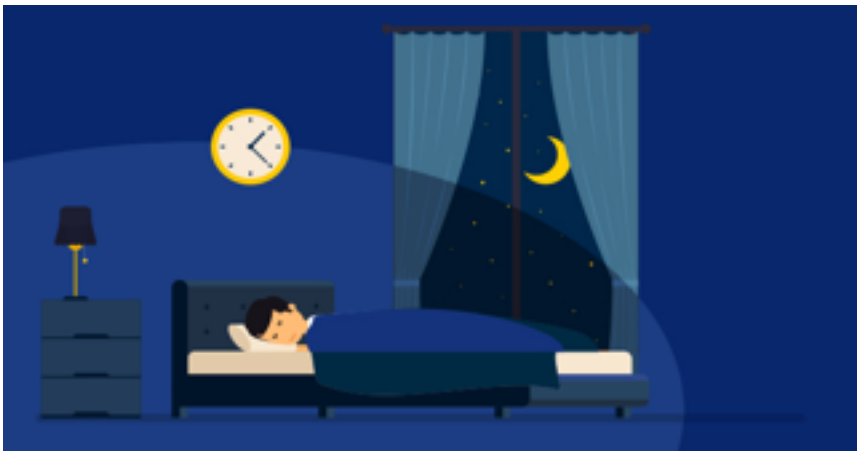
Alzheimer's and Parkinson's disease, Sleep and Mental Health

Ellen Spencer

2023-05-04



Background

Parkinson's disease and Alzheimer's disease are two common neurodegenerative diseases, affecting many individuals worldwide. Alzheimer's disease is characterised by a loss of cognitive functioning and the ability to carry out every day activities. Parkinson's disease is a condition whereby individuals are unable to control their movements, leading to difficulties initiating and coordinating movement, shaking, stiffness and other behavioural changes. Research has extensively been conducted regarding these two diseases, and a common finding suggests that both diseases are associated with various sleep disturbances and disorders (Peter-Derex et al., 2015; Comella, 2007).

Data origins

The raw data that was used for this project was obtained from the open science framework (OSF). The authors, Hayley Owens and Eric Youngstrom, uploaded the data from their project titled 'Length of sleep as a factor of Alzheimer's and Dementia' which was conducted at the University of North Carolina in 2019. These researched conducted a secondary analysis of a subset of data from Waite et al. (2014) which was gathered from interviews and biomeasurements. The aim of Owens' and Youngstrom's research project was to assess the relationship between sleep, Alzheimer's disease, dementia and Parkinson's disease.

Research Question

This aim of this project was to explore how length of sleep differed between those with neither disease, Alzheimer's disease, Parkinson's disease and both diseases. Furthermore, this project also aimed to assess how mental health influenced length of sleep in these diseases.

Setup of the project

Firstly I loaded the libraries that were to be used for this project.

```
# Load libraries that the project will use
library(here)
library(tidyverse)
library(RColorBrewer)
```

Then I loaded the data into R and assigned it to a data frame.

```
# Use library(here) to import the data and assign it to a data frame
df <- read_csv(here("Data", "raw_data.csv"))
```

```
# Check data frame
head(df)
```

```
## # A tibble: 6 x 13
##    ...1     ID   AGE DEMENTIA ALZHEIMER GENDER     EDUC  RACE_~1 PARKI~2 MNTLH~3
##   <dbl>  <dbl> <dbl> <chr>    <chr>     <chr>      <chr> <chr>   <chr>   <chr>
## 1     1 100033    84 (0) no   (0) no    (2) female (3) ~ (1) wh~ (0) no  (5) ex~
## 2     2 100033    84 (0) no   (0) no    (2) female (3) ~ (1) wh~ (0) no  (5) ex~
## 3     3 100067    65 (0) no   (0) no    (2) female (2) ~ (3) as~ (0) no  (3) go~
## 4     4 100067    65 (0) no   (0) no    (2) female (2) ~ (3) as~ (0) no  (3) go~
## 5     5 100067    65 (0) no   (0) no    (2) female (2) ~ (3) as~ (0) no  (3) go~
## 6     6 100149    70 (0) no   (0) no    (2) female (2) ~ (1) wh~ (0) no  (5) ex~
## # ... with 3 more variables: ASSUMED_SLEEP <dbl>, WASO <dbl>,
## #   ACTIGRAPH_SLEEP <dbl>, and abbreviated variable names 1: RACE_RECODE,
## #   2: PARKINSON, 3: MNTLHLTH
```

Data wrangling

To begin to clean the data, I created a new data frame which contained only the columns of interest to the project which were renamed for ease. This new data frame also had non-numerical information discarded from the Alzheimer's and Parkinson's disease columns, and numerical information removed from the mental health column.

```
# Make new data frame looking only at columns of interest which have been renamed for ease
# and have had non-numerical information removed from the disease state columns, and
# numerical information discarded from the mental health column
df1 <- df %>%
  rename(ppt_id = "ID", dementia = "DEMENTIA", alzheimers = "ALZHEIMER", mental_health =
           "MNTLHLTH", parkinsons = "PARKINSON", sleep_length = "ACTIGRAPH_SLEEP") %>%
  select(ppt_id, alzheimers, parkinsons, mental_health, sleep_length) %>%
  mutate(alzheimers = parse_number(str_sub(alzheimers, start = 1))) %>%
  mutate(parkinsons = parse_number(str_sub(parkinsons, start = 1))) %>%
  mutate(mental_health = str_sub(mental_health, start = 5, end = 13))

# Check the first few lines of the data
head(df1)
```

```
## # A tibble: 6 x 5
##   ppt_id alzheimers parkinsons mental_health sleep_length
##    <dbl>      <dbl>      <dbl> <chr>                <dbl>
## 1 100033          0          0 excellent            148.
## 2 100033          0          0 excellent            228
## 3 100067          0          0 good                 444.
## 4 100067          0          0 good                 446.
## 5 100067          0          0 good                 390.
## 6 100149          0          0 excellent            350
```

Next I removed all NA values from the data frame.

```
# Remove NA values from the data.
which(is.na(df1), arr.ind = TRUE)
df1 <- na.omit(df1)
```

Following this, I then created a new 'disease state' column within the data frame, where the four potential disease states present in the data were represented numerically, with 0 representing those with neither disease, 1 representing Alzheimer's disease, 2 representing Parkinson's disease and 3 representing a combination of the two diseases.

```
# Change values in Parkinson's disease column from '1' to '2'.
df1$parkinsons[df1$parkinsons == '1'] <- 2

# Add the Parkinson's disease column and Alzheimer's disease columns together to make a
# new disease state column, where 0 is neither, 1 is Alzheimer's, 2 is Parkinson's and 3
# is both.
df1$disease_state <- df1$parkinsons + df1$alzheimers
```

Checking the data

Using the filter function, I then checked that each value within the disease state column represented the correct disease state.

```
# Check to see that disease state value 1 represents participants with Alzheimer's disease.
filter(df1, disease_state == '1')
```

```
## # A tibble: 12 x 6
##    ppt_id alzheimers parkinsons mental_health sleep_length disease_state
##     <dbl>      <dbl>      <dbl> <chr>                <dbl>         <dbl>
##  1 120682          1          0 good                 248.              1
##  2 120682          1          0 good                 315               1
##  3 120682          1          0 good                 454.              1
##  4 164312          1          0 excellent            359.              1
##  5 164312          1          0 excellent            355               1
##  6 164312          1          0 excellent            372.              1
##  7 192264          1          0 fair                 219.              1
##  8 192264          1          0 fair                 385               1
##  9 192264          1          0 fair                 512.              1
## 10 195448          1          0 fair                 428.              1
## 11 195448          1          0 fair                 466.              1
## 12 195448          1          0 fair                 438.              1
```

```
# Check to see that disease_state value 2 represents participants with Parkinson's disease.
filter(df1, disease_state == '2')
```

```
## # A tibble: 21 x 6
##    ppt_id alzheimers parkinsons mental_health sleep_length disease_state
##     <dbl>      <dbl>      <dbl> <chr>                <dbl>         <dbl>
##  1 105689          0          2 good                   494             2
##  2 105689          0          2 good                   764             2
##  3 105689          0          2 good                   529.            2
##  4 113563          0          2 excellent              485             2
##  5 113563          0          2 excellent              468.            2
##  6 113563          0          2 excellent              420.            2
##  7 123507          0          2 fair                   477.            2
##  8 123507          0          2 fair                   484             2
##  9 123507          0          2 fair                   480.            2
## 10 166119          0          2 good                   400             2
## # ... with 11 more rows
```

```
# Check to see that disease_state value 3 represents participants with both diseases.
filter(df1, disease_state == '3')
```

```
## # A tibble: 3 x 6
##   ppt_id alzheimers parkinsons mental_health sleep_length disease_state
##    <dbl>      <dbl>      <dbl> <chr>                <dbl>         <dbl>
## 1 163874          1          2 fair                  226.            3
## 2 163874          1          2 fair                  191.            3
## 3 163874          1          2 fair                  370.            3
```

Preparing the data for visualisation

In order to prepare the data for the visualisation, I created a new data frame which had the mean length of sleep grouped by disease state and mental health.

```
# Make new data frame grouping the mean length of sleep by disease state and mental health.
df2 <- df1 %>%
  group_by(disease_state, mental_health) %>%
  summarise(mean_sleep_length = mean(sleep_length))
```

