

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

Loading in required packages:

```
library(tidyverse)
library(janitor) # examine and clean data
library(gt) # create tables
library(here) # locate files
library(sf) # work with spatial vector data
library(lubridate) # manipulate dates
```

Read in prescription data sets: 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() # cleans variable names

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

1. Join both prescription data sets together
2. Filter for SSRI depressant prescriptions: Fluoxetine, Citalopram, Escitalopram, Paroxetine, Sertraline
3. Create a table presenting the top 5 most prescribed SSRIs

```
all_prescriptions <- data_dec2019 %>%
  rename(hbt = "hbt2014") %>% # ensures both health board variables have the same name for joining
  full_join(data_dec2020) # full join to retain all data
```

```

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")) %>% # ensures the month is always present
  group_by(bnf_item_description, paid_date_month) %>%
  summarise(quantity_sum = sum(number_of_paid_items)) %>% # create total number of prescriptions for all SSRI prescriptions
  arrange(-quantity_sum) %>%
  head(n = 10)

ssri_all_prescriptions_table%>%
  gt(groupname_col = "bnf_item_description",
    row_group_as_column = TRUE) %>% # groups by SSRI prescription and put into a column
  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:...") %>%
  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:...

Table of top 5 SSRI prescription: This table shows that for each SSRI prescription shown, december 2020 always had a greater number of prescriptions compared to december 2019.

1. Load in Scottish NHS health boards and general census data - join together
2. Join with SSRI prescription data

```

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(general_health == "All people") %>% # filter the data so that we get the population of the ent
  select(hb_name, hb_population) %>%
  mutate(hb_name = paste("NHS", hb_name)) # change health board names so they match the prescription da

## 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") %>% # left join to only add matching data from general h
  select(hb, hb_name, hb_population) %>%
  rename(hbt = "hb")

hb_ssri_all_prescriprions <- ssri_all_prescriptions %>%
  left_join(joined_hb_data, by = "hbt") %>% # left join to only add matching data from the new joined h
  na.omit() %>% # ignores any NA values
  select(-hbt)

```

Obtain the sum of each of the 5 SSRI prescription types (e.g. all Fluoxetine prescriptions):

```

ssri_sums_item <- hb_ssri_all_prescriprions %>%
  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")

# Calculate total sum of SSRI prescriptions per 1000 people
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)

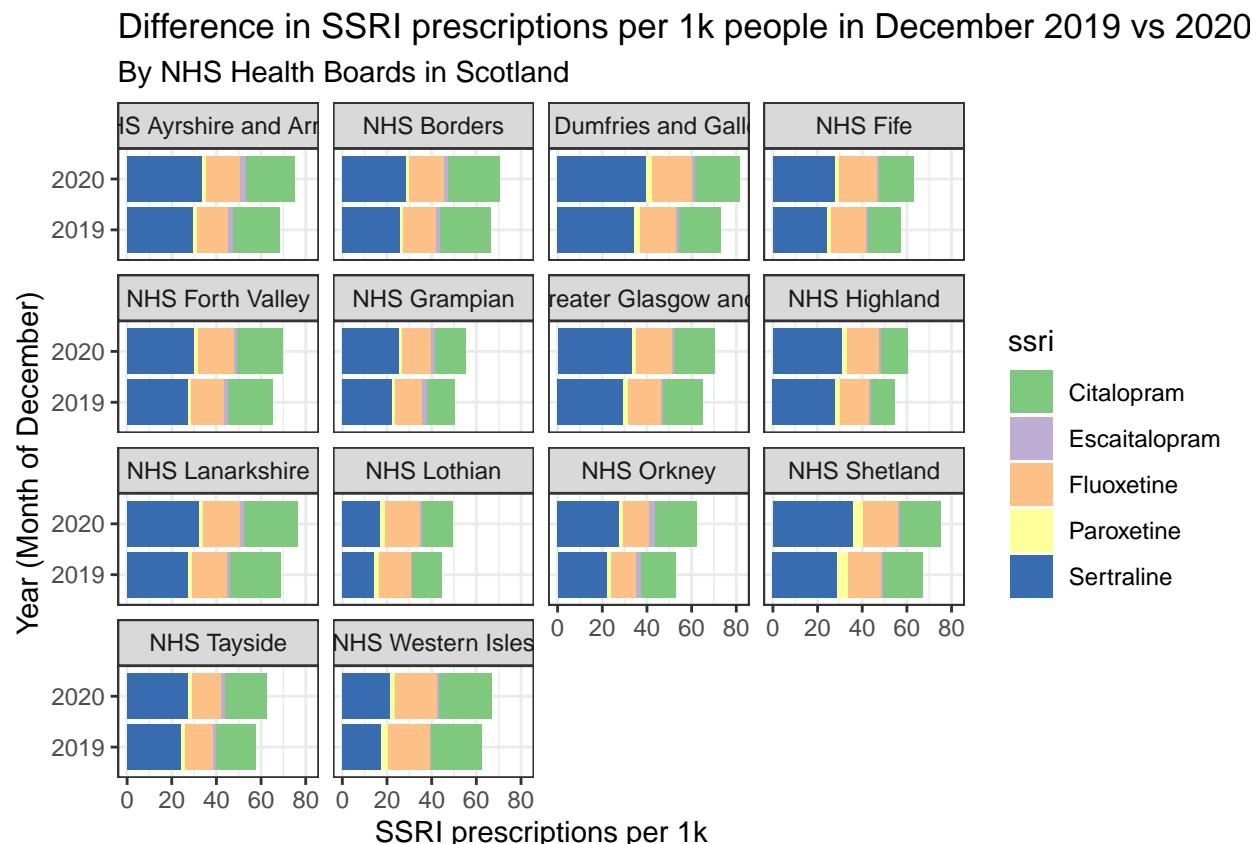
```

Present a graph SSRI prescriptions comparing numbers prescribed in december 2019 vs 2020, including all scottish health boards:

