elena\_skin\_analysis

Peter Higgins

3/15/2021

## Read in data

Start with FCP and CRP

fcp <- readxl::read\_excel(here("Oct FCP and CRP for charts.xlsx"), sheet = "FCP") %>%   
 group\_by(mrn) %>%   
 arrange(mrn, desc(time\_point)) %>%   
 mutate(delta\_fcp = value - lag(value)) %>%   
 mutate(pct\_change\_fcp = 100\*delta\_fcp/lag(value)) %>%   
 ungroup() %>%   
 select(group:pct\_change\_fcp)  
  
crp <- readxl::read\_excel(here("Oct FCP and CRP for charts.xlsx"), sheet = "CRP") %>%   
 group\_by(mrn) %>%   
 arrange(mrn, desc(time\_point)) %>%   
 mutate(delta\_crp = value - lag(value)) %>%   
 mutate(pct\_change\_crp = 100\*delta\_crp/lag(value)) %>%   
 ungroup() %>%   
 select(group:pct\_change\_crp)

Now read in Likert values

path\_likert <- readxl::read\_excel(here("Oct Likert Numbers.xlsx"), sheet = "path") %>%   
 select(group:path\_likert)  
  
endo\_likert <- readxl::read\_excel(here("Oct Likert Numbers.xlsx"), sheet = "endo") %>%   
 select(group:endo\_likert)  
  
imaging\_likert <- readxl::read\_excel(here("Oct Likert Numbers.xlsx"), sheet = "imaging") %>%   
 select(group:imaging\_likert)

Now read in demographics

demog <- readxl::read\_excel(here("Oct demographics organized.xlsx")) %>% clean\_names() %>%   
 select(group\_type:stelara\_start) %>%   
 mutate(dob = as.Date(dob),  
 stelara\_start = as.Date(stelara\_start)) %>%   
 rename(group = group\_type)

## Demographic table with gtsummary and flextable

demog %>%   
 select(-dob, -stelara\_start) %>%   
 mutate(sex = case\_when(sex == "F" ~ "Female",  
 sex =="M" ~ "Male")) %>%   
 mutate(tobacco = case\_when(tobacco == "Y" ~ "Yes",  
 tobacco == "N" ~ "No")) %>%   
 mutate(race = case\_when(race == "A" ~ "Asian or Pacific Islander",  
 race == "AA" ~ "African-American",  
 race == "C" ~ "Caucasian",  
 race == "HL" ~ "Caucasian",  
 race == "O" ~ "Other",  
 race == "U" ~ "Unknown")) %>%   
 tbl\_summary(  
 by = group, # split table by group  
 missing = "no", # don't list missing data separately  
 label = c(sex ~ "Sex", race ~ "Race",  
 tobacco ~ "Tobacco Use",  
 age\_at\_start ~ "Age"),  
 type = c(sex, race, tobacco) ~ "categorical") %>%  
 add\_n() %>% # add column with total number of non-missing observations  
 add\_p() %>% # test for a difference between groups  
 modify\_header(label = "\*\*Variable\*\*") %>% # update the column header  
 bold\_labels() %>%   
 as\_flex\_table()

| Variable | N | NOSKIN, N = 3161 | SKIN, N = 791 | p-value2 |
| --- | --- | --- | --- | --- |
| **Sex** | 395 |  |  | 0.2 |
| Female |  | 184 (58%) | 52 (66%) |  |
| Male |  | 132 (42%) | 27 (34%) |  |
| **Race** | 395 |  |  | >0.9 |
| African-American |  | 24 (7.6%) | 6 (7.6%) |  |
| Asian or Pacific Islander |  | 2 (0.6%) | 0 (0%) |  |
| Caucasian |  | 287 (91%) | 73 (92%) |  |
| Other |  | 1 (0.3%) | 0 (0%) |  |
| Unknown |  | 2 (0.6%) | 0 (0%) |  |
| **Tobacco Use** | 395 |  |  | 0.7 |
| No |  | 180 (57%) | 43 (54%) |  |
| Yes |  | 136 (43%) | 36 (46%) |  |
| **Age** | 395 | 39 (28, 52) | 41 (30, 50) | 0.8 |
| 1n (%); Median (IQR) | | | | |
| 2Pearson's Chi-squared test; Fisher's exact test; Wilcoxon rank sum test | | | | |

