Investigating the Realtionship Between the Prescription of Medication Used to Treat Alzhimer's Disease Symptoms and Aging Populations Across Scotland in 2021

28-11-2022

```
library(knitr)
library(formatR)
library(dplyr)
library(tidyverse)
library(ggplot2)
library(gt)
library(stringr)
library(janitor)
library(readxl)
library(readr)
```

Introduction

Alzheimer's Disease is a progressive neurodegenerative disorder that causes brain atrophy and the loss of brain cells. This leads to cognitive decline that effects thinking, behavioral and social skills. Alzheimer's and other Dementia's were responsible for 117 of 100,000 deaths in Scotland in 2021, making it one of the highest causes of deaths, with similar instances of deaths due to coronary heart disease [1]. This shows that Alzheimers disease, which is the main cause of dementia, is a massive issue when it comes to public health in Scotland and will be putting pressure on the NHS. Though the root cause of the disease is still currently unknown, what is known is that there are a number of risk factors involved in developing the disease. Age is huge risk factor when it comes to Alzheimers Disease as after age 65, the risk of developing the disease doubles every 5 years. However there are other risk factors known to increase an individuals chance of developing Alzheimer's such as family history, high blood pressure, smoking, high alcohol consumption and low levels of cognitive engagement.

There is also currently no cure for Alzheimers disease, however there are a number of medications prescribed to help treat symptoms and help regain some of the cognitive function lost as the disease progresses. The three main medications prescribed are Acetylcholinesterase inhibitors which work by inhibiting the cholinesterase enzyme from breaking down Acetylcholine, a neurotransmitter who's duration of action is increased by raising levels in this way. These medications are as follows;

- Donepenzil
- Rivastigmine
- Galantamine All three can be used to treat mild to moderate Alzheimers Disease but Donepenzil is also used to treat more severe instances of the disease.

As age is the biggest risk factor for developing Alzheimer's Disease, this article aims to asses if NHS Health Boards with higher proportions of older people have higher rates of prescription of the three main medications used to treat Alzheimer's Disease symptoms This is a useful analysis as it could give an indication as to

whether Alzheimer Diseases is more prevalent in certain areas of Scotland and assess if in the future we could use proportion of older people in a population to predict where to focus resources for treatment of the disease.

NHS Data

This article will assess the data for 2021 available from the NHS website NHS Prescription in the Community website [2]. This data set contains all data relating to medications, including what, where and how much medication were prescribed by the NHS in Scotland in 2021.

Importing and Joining Data

```
Jan21 <- read_csv(url("https://www.opendata.nhs.scot/dataset/84393984-14e9-4b0d-a797-b288db64d088/resou
  clean_names %>% #to standerdise naming conventions
  write_csv("Jan21.csv") # saves the data locally
```

The data can also be downloaded and read in which is what was done for the rest of the data;

```
Feb21 <- read.csv("Feb21.csv")
Mar21 <- read.csv("Mar21.csv")
Apr21 <- read.csv("Apr21.csv")
May21 <- read.csv("May21.csv")
Jun21 <- read.csv("Jun21.csv")
Jul21 <- read.csv("Jul21.csv")
Aug21 <- read.csv("Aug21.csv")
Sep21 <- read.csv("Sep21.csv")
Oct21 <- read.csv("Oct21.csv")
Nov21 <- read.csv("Nov21.csv")
Dec21 <- read.csv("Dec21.csv")
```

After importing the data for the individual months it can then be joined together to make a data set for the year;

```
Year21 <- rbind(Jan21, Feb21, Mar21, Apr21, May21, Jun21, Jul21, Aug21, Sep21, Oct21, Nov21, Dec21) %>%
  write_csv("Year2021.csv") # saves it locally so it can be read in later
```

This creates a dataset that includes all the data for 2021 that has 13,695,460 observations and 9 variables. The variables we are interested in for this analysis are;

- bnf_item_description containing the product name, formulation and strength
- paid_quantity containing the quantity of an individual item prescribed +hb_name contains the Health Board Names