```

ssri_per_1k %>%
  ggplot(aes(x = reorder(factor(paid_date_month), ssri_per_1k),
             y = ssri_per_1k, fill = ssri)) +
  geom_bar(stat = "identity", position = "stack") +
  facet_wrap(~hb_name) +
  coord_flip() + # Flip the axes of the plot
  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()

```



Load in geospatial data:

```

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

```

Join geospatial data with SSRI prescriptions per 1000 people:

```

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

joined_geo_sums <- st_as_sf(joined_geo_sums) # ensure joined data set is in the form sf for geosptial p

```

Plot geosptial SSRI prescription data:

```

geo_map <- joined_geo_sums %>%
  ggplot(aes(fill = ssri_per_1k)) +
  geom_sf(colour = "white", linewidth = 0.2) +
  scale_fill_distiller(palette = "PuRd", 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))

geo_map

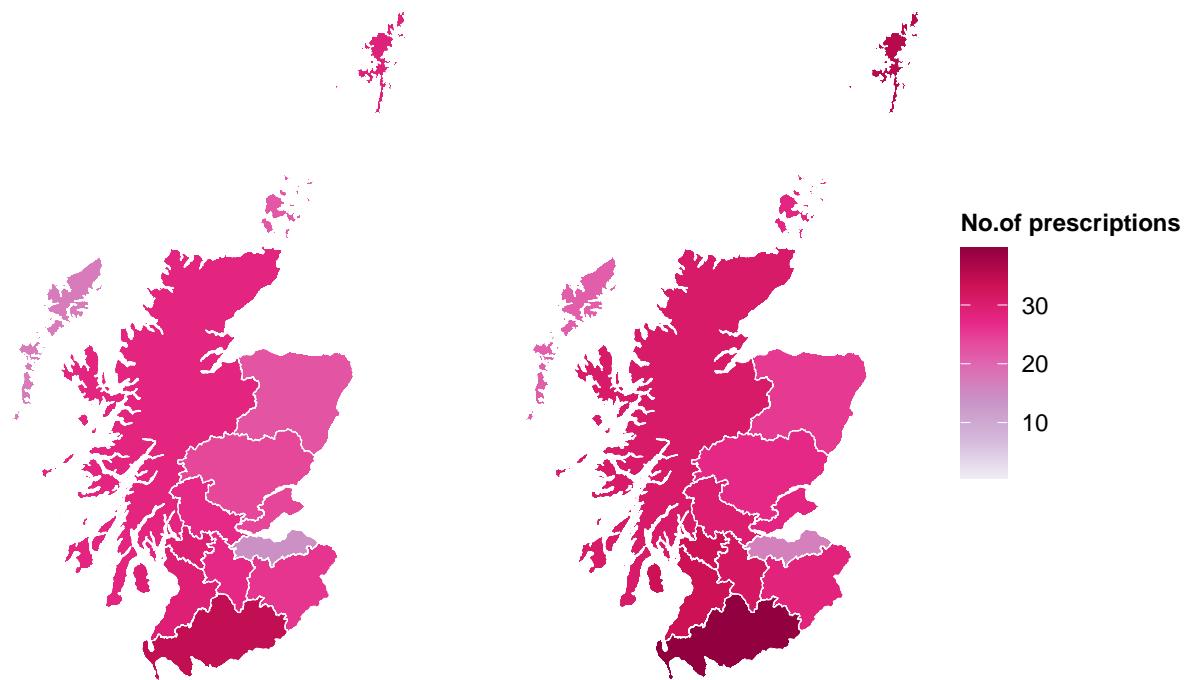
```

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

Prescriptions from December 2019 and 2020

2019

2020



Conclusions

Future and next steps:

Use of AI: