PE Summary Statistics

Aidan Mullan

2/7/2019

Summary Statistics

The first two variables we will look at are the Effective Dose of radiation given to patients and the amount of contrast given before the patient's scan. A focus will be placed on how these two measures differ between age groups (under 18, 18-35, older than 35) and BMI groups (under 25, 25-40, above 40).

Note: Any differences or patterns in the data described here are neither definitively significant nor non-significant unless specifically stated. Significance in variable effect will be determined by later regression.

Effective Radiation Dose and Contrast Given to Patients

Figure 1 shows the distribution of the dose of radiation given to patients during their scan. The average dose administered was 7.23 sieverts. There are a number of patients receiving a substantially higher dose than this average. 522 patients received more than 10 sieverts, 163 received more than 20, and 32 patients received an effective does of more than 40 sieverts. The largest dose received by any patient was 104.73 sieverts.

Figure 1a: Histogram of Effective Dose

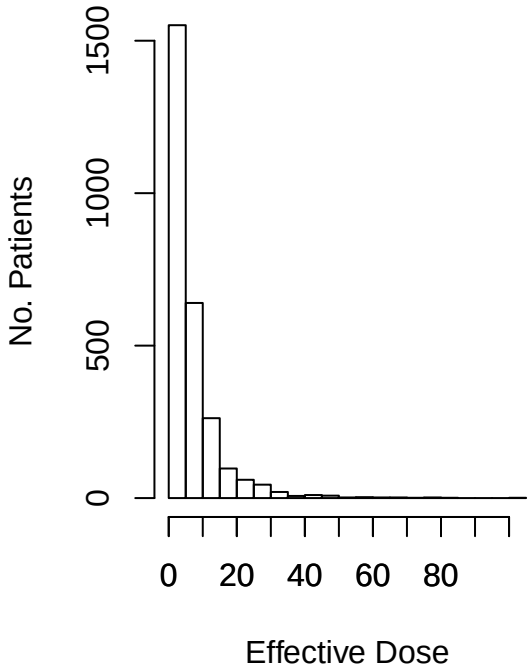
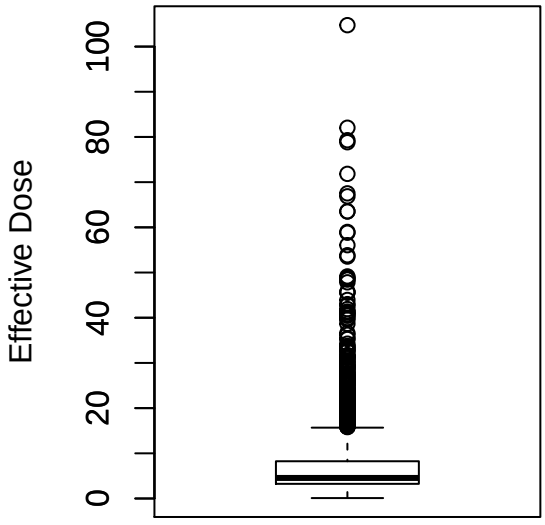