Prescription of Alzheimers Medication In Scotland

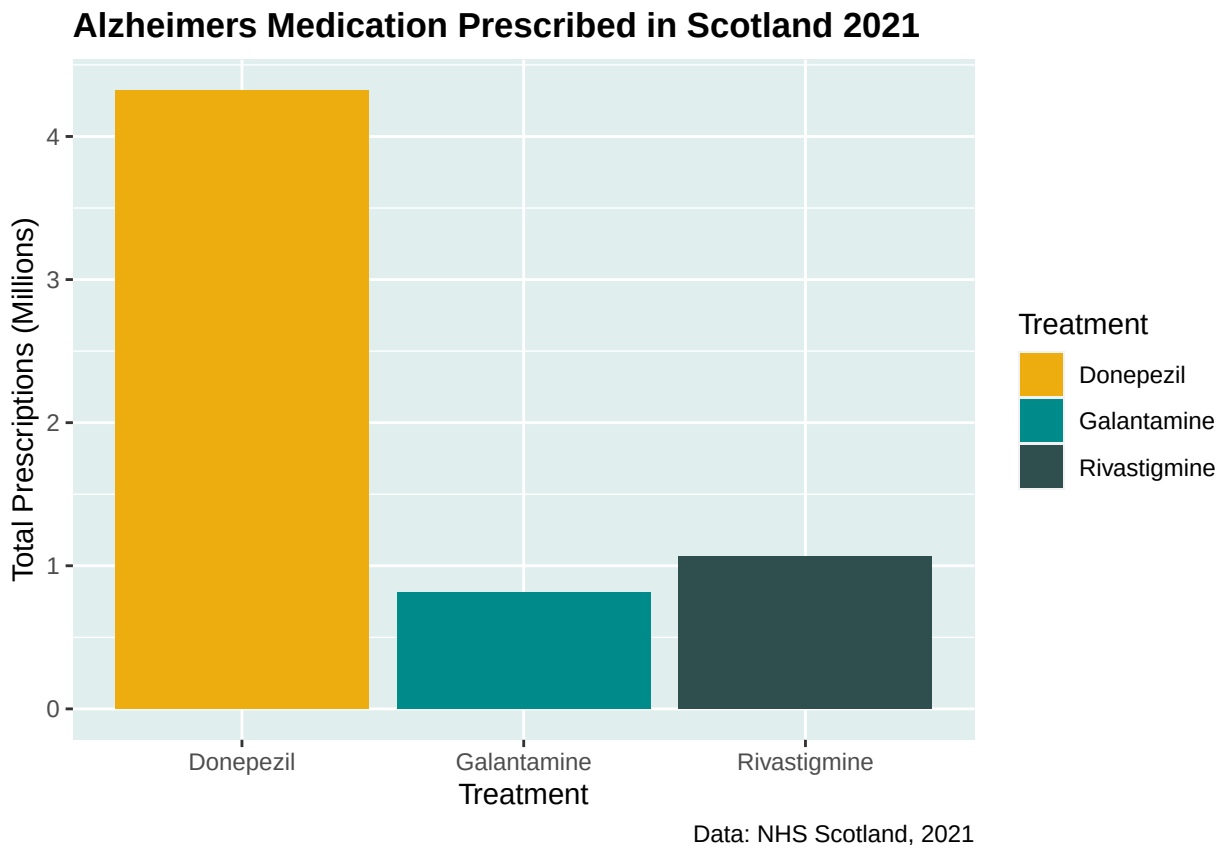
When filtering our data for the three main Alzheimer medications prescribed, it has to be noted that these medications often have more than one name. Donepezil is also known as Aricept, Rivastigmine is known as Exelon and Galantamine is also known as Reminyl [3]. For the purposes of this analysis, the medications under the two names will be renamed under the same name;

```
Alz_Medication <- Year21 %>%
    filter(str_detect(bnf_item_description, "GALANTAMINE|RIVASTIGMINE|ARICEPT|DONEPEZIL|EXELON|REMINYL")
    # filters out the Alzheimers medications prescribed
mutate(Treatment = case_when(str_detect(bnf_item_description, "DONEPEZIL|ARICEPT") ~
    "Donepezil", str_detect(bnf_item_description, "RIVASTIGMINE|EXELON") ~ "Rivastigmine",
    TRUE ~ "Galantamine"))
# adds a new column to the data joining the treatments known under two names
```

