What data we are working with?

10788 mice

Look at excel sheet we sent to collaborators for details

CDU lymphoma nontumor tumor

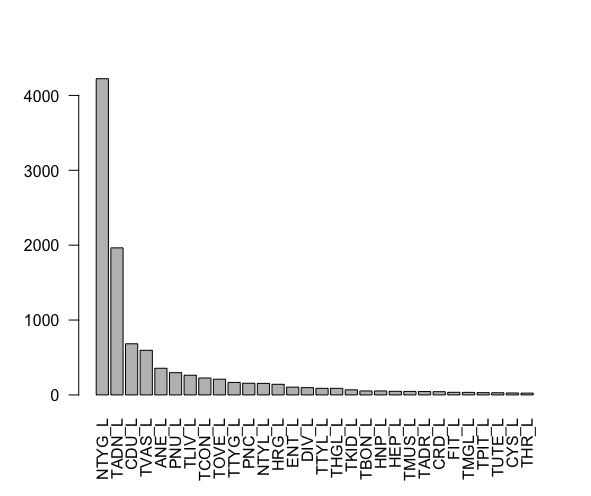
683 4691 1586 3828

CDU lymphoma nontumor tumor

6.33 43.48 14.70 35.48

Which causes of death are most common? Total, within tumors, within non-tumors, within lymphoma

Top 30 causes of death total:



Deaths caused by:

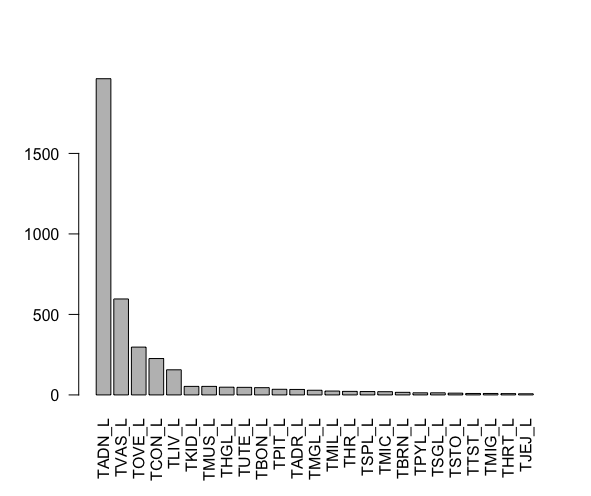
35.48% - tumor

21.03% - non-tumor

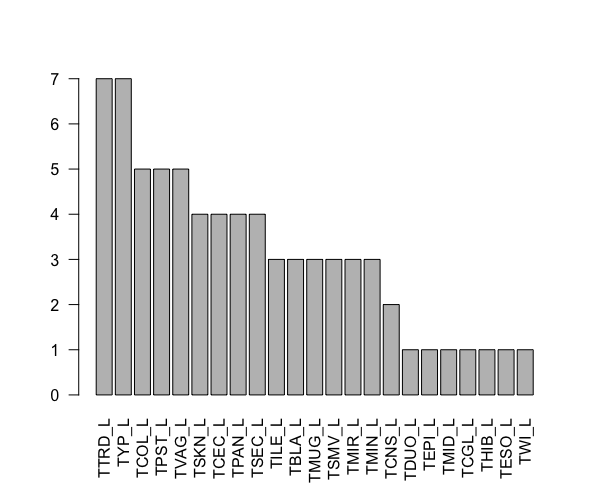
43.48% - lymphoma

- cause of death unknown

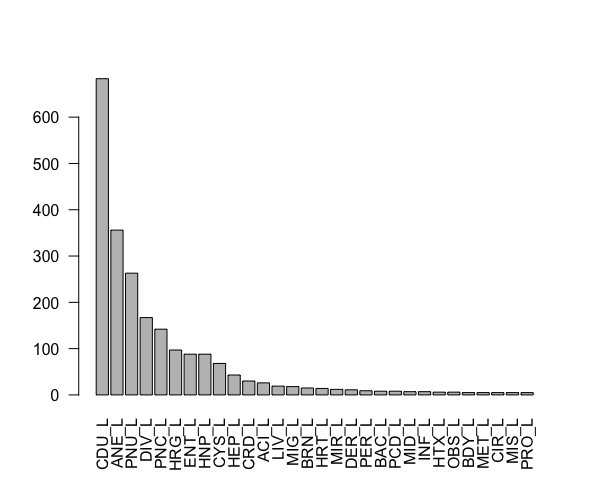
Top 25 tumors that caused death (48 tumors as cause of death)



TADN – lung, TVAS – vascular, TOVE – ovarian ,TCON – connective tissue tumors other than lymphoreticular and vascular tumors, TLIV – liver and gallbladder tumors

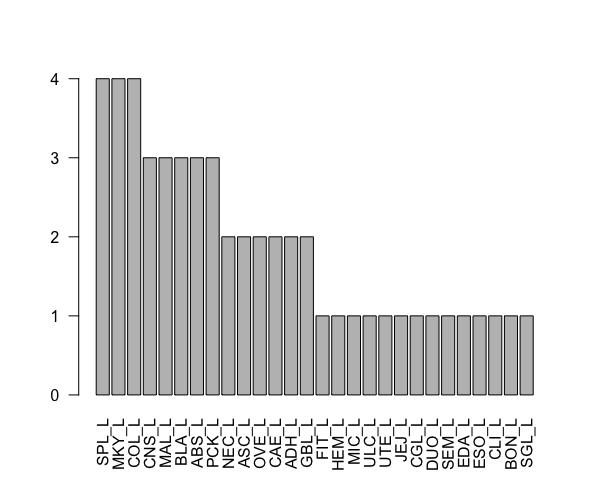
Bottom causes of death from tumors:

Top 30 causes of death – not from tumors (58 total types):

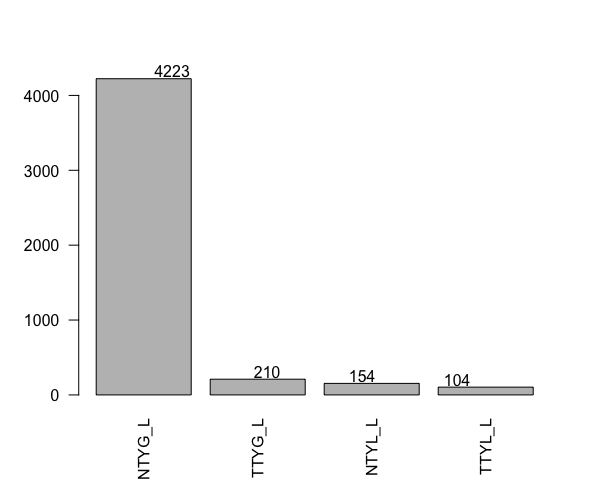


CDU – Cause of death undetermined, ANE – anemia, PNU – pneumonia, DIV – diverticulum, PNC – pneumonitits, HRG – hemorrhage, ENT – enteritis, HNP – hydronephorsis, CYS – ovarian cyst

Bottom causes of death – not tumors

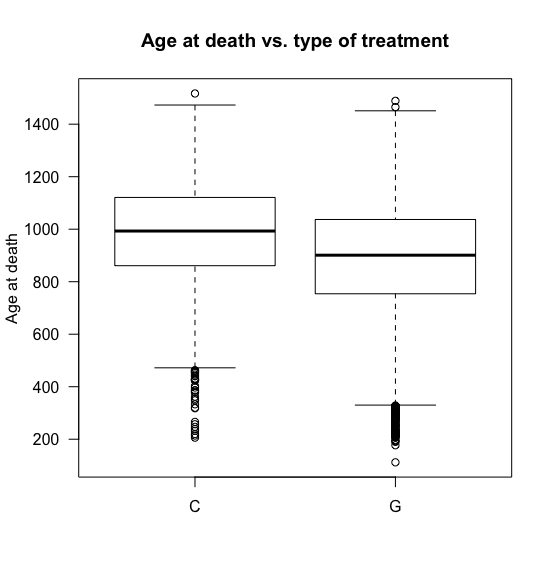


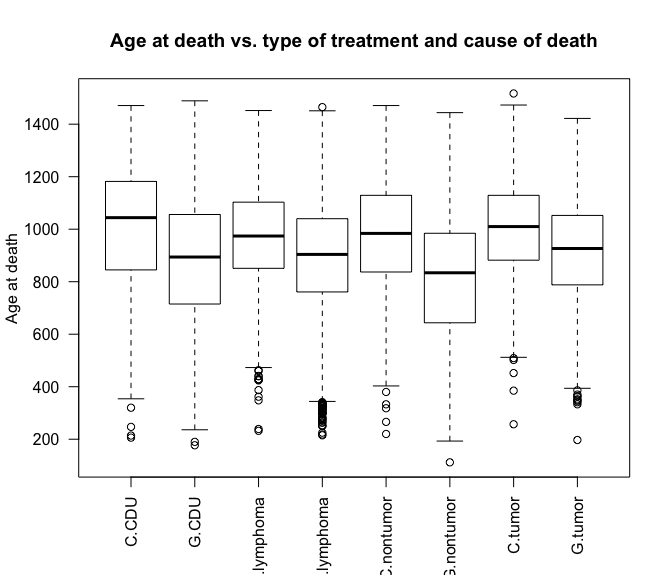
All 4 types of lymphoma (non-thymic/thymic and generalized/localized)