Figure 1b: Boxplot of Effective Dose




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.098   3.248   4.550   7.227   8.246 104.734
```

The distribution in the amount of contrast given to patients prior to their CT scan appears to have a pseudo-normal shape centered around 90, as seen in Figure 2. The average amount of contrast used for a

patient was 91.32 units, with a minimum of 5 units and a maximum of 186 units. There were 310 patients that received more than 100 units of contrast, while 20 patients received below 50 units.

Figure 2a: Histogram of Contrast Given

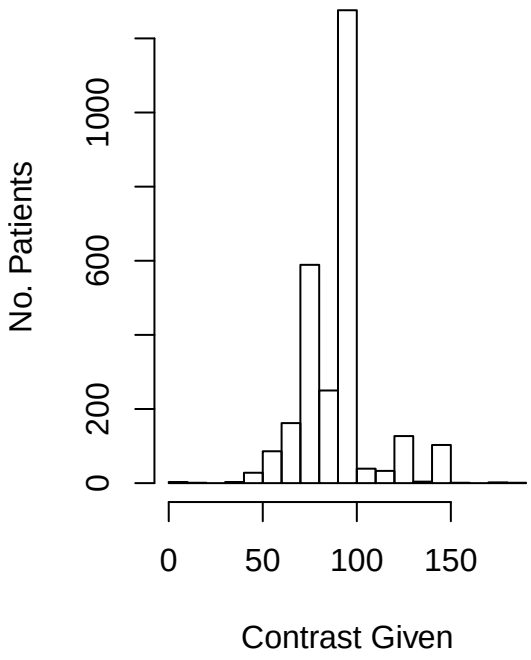
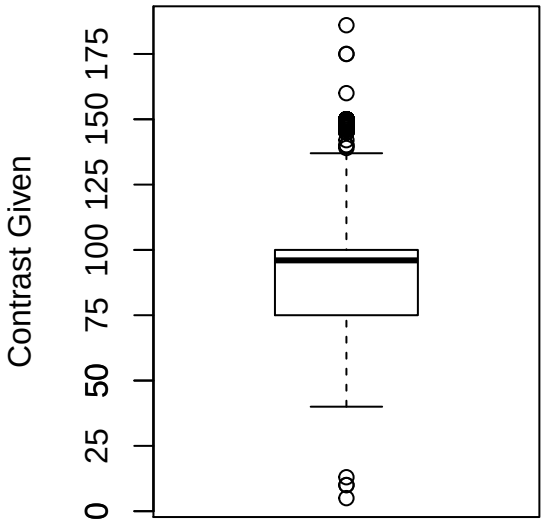


Figure 2b: Boxplot of Contrast Given




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    5.00   75.00   96.00   91.32  100.00  186.00       5
```

Comparison with Age and BMI

Both the dose of radiation and the amount of contrast given to a patient rely heavily on the patient's age and body composition. As such, we are interested in observing any differences in effective dose and contrast across these two variables. For our purposes, age has been divided into 3 groups: under 18, between 18 and 35, and older than 35. Similarly, BMI had been divided into 3 groups as well: under 25, between 25 and 40, and above 40. Table 1 shows the breakdown of patients that fall into each subgroup.

Table 1: No. of Patients by Age and BMI groups

		BMI			
		<25	25-40	>40	
AGE	<18	12 (0.4)	10 (0.3)	2 (0.1)	24 (0.9)
	18-35	95 (3.6)	145 (5.4)	61 (2.3)	301 (11.3)
	>35	650 (24.3)	1336 (50.0)	363 (13.6)	2349 (87.8)
		757 (28.3)	1491 (55.8)	426 (15.9)	2674 (100.0)

Note: Percentage of total patients is given in parentheses

Here we see that over 87% of all patients examined were over 35 years of age. Moreover, approximately half of all patients examined were both over 35 years of age and above a BMI between 25 and 40. This suggests that just looking at effective dose and contrast without accounting for age and BMI will skew our results

toward older, overweight but not obese patients. Tables 2 and 3 give the average effective dose and contrast administered, broken down into these same age and BMI categories.

Table 2: Average Effective Dose by Age and BMI groups

		BMI			
		<25	25-40	>40	
AGE	<18	1.540	4.416	10.633	3.412
	18-35	4.008	6.086	13.983	7.002
	>35	4.574	7.060	13.194	7.296
		4.455	6.948	13.295	7.227

Table 3: Average Contrast Given by Age and BMI groups

		BMI			
		<25	25-40	>40	
AGE	<18	59.08	69.60	95.50	66.28
	18-35	84.61	89.18	102.83	90.56
	>35	87.85	90.59	103.80	91.68
		86.98	102.94	90.31	91.32

From Table 3, it appears that the effective dose of radiation administered increases with age and BMI, and it seems as if BMI has a greater impact on dose than age (This effect will be concretely determined in the later modeling). In Table 4, we see a drastic increase in the amount of contrast used between the under 18 age group and the 18-35 group. There may be a slight increase between this middle group and patients over 35, but the difference appears to be minimal. There does not seem to be any major difference in contrast used across varying BMI groups.

Effect of Scanner Type

The various CT scans were performed using one of four different types of CT scanners, each with different output levels of radiation and varying prerequisites that must be met before using a specific scanner. Table 4 shows the average effective dose and contrast given to patients before a scan on each type of scanner. The "DS_Force" and "DS_Drive" scanners were consolodated into a single variable labelled "DS_Scanner".

