Supplement 3

General health

Peter Kamerman

01 April 2020

Contents

Objective

To describe the relationship between pain characteristics and rating of general health in the analysis cohort, at each time point.

Analysis notes

Definitions of missingness

Data were regarded as **missing** when *pain in the last week* data were not present for one or more of weeks 0, 12, 24, 36, 48. Data also were classified as **missing** when there were inconsistencies in the data across the variables collected within a week.

Definition of data inconsistencies

Pain was defined as *pain in the last week* being 'Yes', and *pain at its worst* being > 0. These two measurements were then the "gatekeeper" measurements, such that the two measurements both had to be positive ('Yes' and '> 0', respectively) in order for there to be any entries for *site of pain* and *site of worst pain*. Were the data were inconsistent (e.g., when there was no *pain in the last week* and *pain at its worst* = 0, but there were entries for *site of pain* and *site of worst pain*), then the *site of pain* and *site of worst pain* entries were marked as **inconsistent**.

Data also were considered **inconsistent** when *pain in the last week* = 'Yes', but *site of worst pain* = 'None'.

Lastly, data were considered **inconsistent** when *site of worst pain* was not listed as one of the pain locations for a given measurement week.

For analysis purposes, missing data in the *site of pain* columns were changed to **'No'** (pain not present in the site). This approach was conservative, but we believed that the approach would have the least effect on the outcome, while still retaining as many participants as possible.

Import data

```
df <- read_rds('data-cleaned/data-ADVANCE.rds') %>%
    select(ranid, interval_name, pain_in_the_last_week, pain_worst,
           general_health, any_missing, interval_numeric)
```

First look

```
head(df)
```

```
## # A tibble: 6 x 7
##   ranid interval_name pain_in_the_las~ pain_worst general_health any_missing
##   <chr> <ord>         <chr>                <dbl>          <dbl> <chr>
## 1 01-0~ 0 weeks       No                       0              4 No
## 2 01-0~ 12 weeks      No                       0              4 No
## 3 01-0~ 24 weeks      No                       0              5 No
## 4 01-0~ 36 weeks      No                       0              5 No
## 5 01-0~ 48 weeks      No                       0              4 No
## 6 01-0~ 0 weeks       No                       0              3 No
## # ... with 1 more variable: interval_numeric <dbl>
```

```
glimpse(df)
```

```
## Rows: 5,265
## Columns: 7
## $ ranid                <chr> "01-0001", "01-0001", "01-0001", "01-0001", "...
## $ interval_name        <ord> 0 weeks, 12 weeks, 24 weeks, 36 weeks, 48 wee...
## $ pain_in_the_last_week <chr> "No", "No", "No", "No", "No", "No", "Yes", "Y...
## $ pain_worst           <dbl> 0, 0, 0, 0, 0, 0, 3, 3, 5, 0, 0, 0, 0, 0, 0, ...
## $ general_health       <dbl> 4, 4, 5, 5, 4, 3, 5, 3, 3, 3, 4, 5, 5, 5, 5, ...
## $ any_missing          <chr> "No", "No", "No", "No", "No", "No", "No", "No...
## $ interval_numeric     <dbl> 0, 12, 24, 36, 48, 0, 12, 24, 36, 48, 0, 12, ...
```

Basic clean

```r
# Clean and process data
df %<>%
    filter(any_missing == 'No') %>%
    select(-any_missing)
```

Quick tabulation

Analysis data set for the period 0 to 48 weeks

```r
# Tabulate data
xtabs(~interval_name, data = df)
```

```
## interval_name
##  0 weeks 12 weeks 24 weeks 36 weeks 48 weeks
##      787      787      787      787      787
```

Analysis

General health

Rating of perceived health status on a 5-point Likert scale (1 = poor, 5 = excellent).