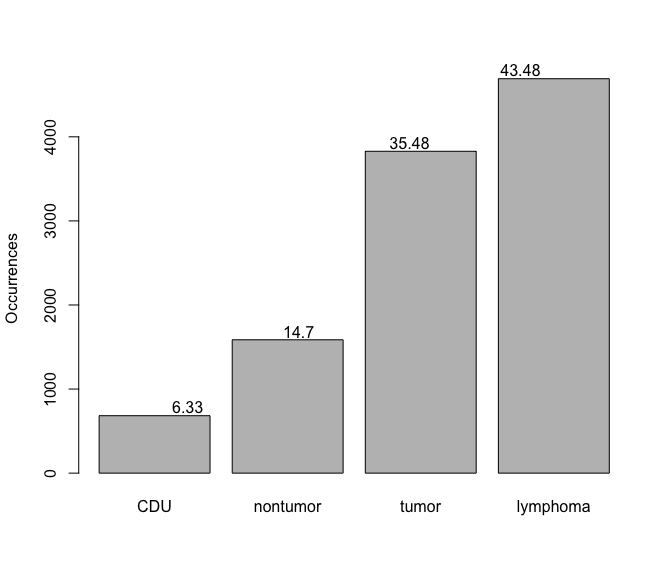


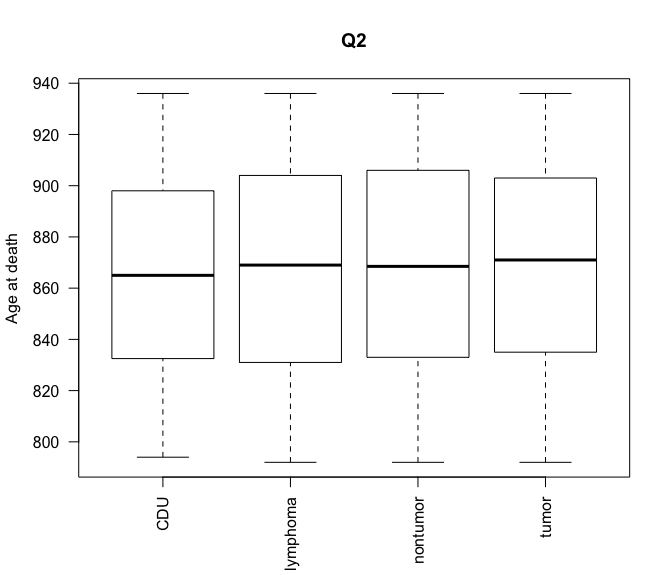
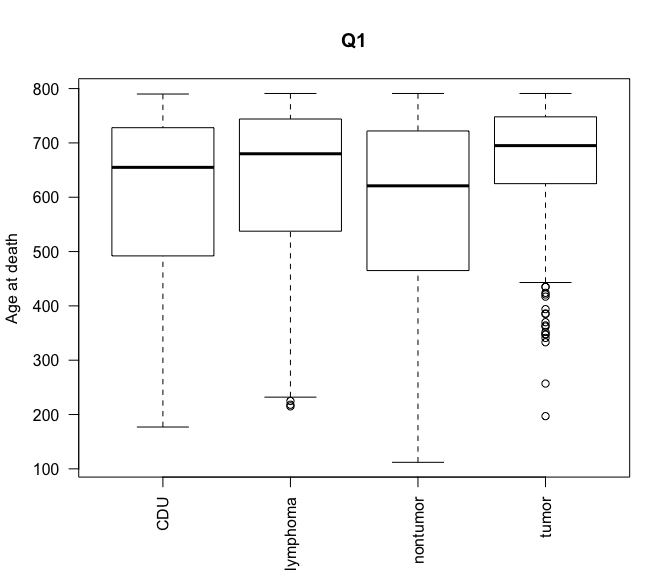
SEPARATE BY QUARTERS

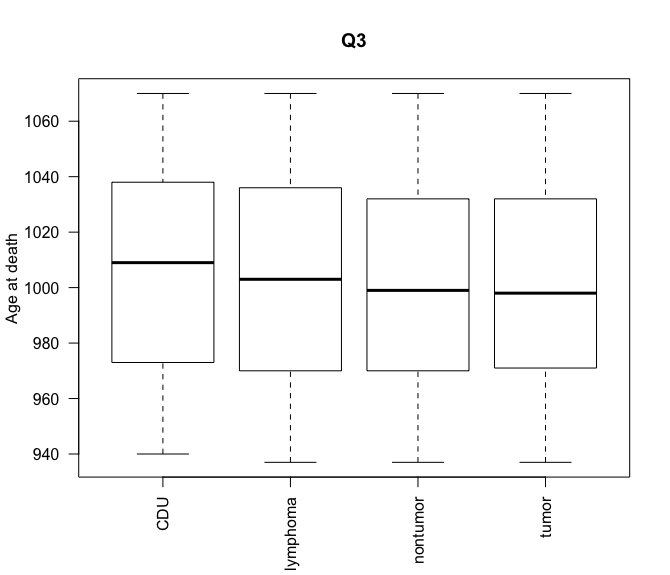
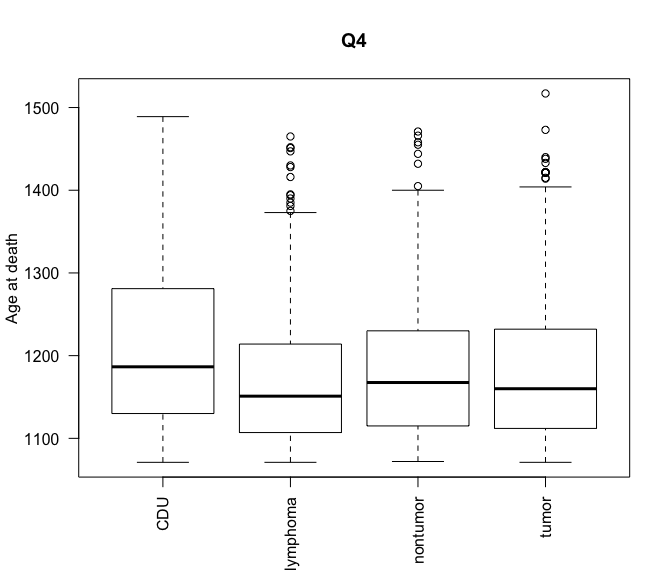


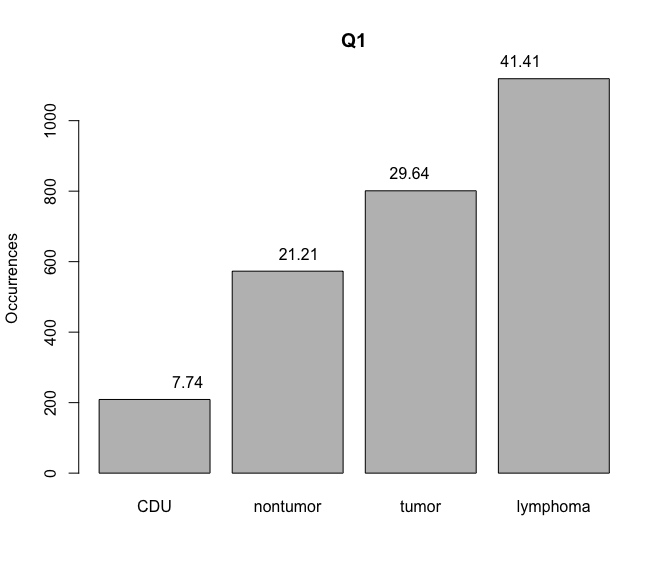
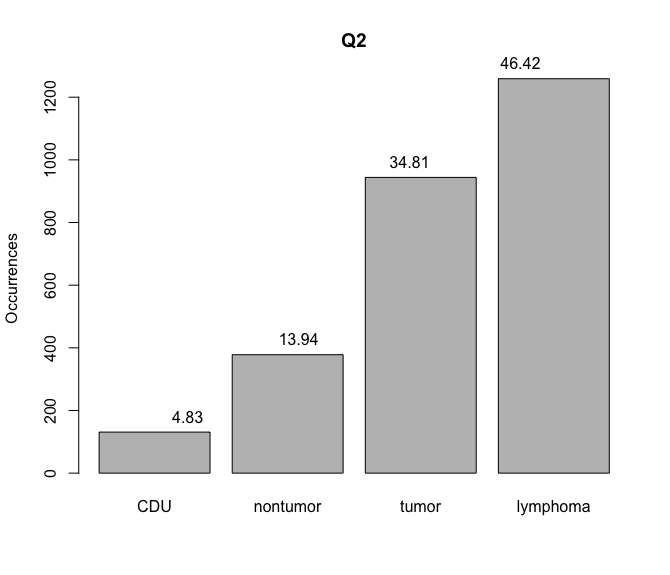