We are now able to calculate the total of each treatment prescribed across the whole of Scotland and plot this data;

```
Alz_Med_Plot_Dt <- Alz_Medication %>%
  select(Treatment, paid_quantity) %>%
  # selects just the columns we need to plot the data
  group_by(Treatment)%>%
  summarise(total_prescribed = sum(paid_quantity)) %>%
  arrange(total_prescribed) %>%
  mutate(total_prescribed = total_prescribed/1000000)
# divide the total prescribed by 1 million for easier visualization

Alz_Med_Plot <- Alz_Med_Plot_Dt %>%
  ggplot(aes(x=Treatment, y= total_prescribed, fill = Treatment))+
  geom_col()+
  theme(panel.background = element_rect(fill = "azure2"))+# sets the colours
  scale_fill_manual(values = c("darkgoldenrod2",
                               "cyan4",
                               "darkslategrey"))+
  labs(title =
    "Alzheimers Medication Prescribed in Scotland 2021",# sets the titles
       caption = "Data: NHS Scotland, 2021",
       y = "Total Prescriptions (Millions)",
       x = "Treatment")+
  theme(plot.title =element_text(face = "bold")) #makes the title bold
Alz_Med_Plot
```



In the data set, paid_quantity is a measure of how many individual treatments were prescribed i.e. if a pack containing 100 tablets was prescribed that would be counted as 100 counts in this column. This is being used in our analysis as it gives us a more accurate indication of how much Alzheimers medication is being prescribed to the population compared with the other measure in the data set, number_of_paid_items which would count 1 packet of 100 tablets as 1 count. It can be seen from the plot that there were millions of treatments distributed for Alzheimer's Disease in Scotland in 2021 and Donepenzil was the most common medication prescribed by a significant amount.Donepenzil is the only of the three medications that is used to treat Alzheimer's Disease in the more severe stages of the disease as well as more moderate stages so the results seen here could be a reflection of that.

Alzheimer Medication Prescribed Per Person Across NHS Health Boards

To get an idea of the amount of Alzheimer's medication prescribed proportional to the population in each health board, we are able to use the most recent census data available to us.

Adding 2011 Census Data for Analysis

In order to include the data from the 2011 Census we have to also include the Health Board names using Health Board lookup data available from the NHS website in our data set;

```
hb_lookup <- read.csv(url("https://www.opendata.nhs.scot/dataset/9f942fdb-e59e-44f5-b534-d6e17229cc7b/r
    clean_names() %>%
```

```
    select(c("hb", "hb_name")) %>%
    # selects the two columns we need from this data set
rename(hbt = hb)
# renames one of the columns so it is the same as the prescription data

HB_Alz_Medication <- Alz_Medication %>%
    full_join(hb_lookup, by = "hbt")  # joins the two data sets
```

Now we can import the 2011 Census data [4] and join it to our Alzheimer's Medication data set for analysis, as well as calculate how much medication is prescribed per person in each health board;

```
population_data <- read_csv("2011_Censusv0.1.csv") %>%
  clean_names() %>%
  rename("hb_name" = x1, "age_group" = x2) %>%
  #renaming column headings
  filter(hb_name != "Scotland", age_group != "All people") %>%
  # removes rows that are not required
  mutate(hb_name = paste("NHS", hb_name), hb_name = str_replace(hb_name, "&", "and")) %>%
  #renaming HB names to be in line with other data set for joining
  select(c("hb_name","all_people_total")) %>% # selects columns required
  group_by(hb_name) %>%
  summarise(all_people_total=sum(all_people_total))
#adds all agegroups together to calculate the total population in each
#healthboard

#Join the 2011 Census Data to our Alzheimer Medication Prescription Data
HB_Alz_Population <- HB_Alz_Medication %>%
  group_by(hb_name) %>%
  summarise(paid_quantity=sum(paid_quantity, na.rm = T)) %>%
  full_join(population_data, by = "hb_name") %>%
  #calculate medication per person
  mutate(treatment_per_person = paid_quantity/ all_people_total)%>%
  select(c("hb_name", "treatment_per_person")) %>%
  arrange(-treatment_per_person)
```

