hip_replacement_operations

XI CHEN

2024-10-30

Aim

Plot 'EQ-5D Index' scores (a combination fo five key criteria concerning patients' self-reported general health) pre and post a hip replacement operation for different age groups.

Load packages

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

Read in data

The data is in the file "Hip Replacement CCG 1819.csv", and it contains patient reported outcomes for hip replacement procedures, form April 2018 to March 2019. It was downloaded from https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms/for-hip-and-knee-replacement-procedures-april-2018-to-march-2019 We also have the data dictionary for this dataset in "proms_data_diciionary.pdf".

```
hip_data <- read.csv("Data/Hip Replacement CCG 1819.csv")
```

Prepare the data

```
glimpse(hip_data)
```

```
## Rows: 28,920
## Columns: 81
## $ Provider.Code                                  <chr> "00C", "00C", "00C", "0~
## $ Procedure                                      <chr> "Hip Replacement", "Hip~
## $ Revision.Flag                                  <int> 0, 0, 1, 1, 0, 0, 0, 0,~
## $ Year                                           <chr> "2018/19", "2018/19", "~
## $ Age.Band                                       <chr> "*", "*", "*", "*", "*"~
## $ Gender                                         <chr> "*", "*", "*", "*", "*"~
## $ Pre.Op.Q.Assisted                              <int> 2, 2, 1, 2, 2, 2, 2, 2,~
```

```
## $ Pre.Op.Q.Assisted.By                                <int> 0, 0, 0, 0, 0, 0, 0, 0,~
## $ Pre.Op.Q.Symptom.Period                             <int> 4, 2, 4, 1, 2, 1, 1, 2,~
## $ Pre.Op.Q.Previous.Surgery                           <int> 2, 1, 1, 1, 2, 2, 1, 2,~
## $ Pre.Op.Q.Living.Arrangements                        <int> 1, 1, 2, 2, 1, 2, 1, 2,~
## $ Pre.Op.Q.Disability                                 <int> 9, 1, 1, 1, 2, 1, 2, 1,~
## $ Heart.Disease                                       <int> 9, 9, 9, 9, 9, 9, 9, 1,~
## $ High.Bp                                             <int> 9, 9, 9, 9, 9, 1, 9, 1,~
## $ Stroke                                              <int> 9, 9, 9, 9, 9, 9, 1, 9,~
## $ Circulation                                         <int> 9, 9, 9, 9, 1, 9, 9, 9,~
## $ Lung.Disease                                        <int> 9, 9, 9, 9, 9, 9, 9, 9,~
## $ Diabetes                                            <int> 9, 9, 9, 9, 9, 9, 9, 1,~
## $ Kidney.Disease                                      <int> 9, 9, 9, 9, 9, 1, 9, 1,~
## $ Nervous.System                                      <int> 9, 9, 9, 9, 9, 9, 9, 9,~
## $ Liver.Disease                                       <int> 9, 9, 9, 9, 9, 9, 1, 9,~
## $ Cancer                                              <int> 9, 9, 9, 9, 9, 9, 1, 9,~
## $ Depression                                          <int> 9, 9, 9, 1, 9, 9, 9, 9,~
## $ Arthritis                                           <int> 9, 1, 1, 1, 1, 1, 9, 9,~
## $ Pre.Op.Q.Mobility                                   <int> 2, 2, 9, 2, 2, 2, 2, 1,~
## $ Pre.Op.Q.Self.Care                                  <int> 1, 2, 9, 1, 2, 1, 1, 2,~
## $ Pre.Op.Q.Activity                                   <int> 9, 3, 9, 3, 3, 2, 2, 2,~
## $ Pre.Op.Q.Discomfort                                 <int> 9, 3, 9, 3, 3, 3, 2, 2,~
## $ Pre.Op.Q.Anxiety                                    <int> 9, 1, 9, 2, 3, 1, 1, 2,~
## $ Pre.Op.Q.EQ5D.Index.Profile                         <int> 21999, 22331, 99999, 21~
## $ Pre.Op.Q.EQ5D.Index                                 <dbl> NA, -0.003, NA, 0.030, ~
## $ Post.Op.Q.Assisted                                  <int> 2, 2, 1, 2, 2, 2, 1, 2,~
## $ Post.Op.Q.Assisted.By                               <int> 9, 9, 1, 9, 9, 9, 1, 9,~
## $ Post.Op.Q.Living.Arrangements                       <int> 1, 1, 2, 2, 1, 2, 1, 9,~
## $ Post.Op.Q.Disability                                <int> 2, 9, 1, 2, 1, 2, 2, 1,~
## $ Post.Op.Q.Mobility                                  <int> 2, 9, 2, 1, 2, 2, 1, 1,~
## $ Post.Op.Q.Self.Care                                 <int> 2, 1, 2, 1, 1, 1, 1, 1,~
## $ Post.Op.Q.Activity                                  <int> 2, 9, 3, 1, 2, 2, 1, 1,~
## $ Post.Op.Q.Discomfort                                <int> 2, 1, 3, 2, 2, 2, 1, 2,~
## $ Post.Op.Q.Anxiety                                   <int> 2, 1, 2, 1, 2, 1, 1, 1,~
## $ Post.Op.Q.Satisfaction                              <int> 2, 3, 2, 1, 3, 1, 1, 9,~
## $ Post.Op.Q.Sucess                                    <int> 1, 1, 1, 1, 2, 2, 1, 9,~
## $ Post.Op.Q.Allergy                                   <int> 2, 2, 2, 2, 2, 9, 9, 9,~
## $ Post.Op.Q.Bleeding                                  <int> 2, 2, 2, 2, 2, 9, 9, 9,~
## $ Post.Op.Q.Wound                                     <int> 2, 2, 1, 2, 2, 9, 9, 9,~
## $ Post.Op.Q.Urine                                     <int> 2, 2, 2, 2, 2, 1, 9, 9,~
## $ Post.Op.Q.Further.Surgery                           <int> 2, 2, 1, 2, 2, 2, 2, 9,~
## $ Post.Op.Q.Readmitted                                <int> 2, 2, 1, 2, 2, 2, 2, 9,~
## $ Post.Op.Q.EQ5D.Index.Profile                        <int> 22222, 91911, 22332, 11~
## $ Post.Op.Q.EQ5D.Index                                <dbl> 0.516, NA, -0.074, 0.79~
## $ Hip.Replacement.EQ5D.Index.Post.Op.Q.Predicted      <dbl> NA, NA, NA, 0.5154424, ~
## $ Pre.Op.Q.EQ.VAS                                     <int> 999, 999, 999, 50, 30, ~
## $ Post.Op.Q.EQ.VAS                                    <int> 70, 999, 80, 90, 70, 60~
## $ Hip.Replacement.EQ.VAS.Post.Op.Q.Predicted          <dbl> NA, NA, NA, 60.05266, 7~
## $ Hip.Replacement.Pre.Op.Q.Pain                       <int> 1, 0, 0, 0, 0, 0, 1, 2,~
## $ Hip.Replacement.Pre.Op.Q.Sudden.Pain                <int> 0, 1, 0, 0, 0, 1, 4, 3,~
## $ Hip.Replacement.Pre.Op.Q.Night.Pain                 <int> 2, 0, 1, 0, 0, 1, 1, 4,~
## $ Hip.Replacement.Pre.Op.Q.Washing                    <int> 3, 1, 1, 2, 2, 4, 4, 0,~
## $ Hip.Replacement.Pre.Op.Q.Transport                  <int> 2, 1, 1, 0, 1, 2, 2, 3,~
## $ Hip.Replacement.Pre.Op.Q.Dressing                   <int> 1, 0, 1, 0, 1, 4, 2, 0,~
## $ Hip.Replacement.Pre.Op.Q.Shopping                   <int> 3, 2, 0, 0, 0, 0, 3, 0,~
```

```
## $ Hip.Replacement.Pre.Op.Q.Walking              <int> 2, 0, 1, 1, 1, 3, 3, 4,~
## $ Hip.Replacement.Pre.Op.Q.Limping              <int> 2, 0, 0, 1, 0, 0, 0, 3,~
## $ Hip.Replacement.Pre.Op.Q.Stairs               <int> 2, 1, 1, 1, 1, 2, 4, 3,~
## $ Hip.Replacement.Pre.Op.Q.Standing             <int> 1, 1, 1, 2, 1, 1, 4, 4,~
## $ Hip.Replacement.Pre.Op.Q.Work                 <int> 1, 1, 0, 1, 0, 0, 4, 2,~
## $ Hip.Replacement.Pre.Op.Q.Score                <int> 20, 8, 7, 8, 7, 18, 32,~
## $ Hip.Replacement.Post.Op.Q.Pain                <int> 3, 4, 2, 2, 4, 2, 2, 9,~
## $ Hip.Replacement.Post.Op.Q.Sudden.Pain         <int> 4, 4, 4, 2, 2, 2, 4, 4,~
## $ Hip.Replacement.Post.Op.Q.Night.Pain          <int> 4, 4, 4, 1, 4, 2, 4, 4,~
## $ Hip.Replacement.Post.Op.Q.Washing             <int> 4, 3, 3, 4, 3, 4, 4, 9,~
## $ Hip.Replacement.Post.Op.Q.Transport           <int> 4, 4, 2, 3, 3, 2, 4, 3,~
## $ Hip.Replacement.Post.Op.Q.Dressing            <int> 2, 4, 3, 3, 4, 4, 3, 9,~
## $ Hip.Replacement.Post.Op.Q.Shopping            <int> 4, 2, 0, 3, 2, 0, 4, 0,~
## $ Hip.Replacement.Post.Op.Q.Walking             <int> 4, 3, 1, 4, 3, 2, 4, 4,~
## $ Hip.Replacement.Post.Op.Q.Limping             <int> 3, 1, 1, 4, 2, 0, 3, 4,~
## $ Hip.Replacement.Post.Op.Q.Stairs              <int> 4, 1, 1, 3, 2, 4, 4, 4,~
## $ Hip.Replacement.Post.Op.Q.Standing            <int> 3, 4, 3, 3, 4, 2, 4, 4,~
## $ Hip.Replacement.Post.Op.Q.Work                <int> 4, 4, 2, 4, 2, 2, 3, 4,~
## $ Hip.Replacement.Post.Op.Q.Score               <int> 43, 38, 26, 36, 35, 26,~
## $ Hip.Replacement.OHS.Post.Op.Q.Predicted       <dbl> 42.20017, 35.29577, 23.~
```