Overall rating by time period

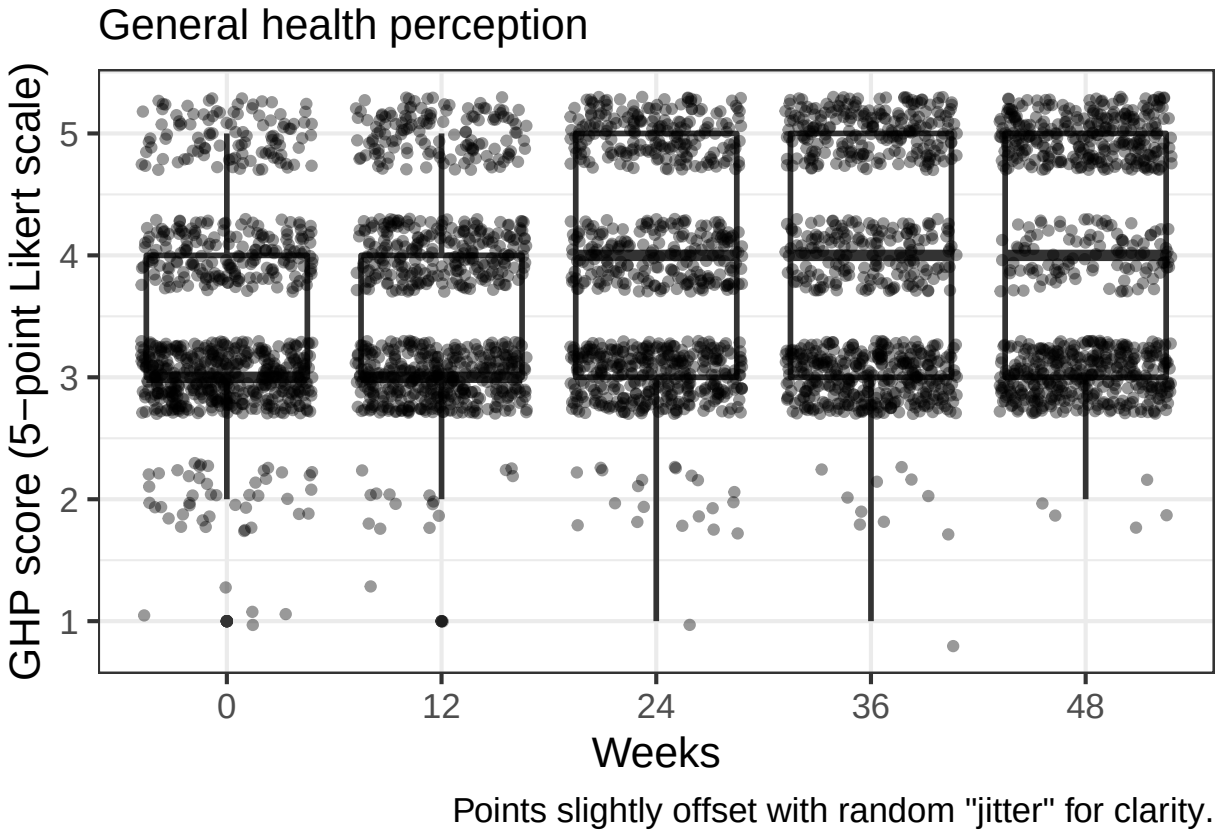
```r
# Tabulate data
df %>%
  select(interval_name, general_health) %>%
  group_by(interval_name) %>%
  skim() %>%
  yank('numeric') %>%
  select(-skim_variable, -hist, -complete_rate) %>%
  kable(caption = '7-number summary of the general health perception score')
```