Table 4: Average Radiation and Contrast by Scanner Type

	64SSWoIR	64SSwIR	128SSwIR	DS Scanner
Effective Dose	10.696 (8.63)	9.305 (10.87)	7.113 (6.87)	6.962 (8.43)
Contrast	62.41 (43.36)	67.01 (40.13)	68.31 (36.48)	62.72 (33.47)

Note: One standard deviation is given in parentheses

It appears, from Table 4, that there are not any apparant differences in administered radiation or contrast between the four types of scanners due to the relatively large standard deviations present. However, this is merely the direct effect of scanner type on contrast and radiation dosage. After accounting for other variables during regression, we may see this effect change.

Positivity Rate

The proportion of tests in which a positive result was achieved in detecting a pulmonary embolism will be referred to as the positivity rate. The overall positivity rate for this data was 10.93%, which corresponds to 296 positive test results. While this is descriptive of the patient base as a whole, the focus of the summary statistics in this section will be on the positivity rate for sub-groups of several key variables.

Positivity and Radiation/Contrast

Figure 3a gives the distribution of radiation administered to patients, group by whether or not their CT scan returned positive results. Figure 3b does the same for the distribution of contrast. In either case, a test result of 0 indicates a negative scan. There does not appear to be any substantial difference in the compounds administered to patients when we group by test result. One thing worth noting here is that when looking at the effective dose of radiation administered, we see more outliers in the case where test results ended up being negative. This would suggest that these patients are being unnecessarilly exposed to large amounts of radiation, the top end being over 100 sieverts.

Figure 3a: Distribution of Effective Dose

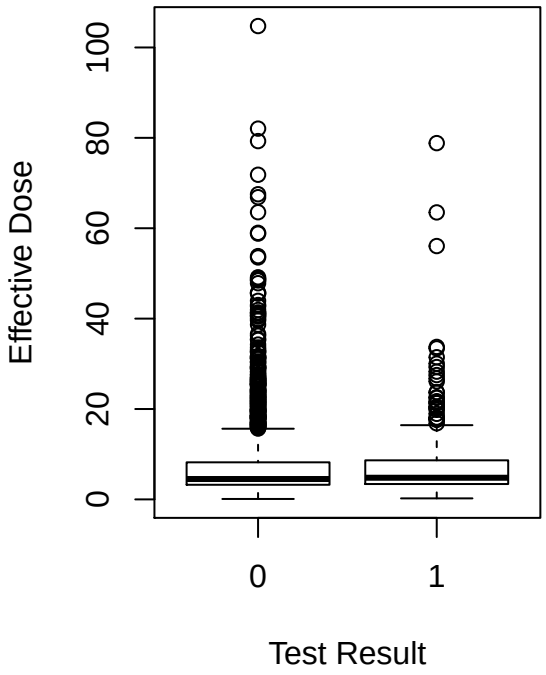
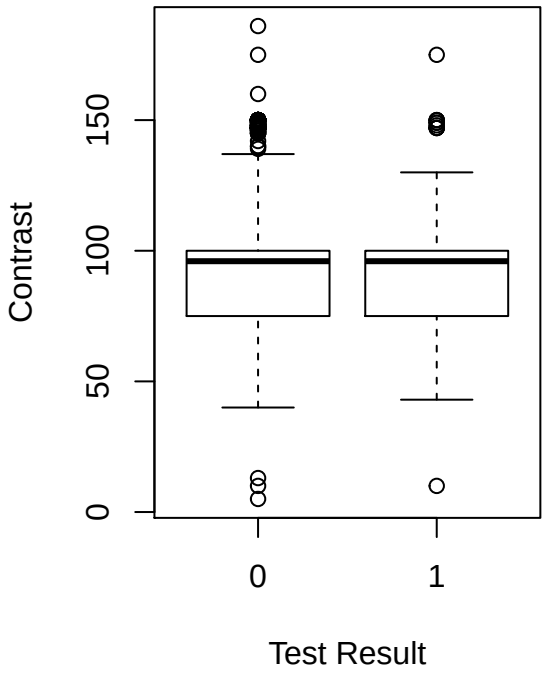


Figure 3b: Distribution of Contrast Administered



Positivity and Type of Procedure

The three types of procedures present in the dataset are chest only, double rule out, and chest, abdomen, and pelvis. There was one case of a triple rule out, which has been excluded from analysis. Table 5 provides both the counts and positivity rate for these types of procedures. Although there appears to be a difference in positivity between Chest Only scans (10.9%) and CAP scans (12.3%), this difference was found to be non-significant through a two-proportion Z test ($Z = 0.708$, $p = 0.239$). Interestingly, all 43 double rule out procedures produced negative results, which may suggest that double rule out procedures are being over-used.