Select age and quality of life score pre and post operation

```
age_EQ5D <- hip_data %>%
  select('Age.Band','Pre.Op.Q.EQ5D.Index','Post.Op.Q.EQ5D.Index') %>%
  rename(Age = 'Age.Band',
         EQ5D_Pre = 'Pre.Op.Q.EQ5D.Index',
         EQ5D_Post = 'Post.Op.Q.EQ5D.Index'
         )

head(age_EQ5D)
```

```
##   Age EQ5D_Pre EQ5D_Post
## 1   *       NA     0.516
## 2   *   -0.003        NA
## 3   *       NA    -0.074
## 4   *    0.030     0.796
## 5   *   -0.239     0.620
## 6   *    0.159     0.691
```

Remove missing values

```
age_EQ5D$Age %>% unique()
```

```
## [1] "*"        "60 to 69" "70 to 79" "80 to 89" "50 to 59" "40 to 49"
```

```
age_EQ5D$Age %>% table()
```

```
## .
##        * 40 to 49 50 to 59 60 to 69 70 to 79 80 to 89
##     2309      275     2998     8303    11157     3878
```

```
age_EQ5D %>% summary()
```

```
##      Age               EQ5D_Pre         EQ5D_Post
```

```
##  Length:28920       Min.   :-0.5940   Min.   :-0.5940
##  Class :character   1st Qu.: 0.0300   1st Qu.: 0.6910
##  Mode  :character   Median : 0.3640   Median : 0.8150
##                     Mean   : 0.3357   Mean   : 0.7975
##                     3rd Qu.: 0.6200   3rd Qu.: 1.0000
##                     Max.   : 1.0000   Max.   : 1.0000
##                     NA's   :1794      NA's   :1104
```

```
age_EQ5D_noNA <- age_EQ5D %>%
  drop_na() %>%
  filter(Age !='*')

table(age_EQ5D_noNA$Age)
```

```
##
## 40 to 49 50 to 59 60 to 69 70 to 79 80 to 89
##      261     2808     7647     9986     3340
```

```
summary(age_EQ5D_noNA)
```

```
##      Age                EQ5D_Pre        EQ5D_Post
##  Length:24042       Min.   :-0.594   Min.   :-0.5940
##  Class :character   1st Qu.: 0.055   1st Qu.: 0.6910
##  Mode  :character   Median : 0.516   Median : 0.8150
##                     Mean   : 0.339   Mean   : 0.7995
##                     3rd Qu.: 0.656   3rd Qu.: 1.0000
##                     Max.   : 1.000   Max.   : 1.0000
```

Check that data is tidy

```
head(age_EQ5D_noNA)
```

```
##        Age EQ5D_Pre EQ5D_Post
## 1 60 to 69   -0.016     0.516
## 2 60 to 69    0.159     0.743
## 3 60 to 69    0.030     0.727
## 4 60 to 69    0.587     0.850
## 5 60 to 69    0.623     0.796
## 6 60 to 69    0.691     1.000
```

```
tidy_age_EQ5D_noNA <- age_EQ5D_noNA %>%
  pivot_longer(c(EQ5D_Pre,EQ5D_Post),
               names_to = 'Time',
               names_prefix = 'EQ5D_',
               values_to = 'EQ5D'
               )
head(tidy_age_EQ5D_noNA)
```

```
## # A tibble: 6 x 3
##   Age      Time    EQ5D
##   <chr>    <chr>  <dbl>
## 1 60 to 69 Pre   -0.016
## 2 60 to 69 Post   0.516
## 3 60 to 69 Pre    0.159
## 4 60 to 69 Post   0.743
## 5 60 to 69 Pre    0.03
```

```
## 6 60 to 69 Post   0.727
```

Plot quality of life pre and operation for each age group

```
tidy_age_EQ5D_noNA$Time <- factor(tidy_age_EQ5D_noNA$Time,levels=c('Pre','Post'))

tidy_age_EQ5D_noNA %>%
  ggplot() +
  geom_boxplot(aes(x = Time, y = EQ5D, colour = Age))
```