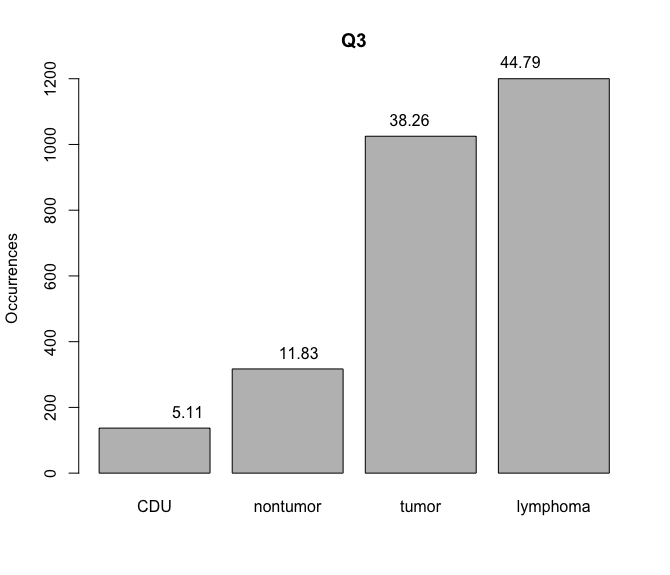


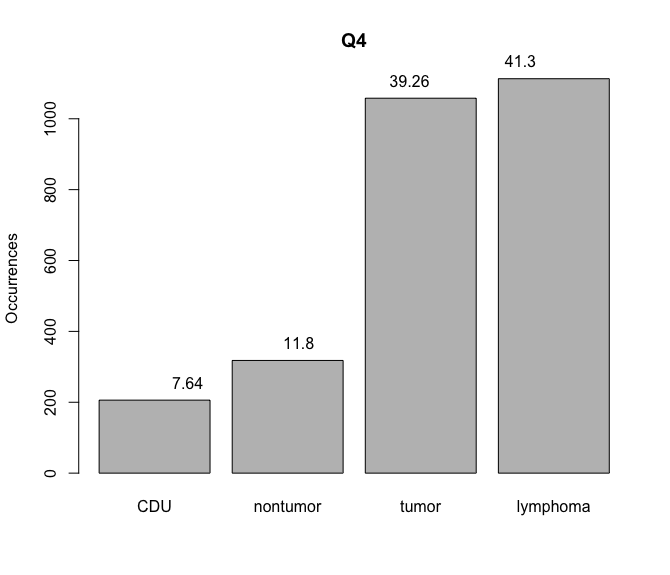


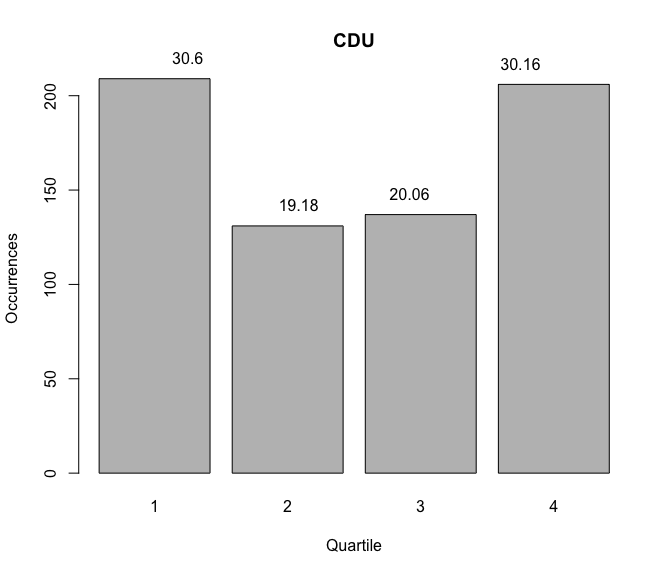


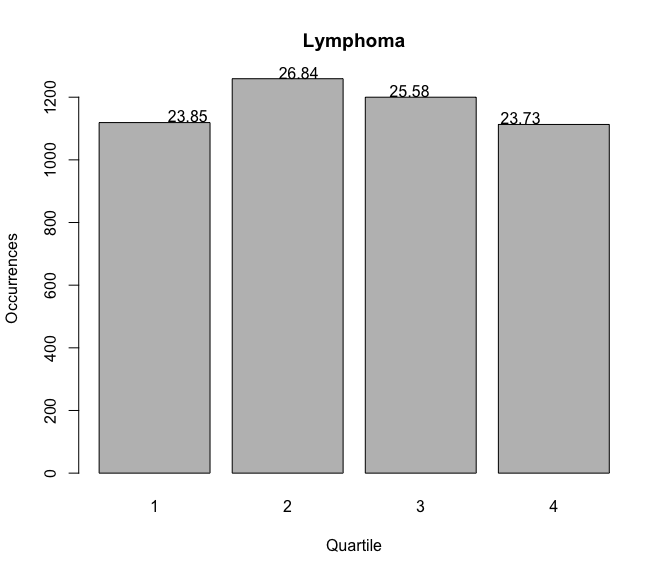
 

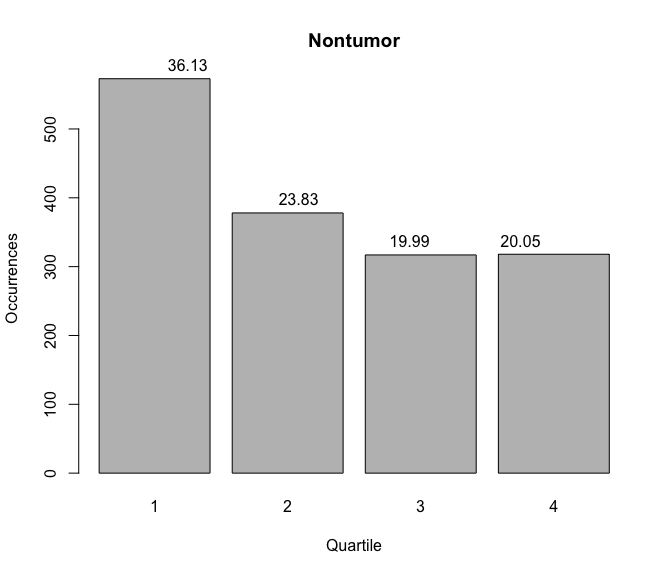
 

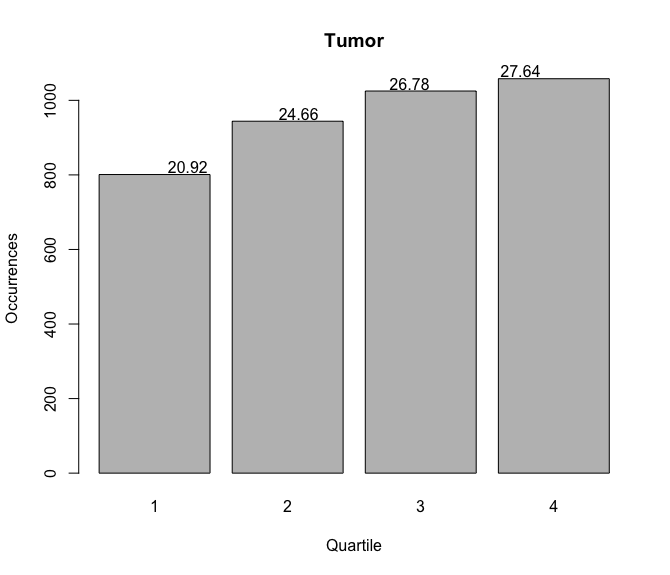


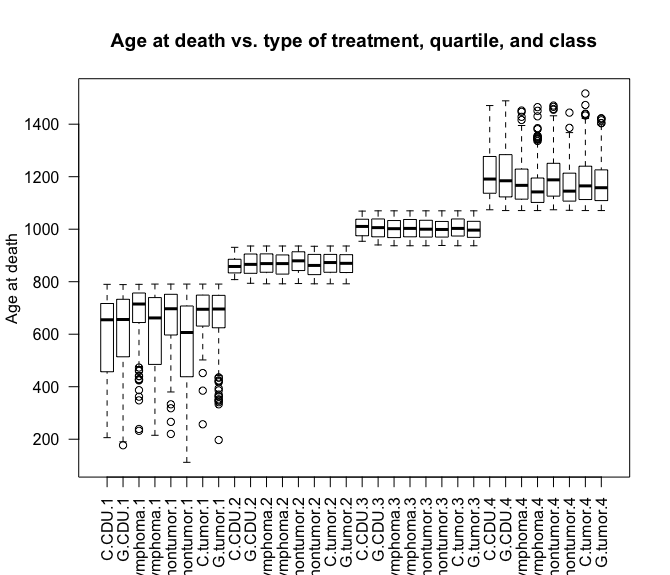












It looks like lymphoma and nontumor actually have the greatest impact on age at death depending on whether animals were irradiated or not.

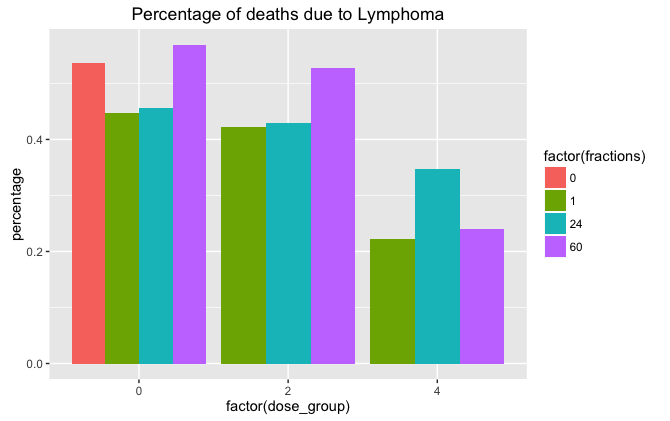
\*\*\*\*\*\*\*\*\*\*\*\*\*\* Now within Lymphoma \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

| **fractions** | **Dose-group** | **N per group** |
| --- | --- | --- |
| 0 | 0 | 164 |
| 1 | 0 | 1199 |
| 1 | 2 | 239 |
| 1 | 4 | 54 |
| 24 | 0 | 629 |
| 24 | 2 | 7 |
| 24 | 4 | 378 |
| 60 | 0 | 541 |
| 60 | 2 | 167 |
| 60 | 4 | 25 |

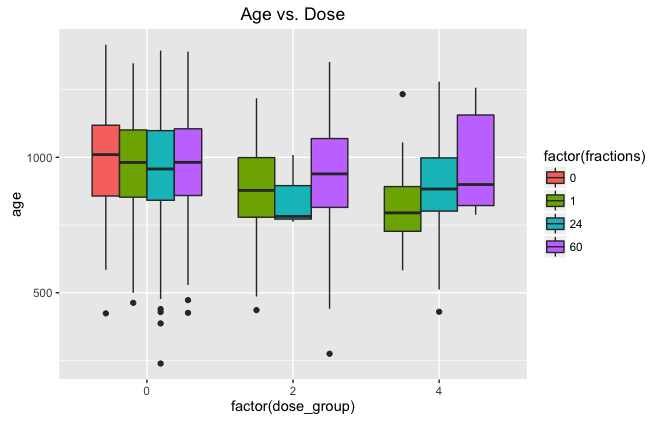
Above is table for all animals with this set of fractions and dose group

| **fractions** | | **dose\_group** | | | **n\_group** | | **percentage** | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | |  | | |  |
| 0 | 0 | | 88 | | 0.5365854 | | |
| 1 | 0 | | 537 | | 0.4478732 | | |
| 1 | 2 | | 101 | | 0.4225941 | | |
| 1 | 4 | | 12 | | 0.2222222 | | |
| 24 | 0 | | 287 | | 0.4562798 | | |
| 24 | 2 | | 3 | | 0.4285714 | | |
| 24 | 4 | | 131 | | 0.3465608 | | |
| 60 | 0 | | 308 | | 0.5693161 | | |
| 60 | 2 | | 88 | | 0.5269461 | | |
| 60 | 4 | | 6 | | 0.2400000 | | |

Above is table for animals that died of thymic lymphoma with this set of conditions



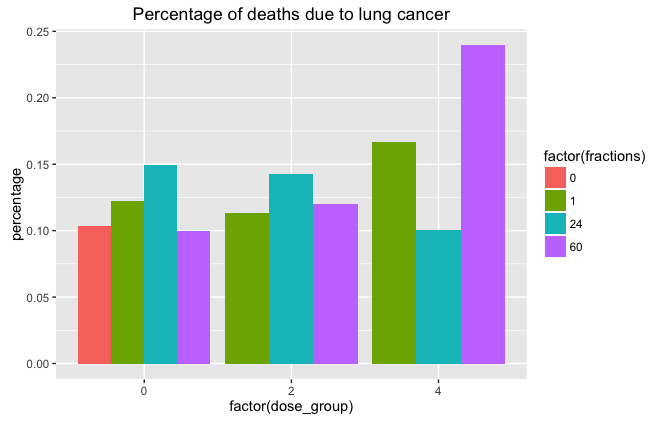
There are many fewer deaths from lymphoma at higher total doses. The 24 fraction data has a very low sample size… must figure out how to determine what cutoff for sample size should be.

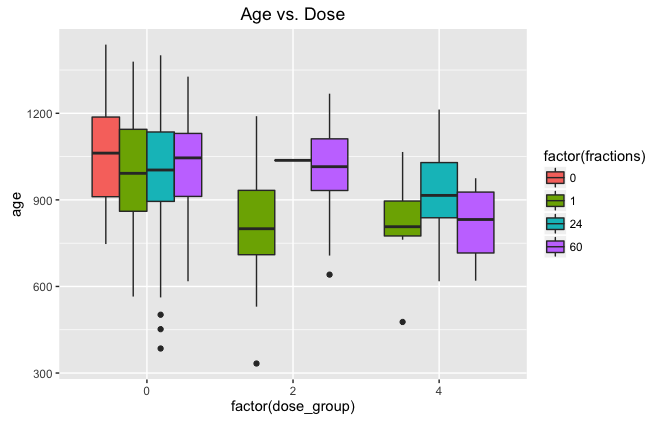


For controls, the age at death is pretty consisted regardless of number of fractions. At higher doses, fractionation helps increase the age at death, but there’s a clear drop in age at death for higher doses. Again, small sample set for 24 fractions that messes up expected trends.

\*\*\*\*\*\*\*\*\*\*\*\*\*\* Now within lung cancer \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

| **Fractions** | **Dose\_group** | **N\_group** | **percentage** |  |
| --- | --- | --- | --- | --- |
| 0 | 0 | 17 | 0.10365854 |
| 1 | 0 | 147 | 0.12260217 |
| 1 | 2 | 27 | 0.11297071 |
| 1 | 4 | 9 | 0.16666667 |
| 24 | 0 | 94 | 0.14944356 |
| 24 | 2 | 1 | 0.14285714 |
| 24 | 4 | 38 | 0.10052910 |
| 60 | 0 | 54 | 0.09981516 |
| 60 | 2 | 20 | 0.11976048 |
| 60 | 4 | 6 | 0.24000000 |





Start getting into more statistical test… begin with correlation matrix below.

age dose\_group fractions

age 1.00000000 -0.2235162 0.06127531

dose\_group -0.22351617 1.0000000 0.11925435

fractions 0.06127531 0.1192544 1.00000000

The largest correlation is between age and dose\_group… the correlation between dose group and fractions is small even with uneven sample sizes. Does this mean I don’t need to be overly concerned with type of Anova?

What question do I want to ask?

I want to control for fractions when interpreting the effect of dose\_group and I want to control for dose\_group when interpreting the effect of fractions… Type III Anova. <https://stackoverflow.com/questions/2624688/2-way-anova-on-unbalanced-dataset>

Using all data from above and lung cancer specific death…

Call:

lm(formula = age ~ factor(dose\_group) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-616.31 -120.69 9.69 135.69 399.69

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1047.06 45.23 23.151 < 2e-16 \*\*\*

factor(dose\_group)2 -108.52 29.95 -3.624 0.000327 \*\*\*

factor(dose\_group)4 -108.82 29.12 -3.737 0.000213 \*\*\*

factor(fractions)1 -65.37 47.52 -1.376 0.169694

factor(fractions)24 -45.74 48.75 -0.938 0.348620

factor(fractions)60 -10.47 50.43 -0.208 0.835676

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 186.5 on 407 degrees of freedom

Multiple R-squared: 0.06977, Adjusted R-squared: 0.05834

F-statistic: 6.105 on 5 and 407 DF, p-value: 1.814e-05

> Anova(uw.model, type = "III")

Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 18637647 1 535.9792 < 2.2e-16 \*\*\*

factor(dose\_group) 874342 2 12.5721 5.035e-06 \*\*\*

factor(fractions) 204804 3 1.9632 0.1189

Residuals 14152642 407

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

--- here we see that dose\_group has a significant impact on age, but the fractions do not. Are the small sample sizes of a couple of groups the issue? Remove and repeat.

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 18637647 1 535.0295 < 2.2e-16 \*\*\*

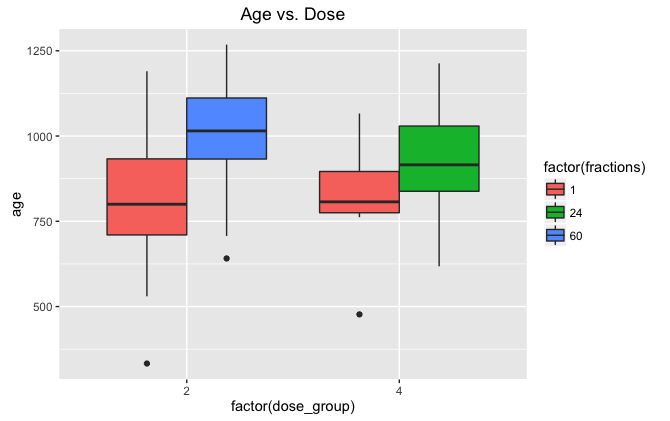
factor(dose\_group) 765496 2 10.9875 2.263e-05 \*\*\*

factor(fractions) 263064 3 2.5173 0.0578 .

Residuals 13933922 400

--- we got closer to having a significant effect for fractions, but still not quite enough. Why is this the case when the graph shows a nice trend?

Issue – fractions should matter for doses 2 and 4 Gy and should NOT matter for dose 0 Gy… check those separately? Yes.



lm(formula = age ~ factor(dose\_group) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-480.44 -81.95 -4.26 109.49 376.56

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 813.44 32.56 24.982 < 2e-16 \*\*\*

factor(dose\_group)4 18.89 65.12 0.290 0.772446

factor(fractions)24 87.43 62.72 1.394 0.166775

factor(fractions)60 188.06 49.92 3.767 0.000294 \*\*\*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 169.2 on 90 degrees of freedom

Multiple R-squared: 0.1526, Adjusted R-squared: 0.1244

F-statistic: 5.404 on 3 and 90 DF, p-value: 0.001827

> Anova(uw.model, type = "III")

Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 17865680 1 624.0873 < 2.2e-16 \*\*\*

factor(dose\_group) 2408 1 0.0841 0.7724461

factor(fractions) 461942 2 8.0683 0.0005984 \*\*\*

Residuals 2576421 90

Findings – factors are significant and dose\_group is not significant when dose\_group = 0 is removed. Factors are NOT significant and dose\_group IS significant when dose\_group = 0 is included. For fractions, this is what we would hope to see. There shouldn’t be a difference for controls, but there should be with higher doses. It’s interesting that the 2 and 4 Gy doses don’t have a significant difference between them…

Run regression and ANOVA with just dose = 0, no difference between number of fractions… good that there’s no significance between fractions.



Call:

lm(formula = age ~ factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-603.76 -132.14 9.21 135.18 412.24

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1047.06 46.00 22.761 <2e-16 \*\*\*

factor(fractions)1 -51.92 48.59 -1.068 0.286

factor(fractions)24 -58.30 49.99 -1.166 0.244

factor(fractions)60 -25.23 52.75 -0.478 0.633

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 189.7 on 308 degrees of freedom

Multiple R-squared: 0.007044, Adjusted R-squared: -0.002627

F-statistic: 0.7283 on 3 and 308 DF, p-value: 0.5357

Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

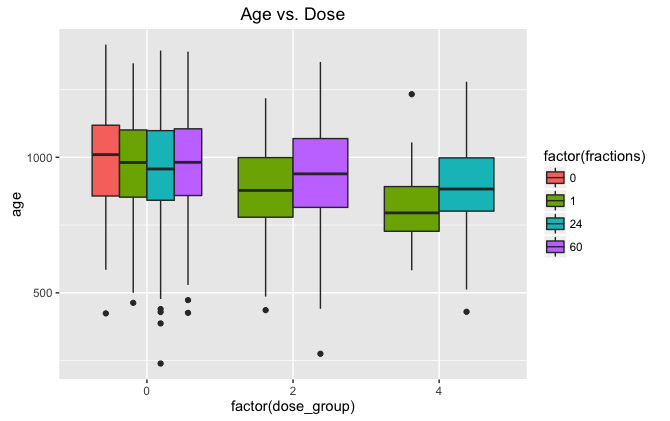
(Intercept) 18637647 1 518.0853 <2e-16 \*\*\*

factor(fractions) 78604 3 0.7283 0.5357

Residuals 11080020 308

Conclusions – Fractions and dose are significant indicators of age at death and the number of fractions INCREASES age and the dose DECREASES age.

Now go back and repeat statistical analysis for lymphoma (ANOVAs)



Including all data except for low n.

lm(formula = age ~ factor(dose\_group) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-727.69 -119.17 -0.17 128.56 435.26

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1001.34 18.71 53.529 < 2e-16 \*\*\*

factor(dose\_group)2 -66.74 14.15 -4.715 2.63e-06 \*\*\*

factor(dose\_group)4 -78.51 17.40 -4.512 6.91e-06 \*\*\*

factor(fractions)1 -31.17 20.06 -1.554 0.120

factor(fractions)24 -34.65 21.29 -1.628 0.104

factor(fractions)60 -17.86 20.92 -0.854 0.393

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 175.5 on 1546 degrees of freedom

Multiple R-squared: 0.03265, Adjusted R-squared: 0.02952

F-statistic: 10.44 on 5 and 1546 DF, p-value: 7.27e-10

Anova Table (Type III tests)

Response: age

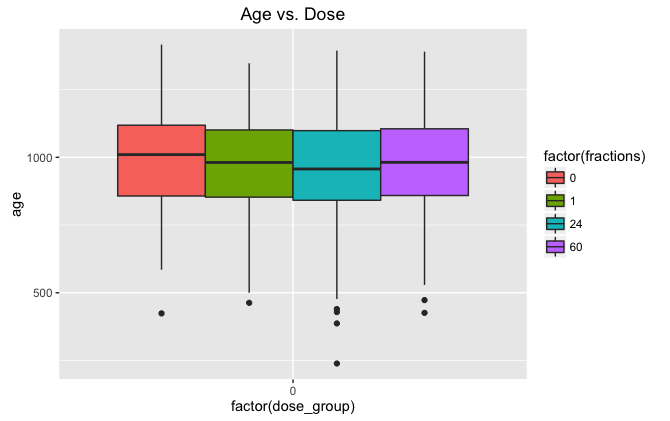
Sum Sq Df F value Pr(>F)

(Intercept) 88236158 1 2865.3527 < 2.2e-16 \*\*\*

factor(dose\_group) 1292260 2 20.9822 1.021e-09 \*\*\*

factor(fractions) 124057 3 1.3429 0.2588

Residuals 47607787 1546



Just for dose = 0 Gy.

Call:

lm(formula = age ~ factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-725.07 -121.81 2.27 127.46 429.93

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1001.34 18.81 53.227 <2e-16 \*\*\*

factor(fractions)1 -26.61 20.30 -1.311 0.1900

factor(fractions)24 -37.27 21.50 -1.733 0.0833 .

factor(fractions)60 -23.37 21.33 -1.095 0.2736

Anova Table (Type III tests)

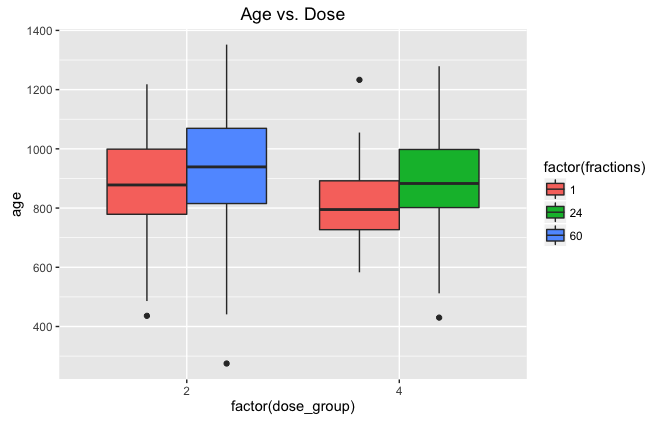
Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 88236158 1 2833.07 <2e-16 \*\*\*

factor(fractions) 98109 3 1.05 0.3695

Residuals 37872349 1216



Now only use 2Gy and 4Gy samples to check for significance of fractionation.

lm(formula = age ~ factor(dose\_group) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-661.02 -101.16 -7.13 112.55 415.98

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 886.63 17.02 52.083 <2e-16 \*\*\*

factor(dose\_group)4 -57.55 52.24 -1.102 0.2714

factor(fractions)24 64.83 51.60 1.256 0.2099

factor(fractions)60 49.39 24.95 1.980 0.0486 \*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 171.1 on 328 degrees of freedom

Multiple R-squared: 0.02014, Adjusted R-squared: 0.01118

F-statistic: 2.248 on 3 and 328 DF, p-value: 0.08266

Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

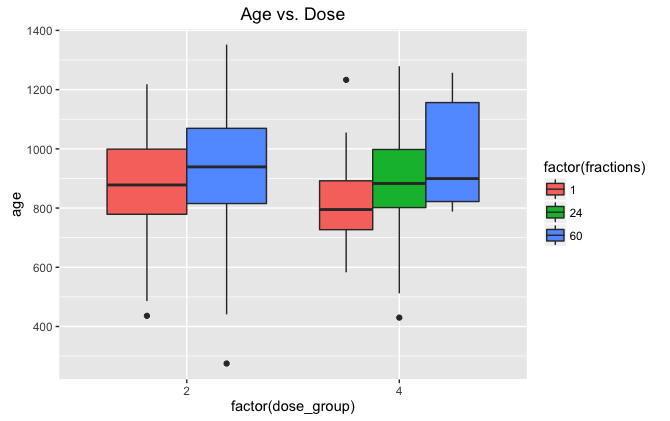
(Intercept) 79398045 1 2712.6311 < 2e-16 \*\*\*

factor(dose\_group) 35524 1 1.2137 0.27141

factor(fractions) 160906 2 2.7487 0.06549 .

Residuals 9600479 328

Fractions are not quite significant… maybe include one of the samples of fractions that has a small n? The larger one is big enough?



Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 81716161 1 2768.7165 < 2e-16 \*\*\*

factor(dose\_group) 8303 1 0.2813 0.59618

factor(fractions) 171253 2 2.9012 0.05634 .

Residuals 9857708 334

So close, but still not significant…

dose\_group fractions n\_group avg\_age

<dbl> <int> <int> <dbl>

1 2 1 101 886.6337

2 2 60 88 936.0227

3 4 1 12 829.0833

4 4 24 131 893.9084

5 4 60 6 980.3333

Fractions are still not significant… think more on this later, but this might just be how it is. Move on to different groups of doses and dose rates.

**Now use bigger ranges for dose groups –**

**0,**

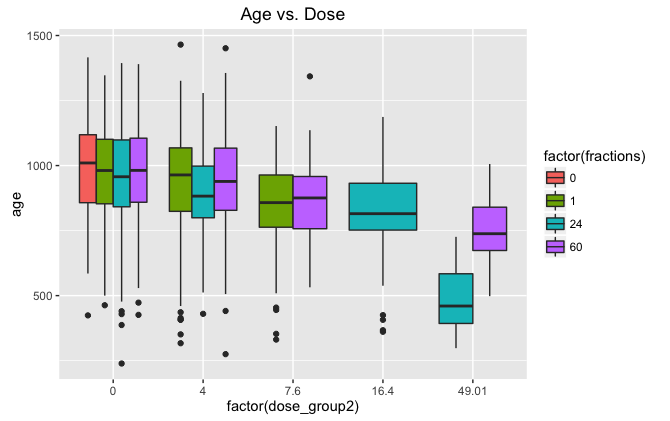
**0 🡪 1.5**

**1.5 🡪 4**

**4 🡪 7.6**

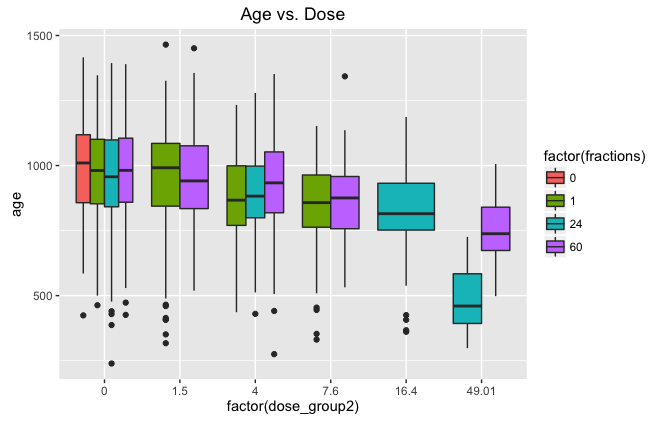
**7.6 🡪 16.4**

**16.4 🡪 inf**



Got a summary for this info and made some comparisons… realized it’s important to separate the 4Gy dose group more… new summary below.

| **fractions** | | **dose\_rate** | | **total\_dose** | **dose\_group2** | | **n** | | **mean\_age** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | |  |
| **1** | 0 | 0.000000 | 0.00 | 0.00 | 88 | | 1001.3409 |
| **2** | 1 | 0.000000 | 0.00 | 0.00 | 537 | | 974.7263 |
| **3** | 1 | 1.078500 | 21.57 | 1.50 | 194 | | 954.0000 |
| **4** | 1 | 2.157500 | 43.15 | 1.50 | 116 | | 963.4138 |
| **5** | 1 | 4.315500 | 86.31 | 1.50 | 129 | | 960.4419 |
| **6** | 1 | 6.857000 | 137.14 | 1.50 | 1 | | 964.0000 |
| **7** | 1 | 9.877500 | 197.55 | 4.00 | 19 | | 887.6316 |
| **8** | 1 | 10.300000 | 206.00 | 4.00 | 82 | | 886.4024 |
| **9** | 1 | 19.995000 | 399.90 | 4.00 | 12 | | 829.0833 |
| **10** | 1 | 27.283500 | 545.67 | 7.60 | 56 | | 878.6964 |
| **11** | 1 | 27.285000 | 545.70 | 7.60 | 24 | | 754.2083 |
| **12** | 24 | 0.000000 | 0.00 | 0.00 | 287 | | 964.0697 |
| **13** | 24 | 0.182889 | 197.55 | 4.00 | 3 | | 851.0000 |
| **14** | 24 | 0.370222 | 399.90 | 4.00 | 131 | | 893.9084 |
| **15** | 24 | 0.716444 | 773.91 | 16.40 | 91 | | 835.7033 |
| **16** | 24 | 0.851556 | 919.68 | 16.40 | 2 | | 396.0000 |
| **17** | 24 | 1.703780 | 1840.30 | 49.01 | 3 | | 326.0000 |
| **18** | 24 | 2.388440 | 2579.70 | 49.01 | 46 | | 497.5217 |
| **19** | 60 | 0.000000 | 0.00 | 0.00 | 308 | | 977.9740 |
| **20** | 60 | 0.083000 | 100.00 | 1.50 | 315 | | 950.9556 |
| **21** | 60 | 0.148000 | 399.90 | 4.00 | 6 | | 980.3333 |
| **22** | 60 | 0.166500 | 200.00 | 4.00 | 88 | | 936.0227 |
| **23** | 60 | 0.250000 | 300.00 | 4.00 | 42 | | 910.8571 |
| **24** | 60 | 0.375000 | 450.00 | 7.60 | 41 | | 894.0000 |
| **25** | 60 | 0.500000 | 600.00 | 7.60 | 45 | | 842.3111 |
| **26** | 60 | 0.681111 | 1839.40 | 49.01 | 55 | | 749.1091 |



New summary:

|  |  |  |  |
| --- | --- | --- | --- |
| 0 | 0.00 | 88 | 1001.3409 |
| **2** | 1 | 0.00 | 537 | 974.7263 |
| **3** | 1 | 1.50 | 440 | 958.3932 |
| **4** | 1 | 4.00 | 113 | 880.5221 |
| **5** | 1 | 7.60 | 80 | 841.3500 |
| **6** | 24 | 0.00 | 287 | 964.0697 |
| **7** | 24 | 4.00 | 134 | 892.9478 |
| **8** | 24 | 16.40 | 93 | 826.2473 |
| **9** | 24 | 49.01 | 49 | 487.0204 |
| **10** | 60 | 0.00 | 308 | 977.9740 |
| **11** | 60 | 1.50 | 315 | 950.9556 |
| **12** | 60 | 4.00 | 136 | 930.2059 |
| **13** | 60 | 7.60 | 86 | 866.9535 |
| **14** | 60 | 49.01 | 55 | 749.1091 |

For each set of fractions, as you increase dose group, average age goes down.

| dose\_group2 | | | fractions | | n | | mean\_age | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | |  |  |  | |
| 1 | 0.00 | | 0 | 88 | 1001.3409 | |
| 2 | 0.00 | | 1 | 537 | 974.7263 | |
| 3 | 0.00 | | 24 | 287 | 964.0697 | |
| 4 | 0.00 | | 60 | 308 | 977.9740 | |
| 5 | 1.50 | | 1 | 440 | 958.3932 | |
| 6 | 1.50 | | 60 | 315 | 950.9556 | |
| 7 | 4.00 | | 1 | 113 | 880.5221 | |
| 8 | 4.00 | | 24 | 134 | 892.9478 | |
| 9 | 4.00 | | 60 | 136 | 930.2059 | |
| 10 | 7.60 | | 1 | 80 | 841.3500 | |
| 11 | 7.60 | | 60 | 86 | 866.9535 | |
| 12 | 16.40 | | 24 | 93 | 826.2473 | |
| 13 | 49.01 | | 24 | 49 | 487.0204 | |
| 14 | 49.01 | | 60 | 55 | 749.1091 | |

Except for 0Gy and 1.5Gy, as you increase the number of fractions for a set dose group, the age goes up. I’d expect the differences for 0Gy and 1.5Gy are not significant, but must do statistical tests now.

Make model:

lm(formula = age ~ factor(dose\_group2) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-710.74 -114.45 2.75 120.55 515.55

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1001.341 18.556 53.964 < 2e-16 \*\*\*

factor(dose\_group2)1.5 -26.225 8.538 -3.072 0.00215 \*\*

factor(dose\_group2)4 -69.061 10.380 -6.653 3.45e-11 \*\*\*

factor(dose\_group2)7.6 -128.314 14.706 -8.725 < 2e-16 \*\*\*

factor(dose\_group2)16.4 -123.492 20.112 -6.140 9.45e-10 \*\*\*

factor(dose\_group2)49.01 -345.236 18.176 -18.994 < 2e-16 \*\*\*

factor(fractions)1 -25.669 19.656 -1.306 0.19170

factor(fractions)24 -51.601 20.567 -2.509 0.01217 \*

factor(fractions)60 -11.663 19.999 -0.583 0.55982

---

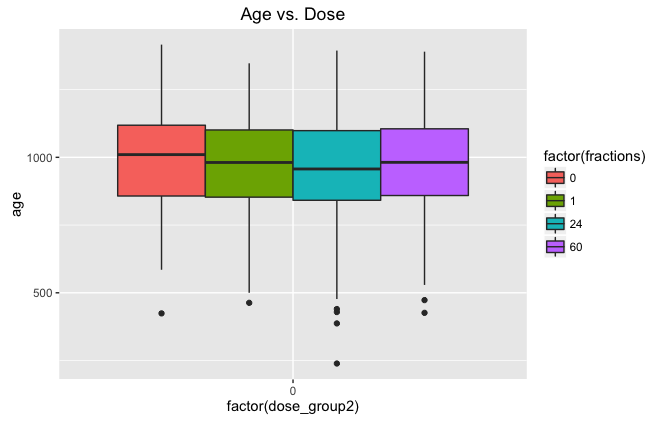
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 174.1 on 2712 degrees of freedom

Multiple R-squared: 0.1569, Adjusted R-squared: 0.1544

F-statistic: 63.08 on 8 and 2712 DF, p-value: < 2.2e-16

Dose\_group2 = 0 is the reference and 0 fractions is the reference. As dose group goes up, the age decreases (except for 16.4 compared to 7.6… it’s close). Issue with factors – 0 factors only works for 0Gy, so of course it shows that age goes down with increased fractions. Separate out data again.



Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 88236158 1 2833.07 <2e-16 \*\*\*

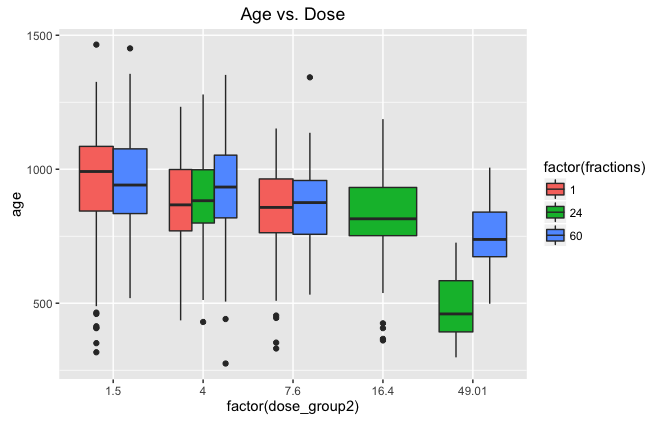
factor(fractions) 98109 3 1.05 0.3695

Residuals 37872349 1216

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Factors not significant for 0Gy.



lm(formula = age ~ factor(dose\_group2) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-659.68 -108.11 7.89 118.52 516.89

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 948.11 7.53 125.919 < 2e-16 \*\*\*

factor(dose\_group2)4 -30.63 12.28 -2.495 0.012694 \*

factor(dose\_group2)7.6 -102.41 14.75 -6.941 5.78e-12 \*\*\*

factor(dose\_group2)16.4 -61.63 24.95 -2.470 0.013614 \*

factor(dose\_group2)49.01 -303.20 19.95 -15.196 < 2e-16 \*\*\*

factor(fractions)24 -60.24 17.43 -3.457 0.000562 \*\*\*

factor(fractions)60 17.21 10.07 1.709 0.087665 .

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 171.7 on 1494 degrees of freedom

Multiple R-squared: 0.2092, Adjusted R-squared: 0.206

F-statistic: 65.86 on 6 and 1494 DF, p-value: < 2.2e-16

Anova Table (Type III tests)

Response: age

Sum Sq Df F value Pr(>F)

(Intercept) 467457940 1 15855.530 < 2.2e-16 \*\*\*

factor(dose\_group2) 8105840 4 68.735 < 2.2e-16 \*\*\*

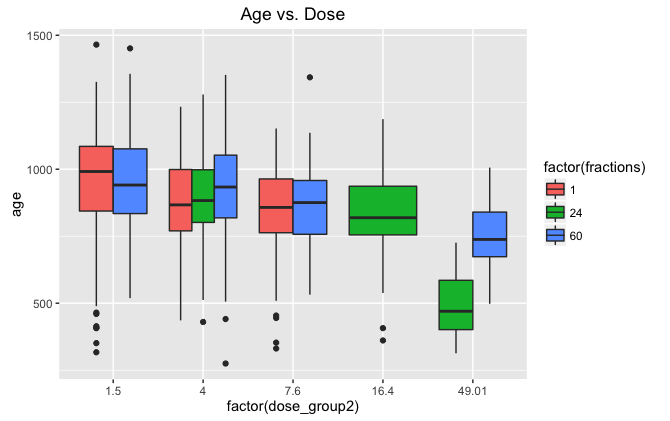
factor(fractions) 662731 2 11.239 1.429e-05 \*\*\*

Residuals 44046597 1494

Dose group and fractions are both significant, but the 24 fractions shows a significant decrease in age compared to 1 fraction. For the 24 fractions, there are two conditions with very low n – they also have very low average ages at death. Remove these and see what happens to data.

New data:

| **fractions** | | **dose\_group2** | | **dose\_rate** | **n\_group** | | **avg\_age** |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| **1** | 1 | 1.50 | 1.078500 | 194 | 954.0000 |
| **2** | 1 | 1.50 | 2.157500 | 116 | 963.4138 |
| **3** | 1 | 1.50 | 4.315500 | 129 | 960.4419 |
| **4** | 1 | 1.50 | 6.857000 | 1 | 964.0000 |
| **5** | 1 | 4.00 | 9.877500 | 19 | 887.6316 |
| **6** | 1 | 4.00 | 10.300000 | 82 | 886.4024 |
| **7** | 1 | 4.00 | 19.995000 | 12 | 829.0833 |
| **8** | 1 | 7.60 | 27.283500 | 56 | 878.6964 |
| **9** | 1 | 7.60 | 27.285000 | 24 | 754.2083 |
| **10** | 24 | 4.00 | 0.370222 | 131 | 893.9084 |
| **11** | 24 | 16.40 | 0.716444 | 91 | 835.7033 |
| **12** | 24 | 49.01 | 2.388440 | 46 | 497.5217 |
| **13** | 60 | 1.50 | 0.083000 | 315 | 950.9556 |
| **14** | 60 | 4.00 | 0.148000 | 6 | 980.3333 |
| **15** | 60 | 4.00 | 0.166500 | 88 | 936.0227 |
| **16** | 60 | 4.00 | 0.250000 | 42 | 910.8571 |
| **17** | 60 | 7.60 | 0.375000 | 41 | 894.0000 |
| **18** | 60 | 7.60 | 0.500000 | 45 | 842.3111 |
| **19** | 60 | 49.01 | 0.681111 | 55 | 749.1091 |



Still significant drop in age for 24 fractions. The only thing that could be causing this is the 16.4Gy dose… but it’s not compared to any other number of fractions so I don’t know why that would be causing the issue. Based on trends, it seems that the 1 fraction would have a lower age for that dose and the 24 fractions would improve it. Remove and see what happens.

This didn’t solve the issue… maybe not being compared to 1 fraction for 16.4Gy and 49.01Gy could be causing the issue. Not sure how to solve this though. Hold for now and come back later. Ask Dr. Malthouse for advice.

Check dose rates – as dose rate increases so does total dose a lot of the time… correlation is actually small, but still an issue.

> cor(df2$dose\_group2, df2$rate\_group)

[1] 0.1231508

> cor(df2$total\_dose, df2$dose\_rate)

[1] 0.1712505

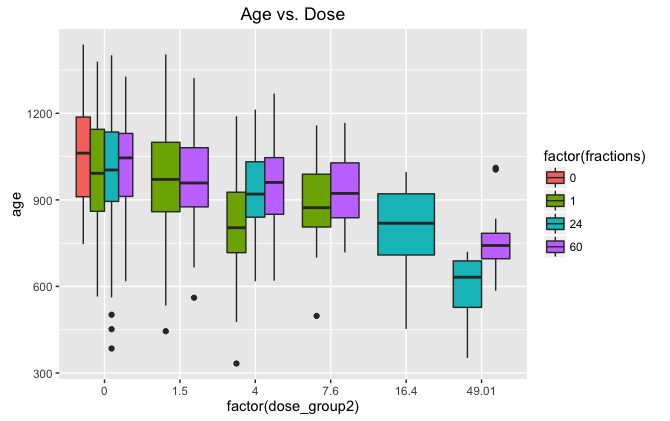
> cor(df2$dose\_rate, df2$fractions)

[1] -0.2917558

> cor(df2$total\_dose, df2$fractions)

[1] 0.1546108

Check dose\_group2 with lung cancer now (before it was lymphoma).



Call:

lm(formula = age ~ factor(dose\_group2) + factor(fractions), data = df2)

Residuals:

Min 1Q Median 3Q Max

-612.43 -105.94 5.05 120.22 446.01

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1047.06 42.85 24.438 < 2e-16 \*\*\*

factor(dose\_group2)1.5 -29.75 18.09 -1.645 0.10038

factor(dose\_group2)4 -107.36 20.09 -5.345 1.24e-07 \*\*\*

factor(dose\_group2)7.6 -102.30 31.37 -3.261 0.00116 \*\*

factor(dose\_group2)16.4 -196.67 33.27 -5.911 5.38e-09 \*\*\*

factor(dose\_group2)49.01 -298.48 35.69 -8.362 3.47e-16 \*\*\*

factor(fractions)1 -59.31 44.75 -1.326 0.18544

factor(fractions)24 -49.63 45.82 -1.083 0.27913

factor(fractions)60 -20.19 46.12 -0.438 0.66168

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 176.7 on 681 degrees of freedom

Multiple R-squared: 0.1527, Adjusted R-squared: 0.1428

F-statistic: 15.35 on 8 and 681 DF, p-value: < 2.2e-16

> Anova(uw.model)

Anova Table (Type II tests)

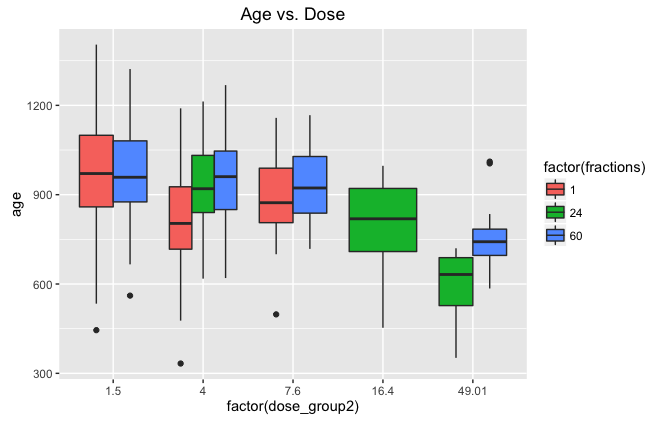
Response: age

Sum Sq Df F value Pr(>F)

factor(dose\_group2) 3542337 5 22.7022 < 2e-16 \*\*\*

factor(fractions) 197212 3 2.1065 0.09811 .

Residuals 21252005 681



Call:

lm(formula = age ~ factor(dose\_group2) + factor(fractions), data = subset(df2,

dose\_group2 != 0))

Residuals:

Min 1Q Median 3Q Max

-531.19 -99.67 1.38 115.03 449.58

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 954.42 14.36 66.453 < 2e-16 \*\*\*

factor(dose\_group2)4 -90.22 23.70 -3.807 0.000164 \*\*\*

factor(dose\_group2)7.6 -72.69 30.02 -2.422 0.015923 \*

factor(dose\_group2)16.4 -198.36 43.46 -4.564 6.84e-06 \*\*\*

factor(dose\_group2)49.01 -279.77 35.55 -7.870 3.95e-14 \*\*\*

factor(fractions)24 44.69 33.06 1.352 0.177172

factor(fractions)60 50.60 21.03 2.406 0.016600 \*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 165.2 on 371 degrees of freedom

Multiple R-squared: 0.1829, Adjusted R-squared: 0.1697

F-statistic: 13.84 on 6 and 371 DF, p-value: 3.242e-14

> Anova(uw.model)

Anova Table (Type II tests)

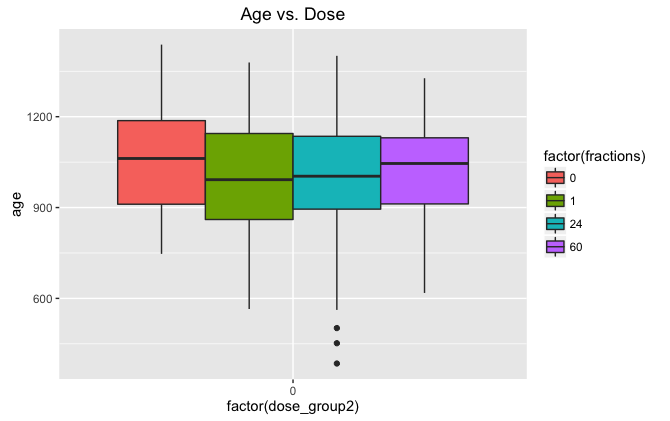
Response: age

Sum Sq Df F value Pr(>F)

factor(dose\_group2) 1933650 4 17.7152 2.548e-13 \*\*\*

factor(fractions) 166762 2 3.0556 0.04828 \*

Residuals 10123832 371



Call:

lm(formula = age ~ factor(fractions), data = subset(df2, dose\_group2 ==

0))

Residuals:

Min 1Q Median 3Q Max

-603.76 -132.14 9.21 135.18 412.24

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1047.06 46.00 22.761 <2e-16 \*\*\*

factor(fractions)1 -51.92 48.59 -1.068 0.286

factor(fractions)24 -58.30 49.99 -1.166 0.244

factor(fractions)60 -25.23 52.75 -0.478 0.633

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 189.7 on 308 degrees of freedom

Multiple R-squared: 0.007044, Adjusted R-squared: -0.002627

F-statistic: 0.7283 on 3 and 308 DF, p-value: 0.5357

> Anova(uw.model)

Anova Table (Type II tests)

Response: age

Sum Sq Df F value Pr(>F)

factor(fractions) 78604 3 0.7283 0.5357

Residuals 11080020 308