Now there is a data set with all the information we require, a table can be made to visualize this;

```
HB_Alz_Population %>%
    gt() %>%
    tab_header(title = "Alzheimer Medication Prescription Across Health Boards in Scotland") %>%
    # table title
cols_label(hb_name = "Health Board", treatment_per_person = "Alzheimer Medication Prescribed per Person
    # column titles
tab_style(locations = cells_column_labels(columns = everything()), style = list(cell_borders(sides = "bo
    weight = px(3)), cell_text(weight = "bold"), cell_fill(color = "cyan4", alpha = 0.5))) %>%
    # makes the underline bold and gives the background colour
tab_style(locations = cells_title(groups = "title"), style = list(cell_text(weight = "bold"),
    cell_fill(color = "darkslategrey", alpha = 0.8))) %>%
    # makes the subheadings
tab_style(location = cells_body(columns = everything()), style = list(cell_fill(color = "azure2",
    alpha = 0.3))) %>%
    tab_footnote(footnote = "Data: NHS Scotland, 2021")
```

Alzheimer Medication Prescription Across Health Boards in Scotland

Health Board	Alzheimer Medication Prescribed per Person
NHS Borders	2.3496004
NHS Western Isles	1.7676275
NHS Tayside	1.5732300
NHS Shetland	1.5555747
NHS Lothian	1.4658908
NHS Forth Valley	1.4350986
NHS Highland	1.3012289
NHS Dumfries and Galloway	1.2238508
NHS Grampian	1.1831579
NHS Fife	1.1105373
NHS Lanarkshire	1.0415149
NHS Orkney	0.9523631
NHS Ayrshire and Arran	0.8854358
NHS Greater Glasgow and Clyde	0.7575464

Data: NHS Scotland, 2021

From this table it is clear to see that proportional to the population there is the highest rates of Alzheimer's Disease medication prescribed in the NHS Borders Health Board.

Assessing the Aging Population Across Health Boards

In order to see how this relates to areas with a more aged population, next we will ascertain the percentage of the population in each health board that is over the age of 70. Individuals over 70 were selected for this analysis as the average age for the onset of Alzheimer's disease is around 70 [5].

Adding Data Required

We will getting the data required from the 2011 Census data, as well as introducing data already wrangled earlier in our analysis.