## Testing FCP and CRP

Using pre, post, and paired t-tests

# decrease in FCP in skin   
  
# before, skin  
fcp %>%   
 filter(group == "SKIN") %>%   
 filter(time\_point == "BEFORE") %>%   
 summarize(mean = sprintf("%0.2f",mean(value)),  
 sd = sprintf("%0.2f",sd(value)))

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>   
## 1 394.01 165.59

# after, skin  
fcp %>%   
 filter(group == "SKIN") %>%   
 filter(time\_point == "AFTER") %>%   
 summarize(mean = sprintf("%0.2f",mean(value)),  
 sd = sprintf("%0.2f",sd(value)))

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>   
## 1 164.12 142.46

# t-test pre-post, skin  
fcp %>%   
 filter(group == "SKIN") %>%   
 rstatix::t\_test(value ~ time\_point, detailed = TRUE) %>%   
 print(width = Inf)

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2 n1  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr> <int>  
## 1 -230. 164. 394. value AFTER BEFORE 7  
## n2 statistic p df conf.low conf.high method  
## \* <int> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   
## 1 7 -2.78 0.0168 11.7 -410. -49.6 T-test  
## alternative  
## \* <chr>   
## 1 two.sided

# before, no skin  
fcp %>%   
 filter(group == "NOSKIN") %>%   
 filter(time\_point == "BEFORE") %>%   
 summarize(mean = sprintf("%0.2f",mean(value)),  
 sd = sprintf("%0.2f",sd(value)))

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>   
## 1 364.87 277.29

# after, no skin  
fcp %>%   
 filter(group == "NOSKIN") %>%   
 filter(time\_point == "AFTER") %>%   
 summarize(mean = sprintf("%0.2f",mean(value)),  
 sd = sprintf("%0.2f",sd(value)))

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>   
## 1 264.70 412.51

# t-test pre-post, no skin  
fcp %>%   
 filter(group == "NOSKIN") %>%   
 rstatix::t\_test(value ~ time\_point, detailed = TRUE) %>%   
 print(width = Inf)

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2 n1  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr> <int>  
## 1 -100. 265. 365. value AFTER BEFORE 37  
## n2 statistic p df conf.low conf.high method  
## \* <int> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   
## 1 37 -1.23 0.225 63.0 -263. 63.1 T-test  
## alternative  
## \* <chr>   
## 1 two.sided

# pct change in fcp  
  
fcp %>%   
 filter(!is.na(pct\_change\_fcp)) %>%   
 rstatix::t\_test(pct\_change\_fcp ~ group, detailed = TRUE) %>%   
 print(width = Inf)

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr>   
## 1 50.2 -10.7 -60.9 pct\_change\_fcp NOSKIN SKIN   
## n1 n2 statistic p df conf.low conf.high  
## \* <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 37 7 2.20 0.0337 41.5 4.08 96.4  
## method alternative  
## \* <chr> <chr>   
## 1 T-test two.sided

Plots for FCP and CRP

fcp %>%   
 mutate(id = rep(seq\_along(1:44), each =2)) %>%   
 ggplot(aes(x = factor(time\_point,   
 levels = c("BEFORE", "AFTER")), y = value,   
 col = group, group = id)) +  
 geom\_line(size = 0.5, alpha = 0.3) +  
 geom\_smooth(aes(group = group),   
 se = FALSE,  
 size = 2)+  
 scale\_color\_manual(values = c("indianred1", "dodgerblue")) +  
 theme\_linedraw() +  
 labs(title = "Fecal Calprotectin Measurements by Group",  
 subtitle = "Before and After 6 months of Ustekinumab Therapy\nWith Individual Patient Values Depicted in Pale Lines, and Group Averages in Bold Lines",  
 x = "Time Point", y = "")+  
 theme(legend.position = c(0.1, 0.2))

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

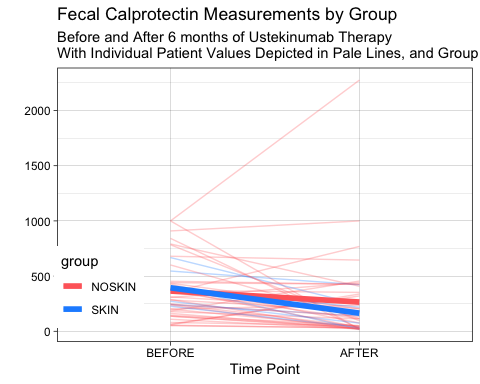
## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01



ggsave(here("fcp\_plot.tiff"))