I then changed the values within the disease state column from numbers to reflect the disease state that they represent. This enabled correct labelling on the x axis for the visualisation.

```
# Change values in disease_state column from numbers to reflect the disease state in words.
for(i in 1:14) {
  if (df2$disease_state[i] == 0){
  df2$disease_state[i] = 'Neither'}
  else if (df2$disease_state[i] == 1){
  df2$disease_state[i] = 'Alzheimers'}
  else if (df2$disease_state[i] == 2){
    df2$disease_state[i] = 'Parkinsons'}
  else if (df2$disease_state[i] == 3){
    df2$disease_state[i] = 'Both'}
  }
```

From this, I then reordered the data frame so that the mental health and disease state columns were in the most appropriate order for the visualisation, rather than the standard alphabetical order.

```
# Arrange the mental health column so it is ordered from poor to excellent.
df2$mental_health <- factor(df2$mental_health, levels = c('poor', 'fair', 'good',
                                                          'very good', 'excellent'))

# Arrange the disease state column so it is ordered Neither, Alzheimers, Parkinsons, Both
df2$disease_state <- factor(df2$disease_state, levels = c('Neither', 'Alzheimers',
                                                          'Parkinsons', 'Both'))
```

Next, I created additional rows within the data frame so that there was data to plot in the visualisation to account for the missing mental health information from those with Alzheimer's disease and both diseases.

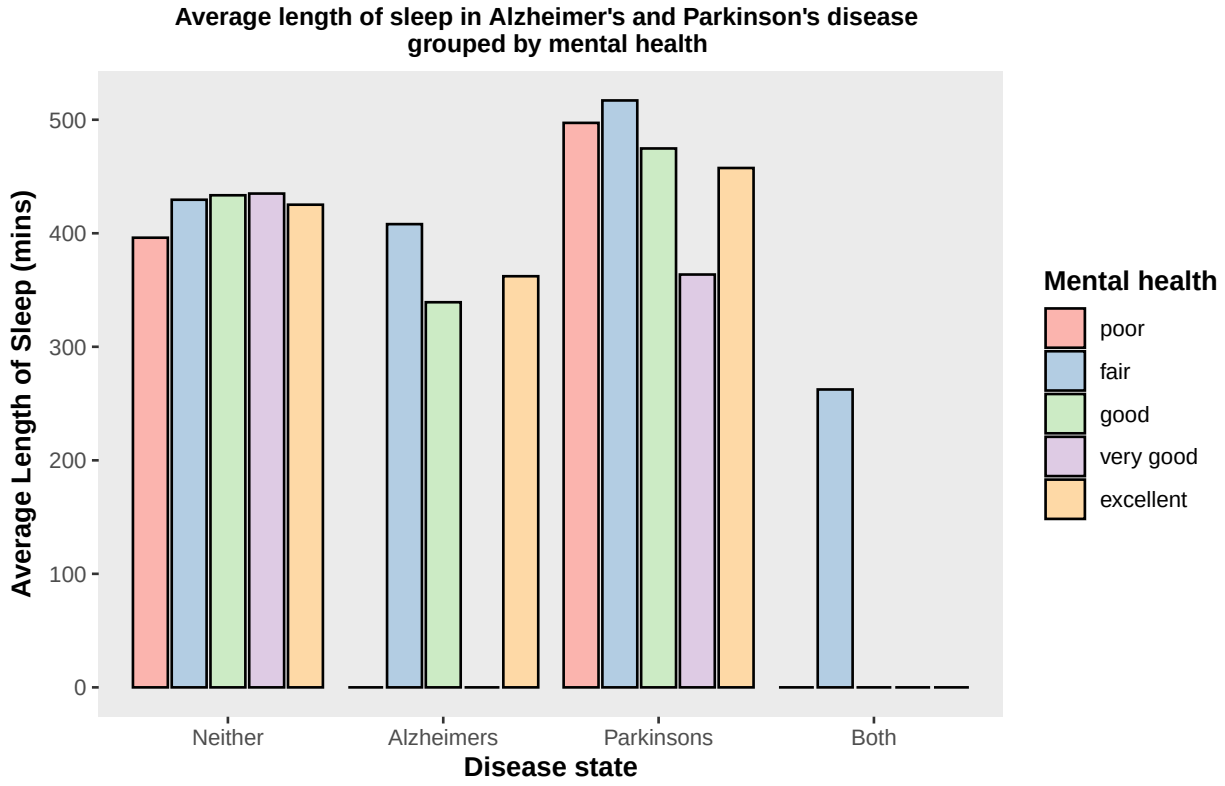
```
# Add rows to the data frame to have data to plot for the 'excellent' and 'very good'
# conditions for the Alzheimers disease state
df2[nrow(df2) + 1,] <- list('Alzheimers', 'very good', 0)
df2[nrow(df2) + 1,] <- list('Alzheimers', 'poor', 0)

# Add rows to the data frame to have data to plot for the 'poor', 'good', 'very good' and
# 'excellent' conditions for the combined disease state
df2[nrow(df2) + 1,] <- list('Both', 'poor', 0)
df2[nrow(df2) + 1,] <- list('Both', 'good', 0)
df2[nrow(df2) + 1,] <- list('Both', 'very good', 0)
df2[nrow(df2) + 1,] <- list('Both', 'excellent', 0)
```