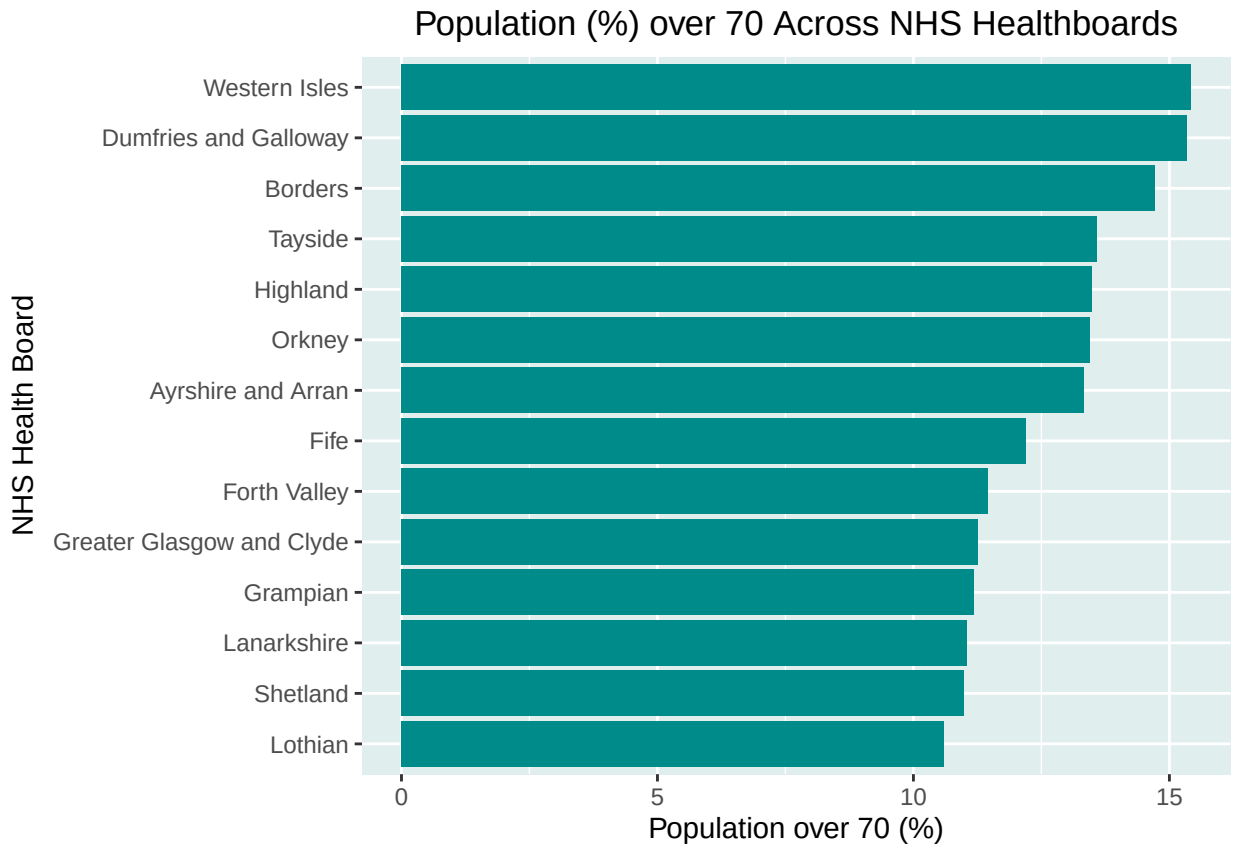
```
# calculate population over 70 from census data
age_population_data <- read_csv("2011_Censusv0.1.csv") %>%
  clean_names() %>%
  rename("hb_name" = x1, "age_group" = x2)  %>% #renames column titles
  mutate(hb_name = paste("NHS", hb_name),
         hb_name = str_replace(hb_name, "&", "and"))%>% # Making sure health boards have same names so
  filter(hb_name != "NHS Scotland", age_group == c("70 to 74", "75 to 79", "80 to 84", "85 to 89", "90
  select(c("hb_name", "all_people_total")) %>% #selects columns required
  group_by(hb_name) %>%
  summarise(over_70=sum(all_people_total))#adds together all people in the age groups over 70 for each

## create a new object joining all data required
age_percentage <- age_population_data %>%
  left_join(population_data,  by = "hb_name") %>% # joins data that has a column with the total populat
  left_join(HB_Alz_Population, by = "hb_name") %>% # joins treatment per person data
  mutate(percent_population_over_70 = (over_70/all_people_total)*100) %>% # calculates
  select(c("hb_name", "percent_population_over_70", "treatment_per_person"))
```

Plotting Data

Using this data we are able to show which Health Boards have the highest proportion of their population over 70 years old.

```
## Age by healthboard
age_HB_plot <- age_percentage %>%
    arrange(percent_population_over_70) %>%
    mutate(hb_name = str_replace(hb_name, "NHS", ""), hb_name = factor(hb_name, levels = unique(hb_name)
    ggplot(aes(x = hb_name, y = percent_population_over_70, hb_name)) + geom_col(fill = "cyan4") +
    coord_flip() + ggtitle("Population (%) over 70 Across NHS Healthboards") + theme(plot.title = eleme
    panel.background = element_rect(fill = "azure2")) + labs(x = "NHS Health Board",
    y = "Population over 70 (%)")
age_HB_plot
```