Table 1: 7-number summary of the general health perception score

interval_name	n_missing	mean	sd	p0	p25	p50	p75	p100
0 weeks	4	3.449553	0.8152615	1	3	3	4	5
12 weeks	1	3.650127	0.7971166	1	3	3	4	5
24 weeks	0	3.762389	0.8712007	1	3	4	5	5
36 weeks	1	3.807888	0.8858341	1	3	4	5	5
48 weeks	2	3.901911	0.9323051	2	3	4	5	5

```r
# Plot data
df %>%
  ggplot(data = .) +
  aes(x = factor(interval_numeric),
      y = general_health) +
  geom_boxplot(size = 1) +
  geom_point(position = position_jitter(height = 0.3),
             shape = 21,
```

```
            fill = '#000000',
            alpha = 0.4) +
  labs(subtitle = 'General health perception',
       y = 'GHP score (5-point Likert scale)',
       x = 'Weeks',
       caption = 'Points slightly offset with random "jitter" for clarity.')
```



Rating by pain status, across time

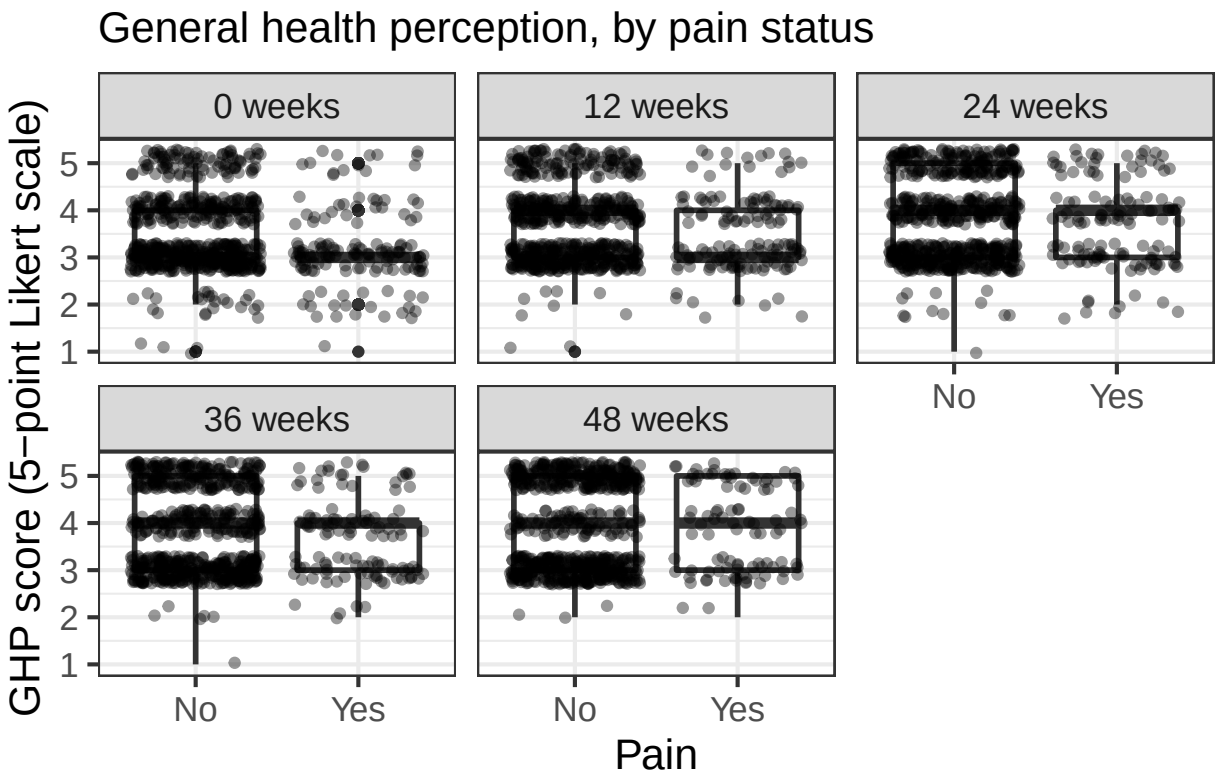
```
# Tabulate data
df %>%
  select(interval_name, pain_in_the_last_week, general_health) %>%
  group_by(pain_in_the_last_week, interval_name) %>%
  skim() %>%
  yank('numeric') %>%
  select(-skim_variable, -hist, -complete_rate) %>%
  kable(caption = '7-number summary of the general health score, by pain status')
```

