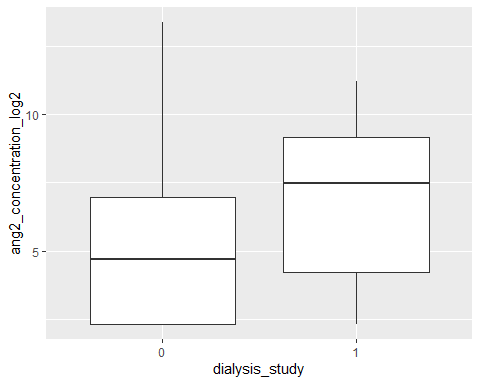
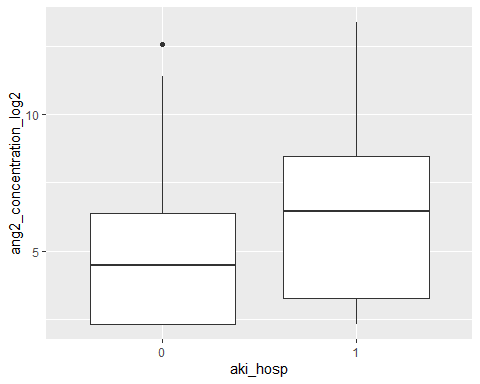
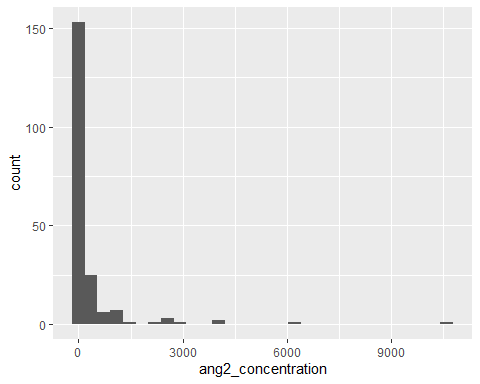
CLOVERS\_Paper

Jordan Lo

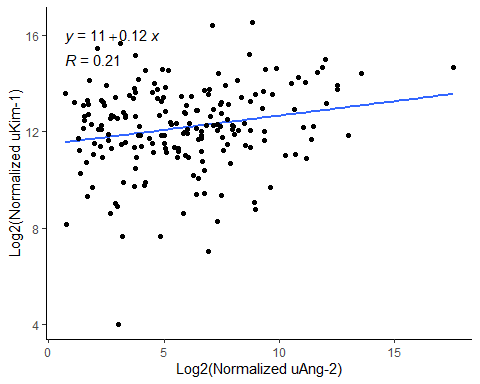
2023-09-08

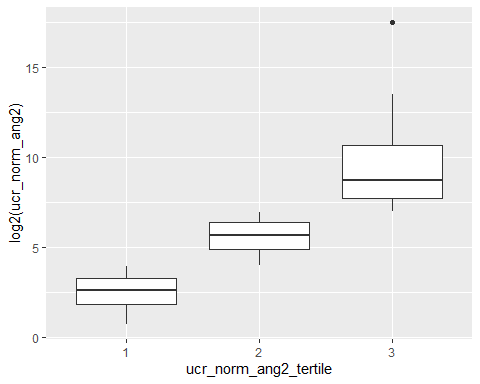
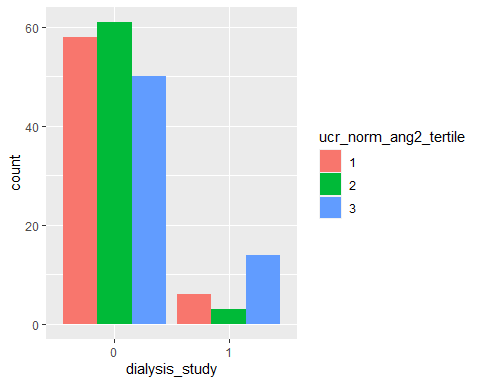
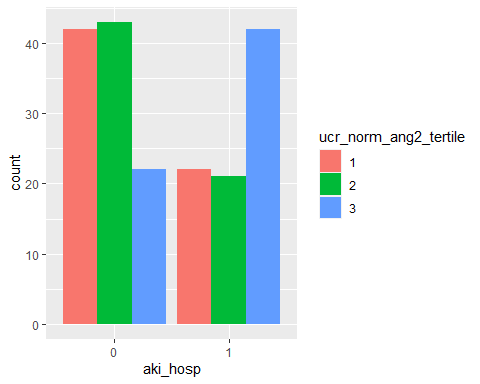
There are 201 patients that remain after merging the two datasets and filtering out those with clinical data and Ang-2 biomarker data available