## Saving 5 x 4 in image  
## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

ggsave(here("fcp\_plot.pdf"))

## Saving 5 x 4 in image  
## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

crp %>%   
 mutate(id = rep(seq\_along(1:139), each =2)) %>%   
 ggplot(aes(x = factor(time\_point,   
 levels = c("BEFORE", "AFTER")), y = value,   
 col = group, group = id)) +  
 geom\_line(size = 0.5, alpha = 0.3) +  
 geom\_smooth(aes(group = group),   
 se = FALSE,  
 size = 2)+  
 scale\_color\_manual(values = c("indianred1", "dodgerblue")) +  
 theme\_linedraw() +  
 labs(title = "C-Reactive Protein Measurements by Group",  
 subtitle = "Before and After 6 months of Ustekinumab Therapy\nWith Individual Patient Values Depicted in Pale Lines, and Group Averages in Bold Lines",  
 x = "Time Point", y = "")+  
 theme(legend.position = c(0.1, 0.2))

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number  
## 6.1634e-31

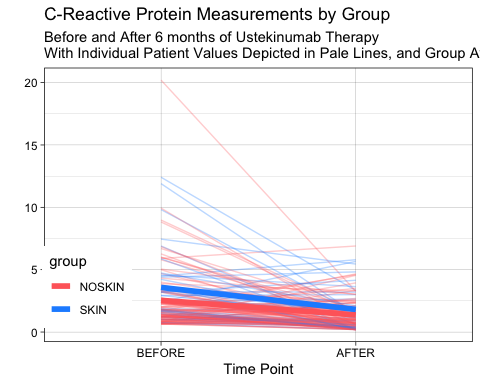
## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01



ggsave(here("crp\_plot.tiff"))

## Saving 5 x 4 in image  
## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number  
## 6.1634e-31

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

ggsave(here("fcp\_plot.pdf"))

## Saving 5 x 4 in image  
## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number  
## 6.1634e-31

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : pseudoinverse used at 0.995

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : neighborhood radius 1.005

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree,  
## parametric = parametric, : There are other near  
## singularities as well. 1.01

Using timepoint t-tests

fcp %>%   
 filter(time\_point == "AFTER") %>%   
 rstatix::t\_test(value ~ group, detailed = TRUE) %>%   
 select(group1:group2, p)

## # A tibble: 1 x 3  
## group1 group2 p  
## <chr> <chr> <dbl>  
## 1 NOSKIN SKIN 0.255

fcp %>%   
 filter(time\_point == "BEFORE") %>%   
 rstatix::t\_test(value ~ group)

## # A tibble: 1 x 8  
## .y. group1 group2 n1 n2 statistic df p  
## \* <chr> <chr> <chr> <int> <int> <dbl> <dbl> <dbl>  
## 1 value NOSKIN SKIN 37 7 -0.376 13.4 0.713

fcp %>%   
 filter(time\_point == "AFTER") %>%   
 rstatix::t\_test(pct\_change\_fcp ~ group, detailed = TRUE) ->  
fcp\_t  
  
fcp\_t %>%   
 select(estimate1:estimate2, group1:conf.high) %>%   
 flextable()

| estimate1 | estimate2 | group1 | group2 | n1 | n2 | statistic | p | df | conf.low | conf.high |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| -10.68071 | -60.91775 | NOSKIN | SKIN | 37 | 7 | 2.197063 | 0.0337 | 41.50392 | 4.076149 | 96.39793 |

Now for CRP

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>  
## 1 3.58 3.02

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>  
## 1 1.81 1.74

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2 n1  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr> <int>  
## 1 -1.76 1.81 3.58 value AFTER BEFORE 34  
## n2 statistic p df conf.low conf.high method  
## \* <int> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   
## 1 34 -2.94 0.00482 52.7 -2.96 -0.561 T-test  
## alternative  
## \* <chr>   
## 1 two.sided

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>  
## 1 2.54 2.59

## # A tibble: 1 x 2  
## mean sd   
## <chr> <chr>  
## 1 1.37 1.18

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2 n1  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr> <int>  
## 1 -1.17 1.37 2.54 value AFTER BEFORE 105  
## n2 statistic p df conf.low conf.high method  
## \* <int> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   
## 1 105 -4.22 0.0000434 146. -1.72 -0.621 T-test  
## alternative  
## \* <chr>   
## 1 two.sided

## # A tibble: 1 x 15  
## estimate estimate1 estimate2 .y. group1 group2  
## \* <dbl> <dbl> <dbl> <chr> <chr> <chr>   
## 1 22.5 -15.6 -38.1 pct\_change\_crp NOSKIN SKIN   
## n1 n2 statistic p df conf.low conf.high  
## \* <int> <int> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 105 34 1.78 0.0787 100. -2.62 47.6  
## method alternative  
## \* <chr> <chr>   
## 1 T-test two.sided

## Testing Likert for Path, Endo, Imaging

path\_likert %>%   
 rstatix::t\_test(path\_likert ~ group, detailed = TRUE)->  
path\_likert\_result  
path\_likert\_result %>%   
 select(estimate1:estimate2, group1:conf.high) %>%   
 print(width = Inf)

## # A tibble: 1 x 11  
## estimate1 estimate2 group1 group2 n1 n2 statistic  
## <dbl> <dbl> <chr> <chr> <int> <int> <dbl>  
## 1 3.24 3.73 NOSKIN SKIN 70 22 -1.85  
## p df conf.low conf.high  
## <dbl> <dbl> <dbl> <dbl>  
## 1 0.074 29.8 -1.02 0.0499

endo\_likert %>%   
 rstatix::t\_test(endo\_likert ~ group, detailed = TRUE) %>%  
 select(estimate1:estimate2, group1:conf.high) %>%   
 print(width = Inf)

## # A tibble: 1 x 11  
## estimate1 estimate2 group1 group2 n1 n2 statistic  
## <dbl> <dbl> <chr> <chr> <int> <int> <dbl>  
## 1 3.38 3.95 NOSKIN SKIN 72 22 -2.52  
## p df conf.low conf.high  
## <dbl> <dbl> <dbl> <dbl>  
## 1 0.0166 34.1 -1.05 -0.112

imaging\_likert %>%   
 rstatix::t\_test(imaging\_likert ~ group, detailed = TRUE) %>%   
 select(estimate1:estimate2, group1:conf.high) %>%   
 print(width = Inf)

## # A tibble: 1 x 11  
## estimate1 estimate2 group1 group2 n1 n2 statistic  
## <dbl> <dbl> <chr> <chr> <int> <int> <dbl>  
## 1 3.04 3.71 NOSKIN SKIN 47 14 -1.78  
## p df conf.low conf.high  
## <dbl> <dbl> <dbl> <dbl>  
## 1 0.0939 16.8 -1.47 0.127

Endo testing for ulcer resolution

# 15 of 22 resolved in SKIN  
# 25 of 59 resolved in NOSKIN  
options(digits = 3)  
  
obsfreq <- matrix(c(7, 15, 24, 25), nrow = 2, ncol =2)  
obsfreq

## [,1] [,2]  
## [1,] 7 24  
## [2,] 15 25

fisher.test(obsfreq)

##   
## Fisher's Exact Test for Count Data  
##   
## data: obsfreq  
## p-value = 0.2  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 0.143 1.556  
## sample estimates:  
## odds ratio   
## 0.491

Endo testing for histopath resolution

# 15 of 20 resolved in SKIN  
# 16 of 53 resolved in NOSKIN  
options(digits = 3)  
  
obsfreq <- matrix(c(5, 15, 37, 16), nrow = 2, ncol =2)  
obsfreq

## [,1] [,2]  
## [1,] 5 37  
## [2,] 15 16

fisher.test(obsfreq)

##   
## Fisher's Exact Test for Count Data  
##   
## data: obsfreq  
## p-value = 0.001  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 0.0358 0.5206  
## sample estimates:  
## odds ratio   
## 0.149

## Plotting Likert Results

Use likert package examples

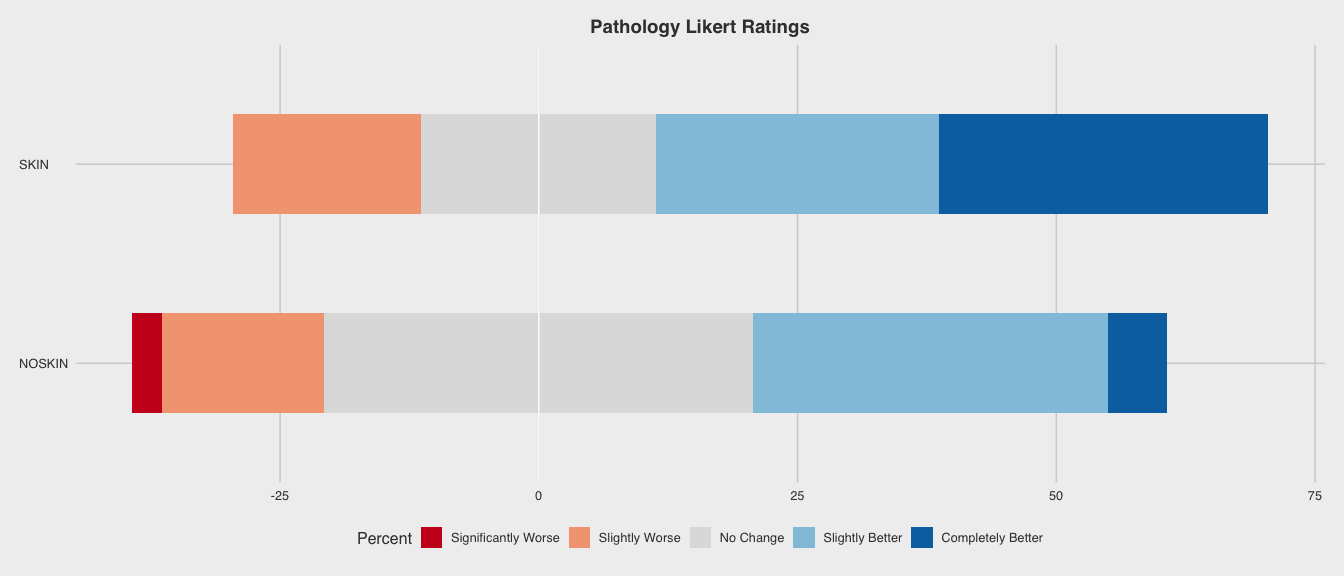
<https://github.com/jbryer/likert>

and example for ggplot here <https://jtr13.github.io/cc19/likert.html>

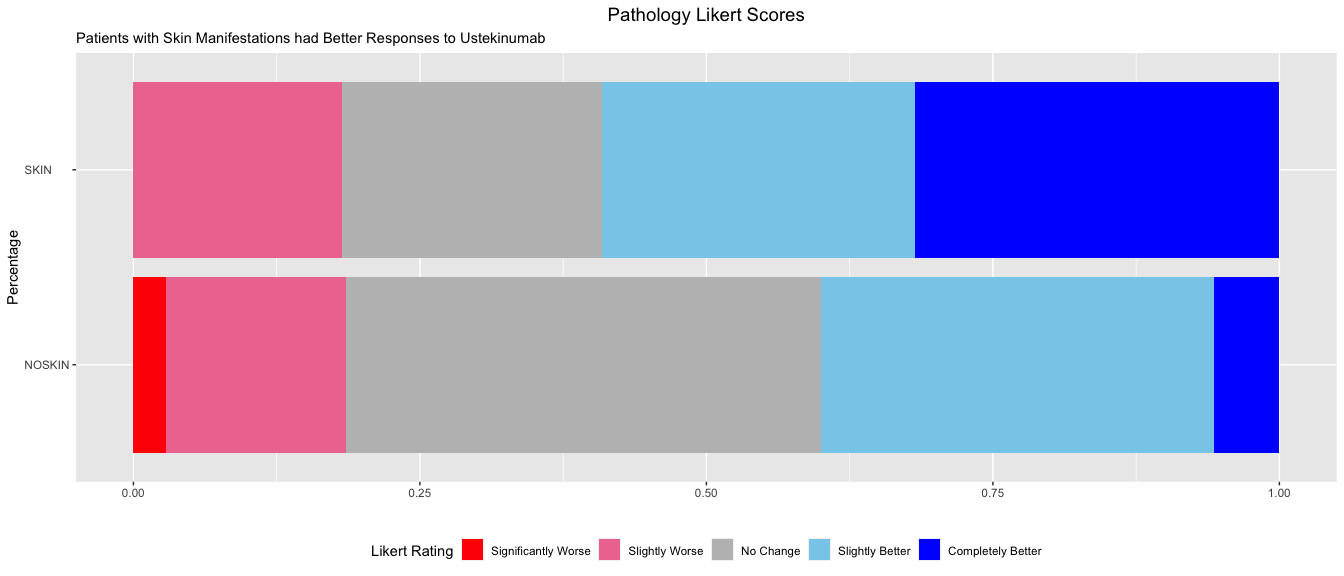
better explained here <http://rnotr.com/likert/ggplot/barometer/likert-plots/>

and generally about likert plots here <http://daydreamingnumbers.com/blog/4-ways-to-visualize-likert-scales/>

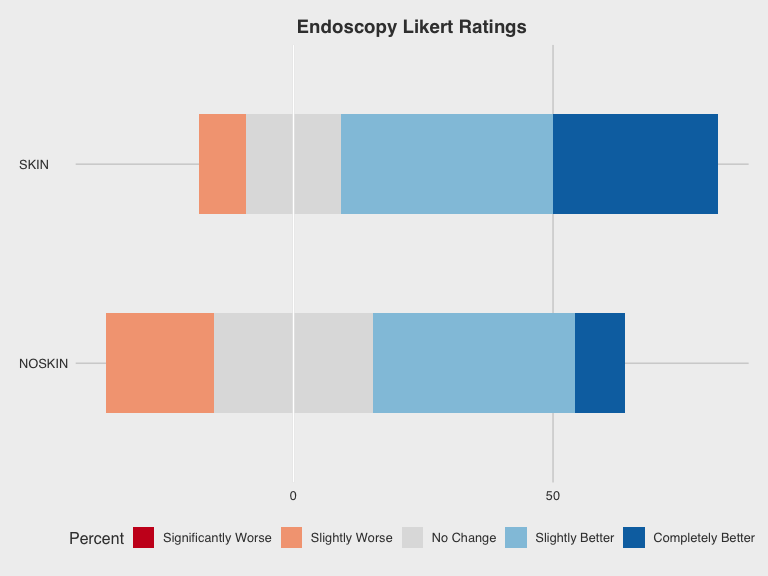
## New names:  
## \* `No Change` -> `No Change...4`  
## \* `No Change` -> `No Change...5`



## Saving 14 x 6 in image

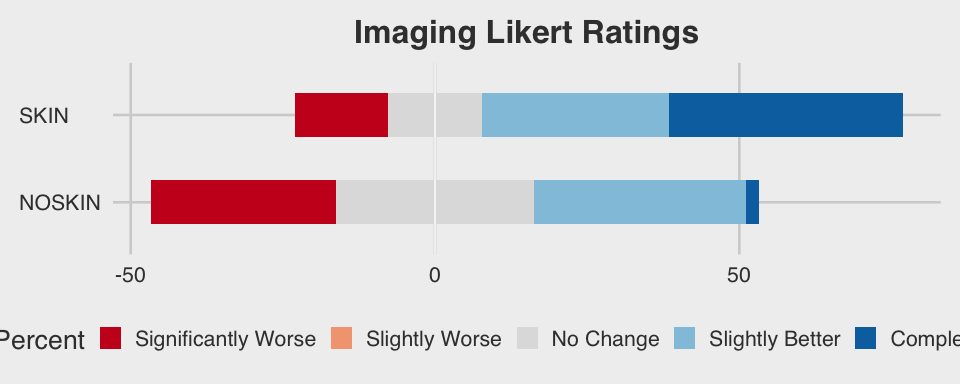


## New names:  
## \* `No Change` -> `No Change...4`  
## \* `No Change` -> `No Change...5`



## Saving 8 x 6 in image

## New names:  
## \* `No Change` -> `No Change...4`  
## \* `No Change` -> `No Change...5`



## Saving 10 x 4 in image