Table 2: 7-number summary of the general health score, by pain status

pain_in_the_last_week	interval_name	n_missing	mean	sd	p0	p25	p50	p75	p100
No	0 weeks	2	3.509464	0.8052388	1	3	3	4	5
No	12 weeks	0	3.684783	0.8014478	1	3	4	4	5
No	24 weeks	0	3.771212	0.8662033	1	3	4	5	5
No	36 weeks	1	3.828869	0.8931933	1	3	4	5	5
No	48 weeks	2	3.904209	0.9382674	2	3	4	5	5
Yes	0 weeks	2	3.194631	0.8110541	1	3	3	3	5
Yes	12 weeks	1	3.492958	0.7602396	2	3	3	4	5
Yes	24 weeks	0	3.716535	0.8988320	2	3	4	4	5

pain_in_the_last_week	interval_name	n_missing	mean	sd	p0	p25	p50	p75	p100
Yes	36 weeks	0	3.684210	0.8341790	2	3	4	4	5
Yes	48 weeks	0	3.885417	0.8928935	2	3	4	5	5

```
# Plot data
df %>%
  ggplot(data = .) +
  aes(x = pain_in_the_last_week,
      y = general_health) +
  geom_boxplot(size = 1) +
  geom_point(position = position_jitter(height = 0.3),
             shape = 21,
            fill = '#000000',
            alpha = 0.4) +
  labs(subtitle = 'General health perception, by pain status',
       y = 'GHP score (5-point Likert scale)',
       x = 'Pain',
       caption = 'Points slightly offset with random "jitter" for clarity.') +
  facet_wrap(~interval_name, ncol = 3)
```