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

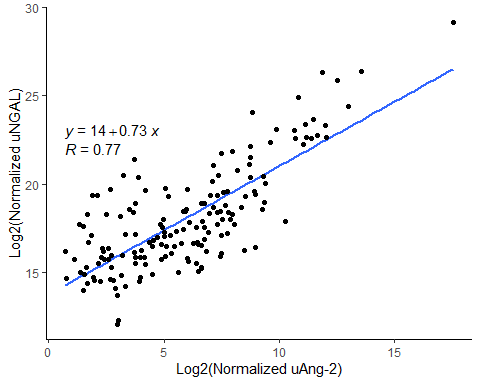


## `geom\_smooth()` using formula = 'y ~ x'

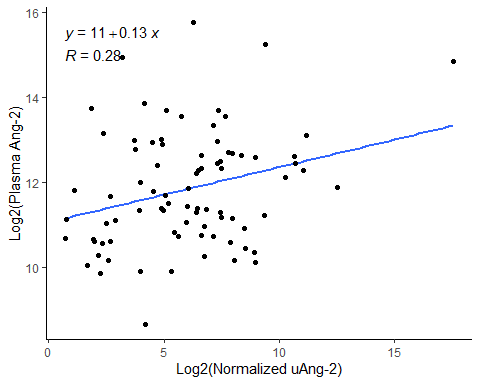
 Looking into the data by tertiles based on Ang-2 normalized to urine creatinine levels



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



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



# finds overlap between ang2 and normalized ang 2 tertiles  
ang2\_tertile <- clinical\_lab\_msd %>%  
 dplyr::select(c("study\_id",  
 "ang2\_tertile")) %>%  
 filter(ang2\_tertile == 3)  
  
norm\_ang2\_tertile <- clinical\_lab\_msd %>%  
 dplyr::select(c("study\_id",  
 "ucr\_norm\_ang2\_tertile")) %>%  
 filter(ucr\_norm\_ang2\_tertile == 3)  
  
tertile\_overlap <- inner\_join(ang2\_tertile,  
 norm\_ang2\_tertile,  
 by = "study\_id")  
  
length(tertile\_overlap$study\_id)

## [1] 53

We see very high overlap between the highest Ang-2 Concentration tertiles and the highest normalized Ang-2 tertiles. 53/64 patients are in the highest tertiles for both.

# runs a poisson regression with covariate adjustment to find relative risk  
  
# aki\_hosp  
# regression on lowest/middle tertile with adjustment  
result <- glm(aki\_hosp ~ tertile\_rr + age + male + bmi + mv + covid\_confirmed + myrace,  
 data = rr\_mt\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ----------------------------------------------------------------  
## (Intercept) | 0.89 | 0.15 | [0.65, 1.24] | -0.68 | 0.497  
## tertile rr | 0.95 | 0.06 | [0.84, 1.08] | -0.76 | 0.446  
## age | 1.00 | 1.73e-03 | [1.00, 1.01] | 1.47 | 0.142  
## male | 1.01 | 0.07 | [0.88, 1.16] | 0.17 | 0.866  
## bmi | 1.00 | 2.75e-03 | [1.00, 1.01] | 1.75 | 0.080  
## mv | 1.08 | 0.07 | [0.96, 1.22] | 1.25 | 0.213  
## covid confirmed | 1.15 | 0.08 | [1.01, 1.31] | 2.10 | 0.036  
## myrace | 0.98 | 0.03 | [0.93, 1.04] | -0.61 | 0.542

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# unadjusted model  
result <- glm(aki\_hosp ~ tertile\_rr,   
 data = rr\_mt\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ---------------------------------------------------------  
## (Intercept) | 1.34 | 0.06 | [1.23, 1.47] | 6.69 | < .001  
## tertile rr | 0.99 | 0.06 | [0.87, 1.12] | -0.19 | 0.852

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# regression on lowest/highest tertile with adjustment   
result <- glm(aki\_hosp ~ tertile\_rr + age + male + bmi + mv + covid\_confirmed + myrace,   
 data = rr\_ht\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ----------------------------------------------------------------  
## (Intercept) | 1.21 | 0.21 | [0.86, 1.70] | 1.09 | 0.274  
## tertile rr | 1.20 | 0.07 | [1.07, 1.35] | 3.19 | 0.001  
## age | 1.00 | 1.77e-03 | [0.99, 1.00] | -1.07 | 0.287  
## male | 1.06 | 0.07 | [0.93, 1.20] | 0.86 | 0.391  
## bmi | 1.00 | 2.96e-03 | [1.00, 1.01] | 1.30 | 0.193  
## mv | 1.06 | 0.07 | [0.93, 1.21] | 0.91 | 0.364  
## covid confirmed | 1.06 | 0.06 | [0.94, 1.19] | 0.96 | 0.339  
## myrace | 0.99 | 0.03 | [0.94, 1.04] | -0.57 | 0.569

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# unadjusted   
result <- glm(aki\_hosp ~ tertile\_rr,   
 data = rr\_ht\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## --------------------------------------------------------  
## (Intercept) | 1.34 | 0.06 | [1.23, 1.47] | 6.69 | < .001  
## tertile rr | 1.23 | 0.07 | [1.10, 1.38] | 3.68 | < .001

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

We do see that even without adjustments, being in the top tertile of the dataset predicts higher chances of developing aki in the hospital.