Visualising the data

The following code makes a grouped bar graph to show the mean length of sleep in minutes for each disease state for each of the five mental health categories.

```
p <- ggplot(df2, aes(x = disease_state, y = mean_sleep_length, fill = mental_health))
p + geom_col(color = 'black', position = position_dodge2(width=.9, preserve="single")) +
  labs(x = substitute(paste(bold("Disease state"))), y =
         substitute(paste(bold("Average Length of Sleep (mins)"))),
       title = "Average length of sleep in Alzheimer's and Parkinson's disease
  grouped by mental health") +
  theme(plot.title = element_text(hjust = 0.5, size = 10, face = "bold")) +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  guides(fill = guide_legend(title = substitute(paste(bold('Mental health'))))) +
  scale_fill_brewer(breaks=c('poor', 'fair', 'good', 'very good', 'excellent'),
  palette = "Pastel1")
```



Discussion

The above visualisation demonstrates that length of sleep on average was much higher for the individuals with Parkinson's disease compared to those with Alzheimer's, both diseases and even those who had neither disease, implying that Parkinson's disease leads to an increase in sleep. On the contrary, those with Alzheimer's disease slept markedly less than those with Parkinson's disease and slightly less than those with neither disease, suggesting that a symptom of this disease state is a reduction in length of sleep. Interestingly, even though there was only data for one mental health category for the patients with both diseases, this visualisation demonstrates that the length of sleep for these individuals was drastically lower than all other disease state conditions, giving the indication that a combination of Alzheimer's and Parkinson's disease leads to a reduction in sleep length.

Furthermore, upon visualising this graph, it is clear that for those with neither disease, there is a pattern of results whereby length of sleep slightly increases as the mental health categories progress from 'poor' to 'excellent'. However, a reversal of this pattern is observed for those with Alzheimer's and Parkinson's disease, whereby the lower mental health categories of 'poor' and 'fair' actually show an increased average length of sleep compared to the higher mental health categories of 'good' and 'very good'. These results therefore suggest that whilst greater reported mental health scores correspond with an increase in length of sleep in individuals with neither disease, the same mental health scores in those with Alzheimer's and Parkinson's instead appear to reduce length of sleep.

Limitations and future directions

However, these interpretations are limited as there were only a small number of individuals within this data set who had Alzheimer's and Parkinson's disease, as well as there only being three individuals who had both diseases. Furthermore, there were only data for three of the five mental health categories for those

Alzheimer's disease and only data for one of the mental health categories for those with both diseases. Thus, there is a clear gap in the understanding of how mental health influences length of sleep in these two disease states which limits the interpretations of the above visualisation. Therefore, the visualisation produced in this project is useful in providing a baseline understanding of the influences of disease state and mental health on average length of sleep, but the aforementioned interpretations needs to be taken with caution. Further research is necessary to recruit a larger sample of individuals with Alzheimer's and Parkinson's disease who report with greater variation in mental health responses. This will enable us to more clearly visualise and understand the influence of mental health on length of sleep within Alzheimer's and Parkinson's disease. Lastly, this data could also be used to look at the influence of other variables, such as education and gender, to assess how these affect length of sleep in Alzheimer's and Parkinson's disease.

The repository for this project (including the markdown file, raw data, codebook and other files) is available online via github.

References

Comella, C. L. (2007). Sleep disorders in Parkinson's disease: An overview. *Movement Disorders, 22*(S17), S367-S373. https://doi.org/10.1002/mds.21682

Lauderdale, D. S., Philip Schumm, L., Kurina, L. M., McClintock, M., Thisted, R. A., Chen, J. H., & Waite, L. (2014). Assessment of sleep in the national social life, health, and aging project. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 69(Suppl_2), S125-S133, https://doi.org/10.1093/geronb/gbu092

Owens, H., & Youngstrom, E. (2019). Length of sleep as a factor of Alzheimer's and Dementia. https://doi.org/10.17605/OSF.IO/7ZSVB

Peter-Derex, L., Yammine, P., Bastuji, H., & Croisile, B. (2015). Sleep and Alzheimer's disease. *Sleep Medicine Reviews, 19*, 29-38. https://doi.org/10.1016/j.smrv.2014.03.007