hb")

hb_ssri_all_prescriprtions <- ssri_all_prescriptions %>%
  left_join(joined_hb_data, by = "hbt") %>%
  na.omit() %>%
  select(-hbt)

ssri_sums_item <- hb_ssri_all_prescriprtions %>%
  group_by(paid_date_month, hb_name, bnf_item_description) %>%
  summarise(quantity_sum = sum(number_of_paid_items))

wider_ssri_sums <- ssri_sums_item %>%
  pivot_wider(names_from = bnf_item_description, values_from = quantity_sum)

ssri_sums_total <- wider_ssri_sums %>%
  ungroup() %>%
  mutate(Citalopram = rowSums(select(., starts_with("CITALOPRAM")), na.rm = TRUE),
         Escitalopram = rowSums(select(., starts_with("ESCITALOPRAM")), na.rm = TRUE),
         Fluoxetine = rowSums(select(., starts_with("FLUOXETINE")), na.rm = TRUE),
         Paroxetine = rowSums(select(., starts_with("PAROXETINE")), na.rm = TRUE),
         Sertraline = rowSums(select(., starts_with("SERTRALINE")), na.rm = TRUE))

longer_ssri_sums_total <- ssri_sums_total %>%
  select(hb_name, paid_date_month, Citalopram, Escitalopram, Fluoxetine, Paroxetine, Sertraline) %>%
  pivot_longer(cols = Citalopram:Sertraline, names_to = "ssri", values_to = "sum")

ssri_per_1k <- longer_ssri_sums_total %>%
  left_join(hb_general_health) %>%
  mutate(ssri_per_1k = (sum/hb_population)*1000,
        paid_date_month = ym(paid_date_month),
        paid_date_month = year(paid_date_month))%>%
  select(-sum, -hb_population)

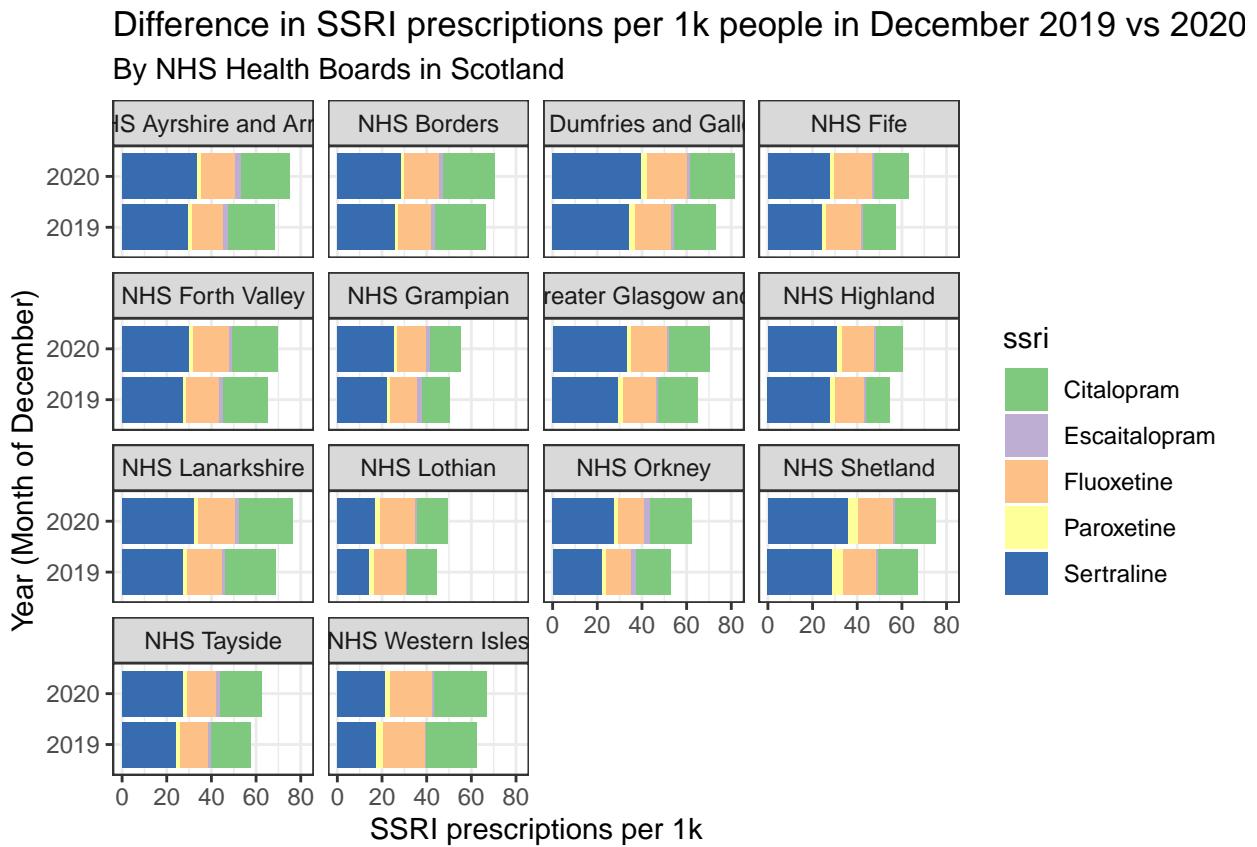
ssri_per_1k %>%
  ggplot(aes(x = reorder(factor(paid_date_month), ssri_per_1k),
             y = ssri_per_1k, fill = ssri)) +
  # We could use geom_col() here
  geom_bar(stat = "identity", position = "stack")+

```

