secure\_ibd

Peter Higgins

4/23/2020

## Secure IBD website

This analysis uses data from a registry, at covidibd.org. the current data are found [here](https://www.covidibd.org/current-data/)

Let’s web-scrape in the data Uh-oh. Turns out it is dynamically generated (probably with javascript). No static scraping for us. Might have to upgrade my hardware and use Docker/splashr in the future. For now, let’s go old-school.

### Pasting with datapasta

url <- "https://www.covidibd.org/current-data/"  
  
# paste without header  
# paste as vector  
  
  
data <- c("Overall\t770\t518\t67%\t251\t33%\t50\t6%\t41\t5%\t29\t4%\t70\t9%", "Age", "0-9 years \t3\t3\t100%\t0\t0%\t0\t0%\t0\t0%\t0\t0%\t0\t0%", "10-19 years \t28\t25\t89%\t3\t11%\t0\t0%\t0\t0%\t0\t0%\t0\t0%", "20-29 years \t155\t125\t81%\t30\t19%\t2\t1%\t1\t1%\t0\t0%\t2\t1%", "30-39 years \t155\t126\t81%\t28\t18%\t5\t3%\t2\t1%\t1\t1%\t5\t3%", "40-49 years \t152\t105\t69%\t47\t31%\t10\t7%\t7\t5%\t2\t1%\t11\t7%", "50-59 years \t120\t67\t56%\t53\t44%\t8\t7%\t8\t7%\t4\t3%\t11\t9%", "60-69 years \t89\t42\t47%\t47\t53%\t21\t24%\t17\t19%\t10\t11%\t25\t28%", "70-79 years \t35\t12\t34%\t23\t66%\t4\t11%\t4\t11%\t3\t9%\t6\t17%", ">=80 years \t30\t10\t33%\t20\t67%\t0\t0%\t2\t7%\t9\t30%\t10\t33%", "Sex", "Male \t405\t259\t64%\t146\t36%\t26\t6%\t21\t5%\t20\t5%\t38\t9%", "Female \t354\t249\t70%\t104\t29%\t24\t7%\t20\t6%\t9\t3%\t32\t9%", "Disease type", "Crohn\u0092s disease \t452\t322\t71%\t129\t29%\t26\t6%\t22\t5%\t11\t2%\t32\t7%", "Ulcerative Colitis/unspecified\t315\t194\t62%\t121\t38%\t24\t8%\t19\t6%\t18\t6%\t38\t12%", "IBD Disease", "Activity", "Remission\t457\t327\t72%\t129\t28%\t26\t6%\t25\t5%\t17\t4%\t39\t9%", "Mild\t142\t103\t73%\t39\t27%\t5\t4%\t3\t2%\t6\t4%\t8\t6%", "Moderate/Severe\t151\t76\t50%\t75\t50%\t17\t11%\t12\t8%\t5\t3%\t21\t14%", "Unknown\t20\t12\t60%\t8\t40%\t2\t10%\t1\t5%\t1\t5%\t2\t10%", "Smoking", "Current smoker\t36\t17\t47%\t19\t53%\t3\t8%\t1\t3%\t4\t11%\t5\t14%", "Non-smoker\t734\t501\t68%\t232\t32%\t47\t6%\t40\t5%\t25\t3%\t65\t9%", "Comorbidities", "0\t503\t387\t77%\t116\t23%\t15\t3%\t11\t2%\t5\t1%\t19\t4%", "1\t164\t102\t62%\t61\t37%\t14\t9%\t11\t7%\t8\t5%\t18\t11%", "2\t55\t16\t29%\t39\t71%\t8\t15%\t10\t18%\t7\t13%\t14\t25%", "3+\t48\t13\t27%\t35\t73%\t13\t27%\t9\t19%\t9\t19%\t19\t40%", "IBD medication\*", "Sulfasalazine/mesalamine\t202\t108\t53%\t94\t47%\t23\t11%\t23\t11%\t16\t8%\t36\t18%", "Budesonide \t23\t12\t52%\t11\t48%\t3\t13%\t3\t13%\t1\t4%\t3\t13%", "Oral/parenteral steroids \t60\t18\t30%\t42\t70%\t13\t22%\t9\t15%\t7\t12%\t17\t28%", "6MP/azathioprine monotherapy\t81\t49\t60%\t32\t40%\t6\t7%\t6\t7%\t2\t2%\t8\t10%", "Methotrexate monotherapy\t5\t2\t40%\t3\t60%\t0\t0%\t0\t0%\t0\t0%\t0\t0%", "Anti-TNF without 6MP/AZA/MTX \t238\t197\t83%\t40\t17%\t6\t3%\t3\t1%\t3\t1%\t7\t3%", "Anti-TNF + 6MP/AZA/MTX \t76\t44\t58%\t32\t42%\t10\t13%\t5\t7%\t2\t3%\t11\t14%", "Anti-integrin \t72\t50\t69%\t22\t31%\t3\t4%\t4\t6%\t1\t1%\t5\t7%", "IL 12/23 inhibitor \t76\t69\t91%\t7\t9%\t2\t3%\t1\t1%\t0\t0%\t2\t3%", "JAK inhibitor \t10\t8\t80%\t2\t20%\t1\t10%\t1\t10%\t1\t10%\t1\t10%", "Other IBD medication \t29\t16\t55%\t13\t45%\t2\t7%\t1\t3%\t0\t0%\t2\t7%")  
  
  
  
data2 <- read\_tsv(data, skip\_empty\_rows = TRUE,  
 col\_names = FALSE) %>%   
 filter(!is.na(X2)) %>%   
 select(X1:X3, X5, X7, X9, X11, X13) %>%   
 purrr::set\_names(c("characteristic", "total", "outpt", "hosp", "icu", "vent", "death", "icu\_vent\_death"))

## Warning: 8 parsing failures.  
## row col expected actual file  
## 2 -- 14 columns 1 columns literal data  
## 12 -- 14 columns 1 columns literal data  
## 15 -- 14 columns 1 columns literal data  
## 18 -- 14 columns 1 columns literal data  
## 19 -- 14 columns 1 columns literal data  
## ... ... .......... ......... ............  
## See problems(...) for more details.

data3 <- data2 %>% slice(25:35)  
data3$characteristic[1] <- "Sulfasalazine or mesalamine"  
data3$characteristic[3] <- "Oral or parenteral steroids"  
data3$characteristic[4] <- "6MP or Azathioprine monotherapy"  
data3$characteristic[6] <- "Anti-TNF monotherapy"  
# data3$characteristic[8] <- "Vedolizumab"  
# data3$characteristic[9] <- "Ustekinumab"  
# data3$characteristic[10] <- "Tofacitinib"  
  
chisq <- data3 %>%   
 mutate(ok = total - icu\_vent\_death) %>%   
 select(characteristic, icu\_vent\_death, ok) %>%   
 column\_to\_rownames('characteristic') %>%   
 chisq.test(correct = TRUE)

## Warning in chisq.test(., correct = TRUE): Chi-squared  
## approximation may be incorrect

chisq

##   
## Pearson's Chi-squared test  
##   
## data: .  
## X-squared = 54.499, df = 10, p-value = 0.00000003916

data4 <- as\_tibble(round(chisq$stdres, 3))  
  
exp\_icu <- chisq$expected %>% as\_tibble()   
obs\_icu <- chisq$observed %>% as\_tibble()   
cbind(data3$characteristic, exp\_icu, obs\_icu) %>%   
 purrr::set\_names("Medication", "\_ICU/Vent/Death", "\_OK",  
 "ICU/Vent/Death", "OK") %>%   
 gt() %>%   
 tab\_spanner(label = "Expected",   
 columns = vars("\_ICU/Vent/Death", "\_OK")) %>%  
 tab\_spanner(label = "Observed",   
 columns = vars("ICU/Vent/Death", "OK")) %>%  
 tab\_header(title = "ICU Outcomes of COVID-19 Infection in IBD Patients by Medication", subtitle = "Expected vs. Observed Case Counts") %>%   
 fmt\_number(columns = 2:3, decimals = 2) %>%   
 cols\_align(align = "center") %>%   
 opt\_row\_striping(row\_striping = TRUE)

ICU Outcomes of COVID-19 Infection in IBD Patients by Medication

Expected vs. Observed Case Counts

Medication

Expected

Observed

\_ICU/Vent/Death

\_OK

ICU/Vent/Death

OK

Sulfasalazine or mesalamine

21.31

180.69

36

166

Budesonide

2.43

20.57

3

20

Oral or parenteral steroids

6.33

53.67

17

43

6MP or Azathioprine monotherapy

8.55

72.45

8

73

Methotrexate monotherapy

0.53

4.47

0

5

Anti-TNF monotherapy

25.11

212.89

7

231

Anti-TNF + 6MP/AZA/MTX

8.02

67.98

11

65

Anti-integrin

7.60

64.40

5

67

IL 12/23 inhibitor

8.02

67.98

2

74

JAK inhibitor

1.06

8.94

1

9

Other IBD medication

3.06

25.94

2

27

chisq\_hosp <- data3 %>%   
 mutate(ok = total - hosp) %>%   
 select(characteristic, hosp, ok) %>%   
 column\_to\_rownames('characteristic') %>%   
 chisq.test(correct = TRUE)

## Warning in chisq.test(., correct = TRUE): Chi-squared  
## approximation may be incorrect

chisq\_hosp

##   
## Pearson's Chi-squared test  
##   
## data: .  
## X-squared = 110.23, df = 10, p-value <  
## 0.00000000000000022

data5 <- as\_tibble(round(chisq\_hosp$stdres, 3))  
  
exp\_hosp <- chisq\_hosp$expected %>% as\_tibble()   
obs\_hosp <- chisq\_hosp$observed %>% as\_tibble()   
cbind(data3$characteristic, exp\_hosp, obs\_hosp) %>%   
 purrr::set\_names("Medication", "\_Hospitalized", "\_Outpatient",  
 "Hospitalized", "Outpatient") %>%   
 gt() %>%   
 tab\_spanner(label = "Expected",   
 columns = vars("\_Hospitalized", "\_Outpatient")) %>%  
 tab\_spanner(label = "Observed",   
 columns = vars("Hospitalized", "Outpatient")) %>%  
 tab\_header(title = "Hospitalization Outcomes of COVID-19 Infection in IBD Patients by Medication", subtitle = "Expected vs. Observed Case Counts") %>%   
 fmt\_number(columns = 2:3, decimals = 2) %>%   
 cols\_align(align = "center") %>%   
 opt\_row\_striping(row\_striping = TRUE)

Hospitalization Outcomes of COVID-19 Infection in IBD Patients by Medication

Expected vs. Observed Case Counts

Medication

Expected

Observed

\_Hospitalized

\_Outpatient

Hospitalized

Outpatient

Sulfasalazine or mesalamine

69.03

132.97

94

108

Budesonide

7.86

15.14

11

12

Oral or parenteral steroids

20.50

39.50

42

18

6MP or Azathioprine monotherapy

27.68

53.32

32

49

Methotrexate monotherapy

1.71

3.29

3

2

Anti-TNF monotherapy

81.33

156.67

40

198

Anti-TNF + 6MP/AZA/MTX

25.97

50.03

32

44

Anti-integrin

24.61

47.39

22

50

IL 12/23 inhibitor

25.97

50.03

7

69

JAK inhibitor

3.42

6.58

2

8

Other IBD medication

9.91

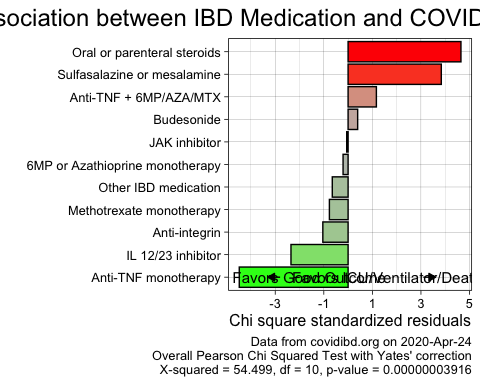
19.09

13

16

## Including Plots for icu/vent/death

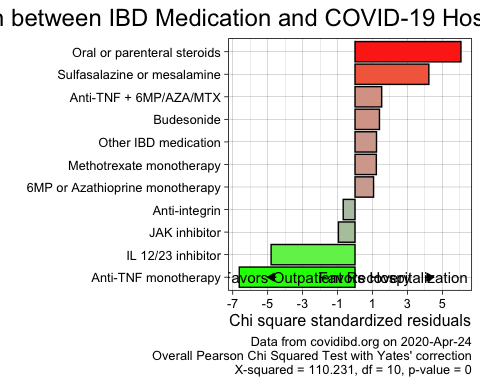
data4 %>%   
 cbind(data3$characteristic) %>%   
 purrr::set\_names('icu\_vent\_death', 'ok', 'medication') %>%   
 mutate(medication = fct\_reorder(medication, icu\_vent\_death)) %>%   
 ggplot(aes(x= icu\_vent\_death,  
 fill = icu\_vent\_death,  
 y = medication)) +  
 geom\_col(color = "black") +  
 labs(x = "Chi square standardized residuals",  
 title = "Association between IBD Medication and COVID-19 Outcomes",  
 caption = glue("Data from covidibd.org on 2020-Apr-24\nOverall Pearson Chi Squared Test with Yates' correction\nX-squared = {round(chisq$statistic[[1]][1], 3)}, df = {chisq$parameter[[1]][1]}, p-value = {round(chisq$p.value[[1]][1], 11)}"),  
 y = "") +  
 annotate("text", x = 1.6, y = 1, size =4,  
 label = "Favors ICU/Ventilator/Death") +  
 annotate("text", x = -1.6, y = 1, size =4,  
 label = "Favors Good Outcome") +  
 theme\_linedraw(base\_size = 12)+  
 theme(legend.position = "none") +  
 annotate(geom = "segment", x = -2.8, y = 1,  
 xend = -3.3, yend = 1,  
 arrow = arrow(length = unit(0.2, "cm"), type = "closed")) +  
 annotate(geom = "segment", x = 3.1, y = 1,  
 xend = 3.6, yend = 1,  
 arrow = arrow(length = unit(0.2, "cm"), type = "closed")) +  
 scale\_fill\_gradient2(low = "green", mid="grey", high="red") +  
 theme(plot.title.position = "plot") +  
 theme(plot.title = element\_text(hjust = 0.2, size = 18)) +  
 scale\_x\_continuous(breaks = seq(-7,7,2))



ggsave("figures/covidibd.png", height = 4, width =9, units = "in", device = "png")

## Including Plot for Hospitalization

data5 %>%   
 cbind(data3$characteristic) %>%   
 purrr::set\_names('hosp', 'ok', 'medication') %>%   
 mutate(medication = fct\_reorder(medication, hosp)) %>%   
 ggplot(aes(x= hosp,  
 fill = hosp,  
 y = medication)) +  
 geom\_col(color = "black") +  
 labs(x = "Chi square standardized residuals",  
 title = "Association between IBD Medication and COVID-19 Hospitalization",  
 caption = glue("Data from covidibd.org on 2020-Apr-24\nOverall Pearson Chi Squared Test with Yates' correction\nX-squared = {round(chisq\_hosp$statistic[[1]][1], 3)}, df = {chisq\_hosp$parameter[[1]][1]}, p-value = {round(chisq\_hosp$p.value[[1]][1], 11)}"),  
 y = "") +  
 annotate("text", x = 2.2, y = 1, size =4,  
 label = "Favors Hospitalization") +  
 annotate("text", x = -2.2, y = 1, size =4,  
 label = "Favors Outpatient Recovery") +  
 theme\_linedraw(base\_size = 12)+  
 theme(legend.position = "none") +  
 annotate(geom = "segment", x = -4.5, y = 1,  
 xend = -5, yend = 1,  
 arrow = arrow(length = unit(0.2, "cm"), type = "closed")) +  
 annotate(geom = "segment", x = 4, y = 1,  
 xend = 4.5, yend = 1,  
 arrow = arrow(length = unit(0.2, "cm"), type = "closed")) +  
 scale\_fill\_gradient2(low = "green", mid="grey", high="red") +  
 theme(plot.title.position = "plot") +  
 theme(plot.title = element\_text(hjust = 0.5, size = 18)) +  
 scale\_x\_continuous(breaks = seq(-7,7,2))



ggsave("figures/covidibdhosp.png", height = 4, width =9, units = "in", device = "png")