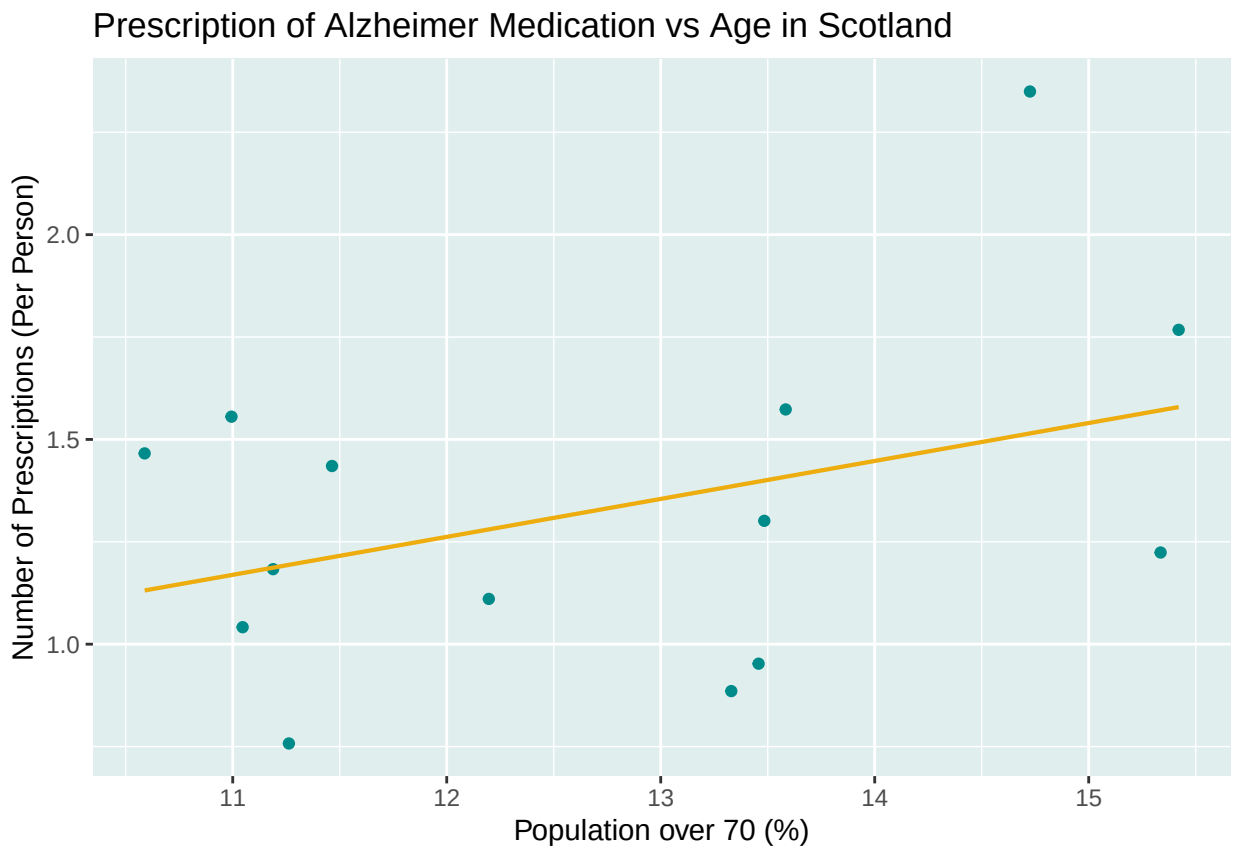


We can see from this analysis that the Western Isles, Dumfries and Galloway and the Boarders have the highest percentage of their population over the age of 70 years old.

Effect of Population Age on Alzheimer Medication Prescribed

Now we have extracted and wrangled the data to ascertain the rate of Alzheimer medication prescribed in relation to the population and the percentage of the population over the age of 70 of each area, it is possible to now plot this data to determine if there is a correlation between these two factors across the whole of Scotland. Each point of the scatterplot represents a NHS Health Board.

```
age_percentage_plot <- age_percentage %>%
    ggplot(aes(x = percent_population_over_70, y = treatment_per_person)) + geom_point(colour = "cyan4"
    theme(panel.background = element_rect(fill = "azure2")) + geom_smooth(method = lm,
    se = FALSE, col = "darkgoldenrod2", size = 0.75) + labs(title = "Prescription of Alzheimer Medicati
    x = "Population over 70 (%)", y = "Number of Prescriptions (Per Person)", fill = "hb_names")
age_percentage_plot
```



It can be seen from plotting this data that and including a line of best fit that there does seem to be a slight correlation between the percentage of the population and the number of Alzheimer's treatments prescribed per person. Though further analysis would have to be done to see if these results are statistically significant, the data seems to suggest that the higher proportion of the population that is over the age of 70, the more Alzheimer medication is prescribed when population is taken into account.