# death\_dialysis  
# # regression on lowest/middle tertile without adjustment  
result <- glm(death\_dialysis ~ tertile\_rr,   
 data = rr\_mt\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ---------------------------------------------------------  
## (Intercept) | 0.27 | 0.06 | [0.18, 0.40] | -6.38 | < .001  
## tertile rr | 1.24 | 0.34 | [0.72, 2.11] | 0.77 | 0.441

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# regression on lowest/middle tertile with adjustment  
result <- glm(death\_dialysis ~ tertile\_rr + age + male + bmi + mv + covid\_confirmed + myrace,   
 data = rr\_mt\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ----------------------------------------------------------------  
## (Intercept) | 0.07 | 0.07 | [0.01, 0.46] | -2.80 | 0.005  
## tertile rr | 1.09 | 0.33 | [0.60, 1.99] | 0.30 | 0.765  
## age | 1.01 | 9.51e-03 | [0.99, 1.03] | 1.36 | 0.175  
## male | 1.09 | 0.37 | [0.55, 2.13] | 0.24 | 0.810  
## bmi | 1.01 | 0.01 | [0.98, 1.04] | 0.44 | 0.661  
## mv | 1.30 | 0.38 | [0.73, 2.32] | 0.90 | 0.366  
## covid confirmed | 1.44 | 0.47 | [0.76, 2.72] | 1.11 | 0.266  
## myrace | 0.92 | 0.12 | [0.71, 1.20] | -0.62 | 0.534

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# regression on lowest/highest tertile without adjustment  
result <- glm(death\_dialysis ~ tertile\_rr,   
 data = rr\_ht\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## ---------------------------------------------------------  
## (Intercept) | 0.27 | 0.06 | [0.18, 0.40] | -6.38 | < .001  
## tertile rr | 2.06 | 0.49 | [1.29, 3.28] | 3.05 | 0.002

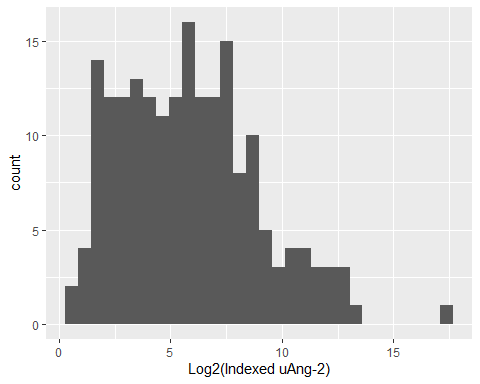
##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

# regression on lowest/highest tertile with adjustment  
result <- glm(death\_dialysis ~ tertile\_rr + age + male + bmi + mv + covid\_confirmed + myrace,   
 data = rr\_ht\_lt,  
 family = "poisson")  
  
model\_parameters(result,  
 exponentiate = TRUE,  
 vcov = "HC",  
 vcov\_args = list(type = "HC0"))

## Parameter | IRR | SE | 95% CI | z | p  
## -----------------------------------------------------------------  
## (Intercept) | 0.10 | 0.07 | [0.03, 0.39] | -3.37 | < .001  
## tertile rr | 1.95 | 0.46 | [1.22, 3.10] | 2.81 | 0.005   
## age | 1.01 | 7.22e-03 | [0.99, 1.02] | 1.25 | 0.210   
## male | 0.78 | 0.17 | [0.51, 1.20] | -1.13 | 0.259   
## bmi | 1.01 | 0.01 | [0.99, 1.03] | 0.99 | 0.324   
## mv | 1.01 | 0.23 | [0.64, 1.58] | 0.03 | 0.979   
## covid confirmed | 1.54 | 0.39 | [0.94, 2.52] | 1.72 | 0.086   
## myrace | 1.03 | 0.10 | [0.85, 1.25] | 0.35 | 0.728

##   
## Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
## using a Wald z-distribution approximation.

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



##   
## Grubbs test for one outlier  
##   
## data: clinical\_lab\_msd$ucr\_norm\_ang2  
## G = 13.7178332, U = 0.0096116, p-value < 2.2e-16  
## alternative hypothesis: highest value 189128.581175237 is an outlier

CHR\_287 is an outlier according to the grubbs tes (p-value < 2.2 e-16)