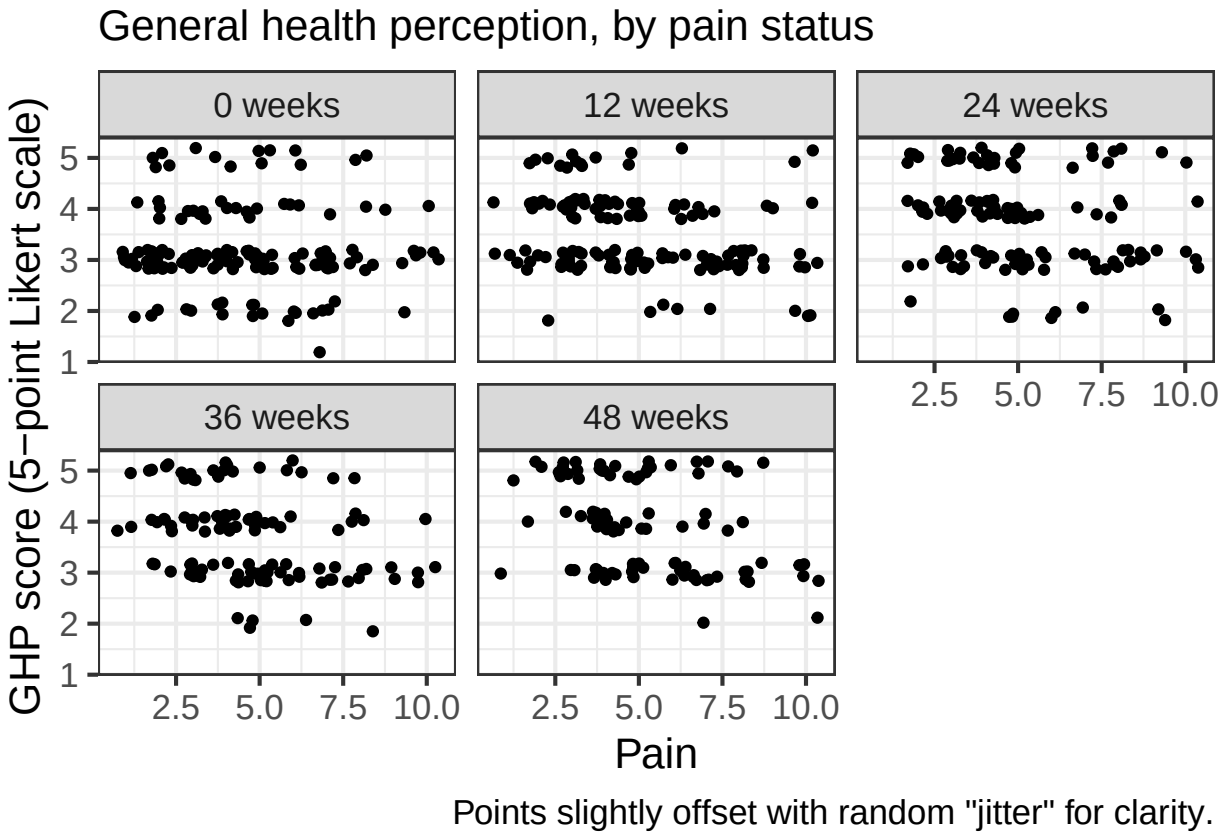


No clear relationship between the presence of pain and GHP rating.

Rating by pain intensity, across time

```
# Plot data
df %>%
  filter(pain_in_the_last_week != 'No') %>%
  ggplot(data = .) +
  aes(x = pain_worst,
      y = general_health) +
```

```
  geom_point(position = position_jitter(height = 0.2)) +
  labs(subtitle = 'General health perception, by pain status',
       y = 'GHP score (5-point Likert scale)',
       x = 'Pain',
       caption = 'Points slightly offset with random "jitter" for clarity.') +
  facet_wrap(~interval_name, ncol = 3)
```



No clear relationship between the pain intensity and GHP rating.

Session information

```
sessionInfo()
```

```
## R version 3.6.1 (2019-07-05)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Catalina 10.15.3
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] knitr_1.27         skimr_2.0.2        magrittr_1.5       forcats_0.4.0
```

```
##  [5] stringr_1.4.0      dplyr_0.8.3        purrr_0.3.3        readr_1.3.1
##  [9] tidyr_1.0.0        tibble_3.0.0       ggplot2_3.3.0.9000 tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_0.2.5 xfun_0.12        repr_1.0.2       haven_2.2.0
##  [5] lattice_0.20-38  colorspace_1.4-1 vctrs_0.2.4      generics_0.0.2
##  [9] htmltools_0.4.0  base64enc_0.1-3  yaml_2.2.1       utf8_1.1.4
## [13] rlang_0.4.5      pillar_1.4.3     withr_2.1.2      glue_1.3.2
## [17] DBI_1.1.0        dbplyr_1.4.2     modelr_0.1.5     readxl_1.3.1
## [21] lifecycle_0.2.0  munsell_0.5.0    gtable_0.3.0     cellranger_1.1.0
## [25] rvest_0.3.5      evaluate_0.14    labeling_0.3     fansi_0.4.1
## [29] highr_0.8        broom_0.5.3      Rcpp_1.0.4       backports_1.1.5
## [33] scales_1.1.0     jsonlite_1.6.1   farver_2.0.3     fs_1.3.1
## [37] hms_0.5.3        digest_0.6.25    stringi_1.4.6    grid_3.6.1
## [41] cli_2.0.2        tools_3.6.1      crayon_1.3.4     pkgconfig_2.0.3
## [45] ellipsis_0.3.0   xml2_1.2.5       reprex_0.3.0     lubridate_1.7.4
## [49] assertthat_0.2.1 rmarkdown_2.1    httr_1.4.1       rstudioapi_0.11
## [53] R6_2.4.1         nlme_3.1-143     compiler_3.6.1
```