```

facet_wrap(~hb_name) +
# Flip the axes of the plot
coord_flip() +
labs(title = "Difference in SSRI prescriptions per 1k people in December 2019 vs 2020",
    subtitle = "By NHS Health Boards in Scotland",
    x = "Year (Month of December)",
    y = "SSRI prescriptions per 1k") +
scale_fill_brewer(palette = "Accent") +
theme_bw()

```



## Geospatial data

- Need shapefile with extension .shp

```

hb_geo_data <- st_read(here("data", "SG_NHS_HealthBoards_2019")) %>%
  clean_names() %>%
  mutate(hb_name = paste("NHS ", hb_name, sep = ""))

```

```

## Reading layer 'SG_NHS_HealthBoards_2019' from data source
##   '/Users/stelladoherty/Data_Science/Final_R_Project/B192311/data/SG_NHS_HealthBoards_2019'
##   using driver 'ESRI Shapefile'
## Simple feature collection with 14 features and 4 fields
## Geometry type: MULTIPOLYGON

```

```

## Dimension:      XY
## Bounding box:  xmin: 5512.998 ymin: 530250.8 xmax: 470332 ymax: 1220302
## Projected CRS: OSGB36 / British National Grid

hb_geo_data %>%
  mutate(hb_name = paste("NHS ", hb_name, sep = ""))

```

```

## Simple feature collection with 14 features and 4 fields
## Geometry type: MULTIPOLYGON
## Dimension:      XY
## Bounding box:  xmin: 5512.998 ymin: 530250.8 xmax: 470332 ymax: 1220302
## Projected CRS: OSGB36 / British National Grid
## First 10 features:
##   hb_code          hb_name shape_leng shape_area
## 1 S08000015 NHS NHS Ayrshire and Arran 679782.3 3408802229
## 2 S08000016           NHS NHS Borders 525406.7 4742684960
## 3 S08000017 NHS NHS Dumfries and Galloway 830301.2 6676314851
## 4 S08000019           NHS NHS Forth Valley 430568.3 2733658721
## 5 S08000020           NHS NHS Grampian 802844.3 8800885268
## 6 S08000022           NHS NHS Highland 8486537.7 33637044520
## 7 S08000024           NHS NHS Lothian 378923.1 1760804006
## 8 S08000025           NHS NHS Orkney 1589228.2 1086210772
## 9 S08000026           NHS NHS Shetland 2211513.9 1656618008
## 10 S08000028          NHS NHS Western Isles 3723205.1 3268473326
##   geometry
## 1 MULTIPOLYGON (((201916.2 60...
## 2 MULTIPOLYGON (((345325.9 57...
## 3 MULTIPOLYGON (((266004.4 54...
## 4 MULTIPOLYGON (((256533 7405...
## 5 MULTIPOLYGON (((383521.2 77...
## 6 MULTIPOLYGON (((173254 6049...
## 7 MULTIPOLYGON (((313586.1 67...
## 8 MULTIPOLYGON (((347378.5 97...
## 9 MULTIPOLYGON (((419927.6 10...
## 10 MULTIPOLYGON (((55528.5 780...

```

## Join geo data

```

joined_geo_sums <- full_join(ssri_per_1k, hb_geo_data, by = "hb_name")
#class(joined_geo_sums)

joined_geo_sums <- st_as_sf(joined_geo_sums)

class(joined_geo_sums)

## [1] "sf"        "tbl_df"     "tbl"        "data.frame"

```

## Plot geo + ggplot

```

joined_geo_sums %>%
  ggplot(aes(fill = ssri_per_1k)) +
  geom_sf(colour = "white", linewidth = 0.2) +
  scale_fill_distiller(palette = "Greens", direction = 1) +
  theme_void() +
  labs(title = "Distribution of SSRIs per 1,000 across Scottish NHS Health Boards",
       subtitle = "Prescriptions from December 2019 and 2020",
       fill = "No.of prescriptions") +
  theme(plot.title = element_text(face = "bold", size = 10),
        plot.subtitle = element_text(size = 9),
        legend.title = element_text(face = "bold", size = 9)) +
  facet_wrap(~factor(paid_date_month))

```

## Distribution of SSRIs per 1,000 across Scottish NHS Health Boards

Prescriptions from December 2019 and 2020

2019

2020