Study Limitations

There are a number of limitations that have to be taken into consideration when it comes to this analysis. The first thing we must consider is that the Census data we are using is from 2011 and the prescription data being used is from 2021. Because the data was collected 10 years apart and the population is always increasing, this could have lead to the analysis over estimating the treatments per person and underestimating the percentage of the population over 70. The second thing we must consider is that not every person with Alzheimer's disease is prescribed these medications to treat their symptoms. Some people cannot tolerate Acetylcholinesterase inhibitors as they have some major side effects, or may have received other therapies so the amount of medication prescribed in this analysis may not directly reflect the prevalence of Alzheimer's Disease in the population or even medication prescribed to treat the disease. Furthermore, Donepenzil is

also used in some cases to treat Parkinsons Disease and Dementia with Lewy Bodies. Though this is more rare and it is majoritively used to treat Alzheimer's Disease it is a factor that needs to be considered.

Discussion and Future Analysis

Our analysis first showed that Donepenzil was prescribed the most, then Rigastigmine and Galantamine the least. This is inline with what would be expected as all three are used to treat moderate symptoms of Alzheimer's however Donepenzil is also used to treat more severe instances of the disease. Next using the 2011 Census Data we were able to show the amount of treatment prescribed per person in each NHS health board, which allowed us to compare prescription of medications used to treat Alzheimer Disease symptoms in each area with population taken into account. This showed that the Borders, Western Isles and Tayside had the highest rates of Alzheimers medication prescribed per person. Further analyzing the Census data meant we were able to ascertain the NHS Health Boards with the highest proportion of their population over 70 were found to be the Western Isles, the Boarders and Dumfries and Galloway. Finally after plotting the percentage of the population over 70 against the amount of treatment prescribed per person,the data suggested after a line of best fit was drawn that there was a correlation between age and amount of medication prescribed to treat Alzheimer Disease symptoms across NHS helth boards in Scotland.

In the future, analysis of the other risk factors could be looked at using this data set to see if there is a link with Alzheimer's Disease medication prescription. One possibility is looking at the effect of deprivation on instances of prescription across NHS health boards as this is currently unknown. Areas which are less deprived usually have a higher life expectancy meaning more people with age as a risk factor for developing AD, however areas that have high rates of deprivation usually will have individuals with many other risk factors such as high alcohol consumption, higher rates of smoking ect.which could effect rates of development of the disease.

Conclusions

This study set out to establish whether there was link between areas with a higher proportion of older people and the amount of medication prescribed to treat Alzheimer Disease symptoms by the NHS in 2021. In the future this analysis could be done with the 2021 census data to give more accurate results, however overall this article showed that NHS Boards with a higher proportion of their population over 70 do have a higher rate of prescription of medication used to treat Alzheimer Disease symptoms. This is an important analysis as it could mean we could use population age in NHS health boards to focus resources where they are needed most.

Bibliogrophy

[1] https://www.nrscotland.gov.uk/files//statistics/dementia-alzheimers/2021/dementia-and-alzheimers-deaths-21-report.pdf

[2] (https://www.opendata.nhs.scot/dataset/prescriptions-in-the-community)

[3] https://www.nhs.uk/conditions/dementia/treatment/#:~:text=Donepezil%20(also%20known%20as%20Aricept,treat%20

[4] https://www.scotlandscensus.gov.uk/webapi/jsf/tableView/tableView.xhtml

[5] https://www.nia.nih.gov/health/what-are-signs-alzheimers-disease#:~:text=For%20most%20people%20with%20Alzheim