Table 5: Positivity Rate by Type of Procedure

	n	Positivity
Chest Only	2369	0.109
Chest, Abdomen, and Pelvis	301	0.123
Double Rule Out	43	0.000

Positivity by Age and BMI

Once again, we are interested in the rate of positive test results after subdividing the patients into the three age and BMI groups used previously. These positivity rates are given in Table 6. From this table it is hard to conclude anything regarding a pattern in positivity rate for Age group or BMI group. There may be a slight decrease in positivity for older age groups, but this may be caused by the relatively small sample size of younger patients.

Table 6: Positivity Rate by Age and BMI groups

		BMI			
		<25	25-40	>40	
AGE	<18	0.083	0.300	0.000	0.160
	18-35	0.116	0.124	0.115	0.117
	>35	0.091	0.118	0.099	0.108
		0.094	0.119	0.101	0.109

Positivity and Location of Patient

We are also interested in any difference in positivity dependent on where the patient came from prior to their CT scan. Both the number of patients from each location and the positivity rate are given in Table 7. We can see here that both the ICU and IN locations have a higher rate of positive test results than either the ED or OUT patients. While this appears to be a substantial difference given the sample size for each group, we will determine significance through later regression.

Table 7: Positivity Rate by Patient Location

	n	Positivity
ED	1537	0.096
OUT	352	0.085
IN	613	0.124
ICU	211	0.153

Positive Test Result Statistics

Beyond simply looking at positivity rates, we are also interested in determining patterns within the group of patients that test positive for a pulmonary embolism. More specifically, we will be looking at differences in the type of pulmonary embolism as well as the type of strain detected by the CT scan. Overall, there were 296 patients who tested positive. One patient was recorded as testing positive but did not have information regarding the type of pulmonary embolism or strain, so they were removed from the following summaries, giving us 295 patients for analysis.

Type of PE

The number of patients testing positive for each type of pulmonary embolism, divided by location of admittance, is given in Table 8. Here we notice that half of all patients testing positive come from the emergency department category. However, as we saw previously, ED had the lowest positivity rate of all locations, suggesting that the reason for having som many positive cases is that most patients receiving CT scans came from ED. We can also see that a majority of pulmonary embolisms detected are classified as peripheral.

Table 8: Type of PE by Location Admitted

		Location				
		ED	OUT	IN	ICU	
PE Type	Central	13 (4.4)	0 (0.0)	2 (0.7)	2 (0.7)	17 (5.8)
	Central and Peripheral	20 (6.8)	1 (0.3)	11 (3.7)	10 (3.4)	42 (14.2)
	Peripheral	115 (39.0)	17 (5.8)	62 (21.0)	42 (14.2)	236 (80.0)
		148 (50.2)	18 (6.1)	75 (25.8)	54 (18.3)	295 (100.0)

Note: Percentage of total patients is given in parentheses

Table 9 shows the breakdown of type of PE by patient age. We find that a vast majority of positive test results come from patients over 35 years of age. Again however, this is not necessarily due to anything except the fact that most CT scans were performed on patients over 35. This is evidenced by the fact that this age range did not have a larger positivity rate than either of the other two.

Table 9: Type of PE by Age Group

		AGE			
		<18	18-35	>35	
PE Type	Central	0 (0.0)	6 (2.0)	11 (3.7)	17 (5.8)
	Central and Peripheral	1 (0.3)	7 (2.4)	34 (11.5)	42 (14.2)
	Peripheral	3 (1.0)	23 (7.8)	210 (71.1)	236 (80.0)
		4 (1.4)	36 (12.2)	256 (86.8)	295 (100)

Note: Percentage of total sample is given in parentheses

Type of Strain

Similar to the type of pulmonary embolism detected, we are also interested in the distribution of whether a strain was detected on the CT scan, grouped by location of patient admittance and age group. These data are given in Table 10 and 11, respectively. Here we see that a larger proportion of positive CT scans showed no strain. Again, we have higher incidence of strain in the ED and age over 35 categories due to the larger number of patients overall who fall into these groups.

Table 10: Detection of Strain by Location Admitted

		Location				
		ED	OUT	IN	ICU	
Strain	No	88 (29.8)	14 (4.7)	51 (17.3)	33 (11.2)	186 (63.1)
	Yes	60 (20.3)	4 (1.4)	24 (8.1)	21 (7.1)	109 (36.9)
		148 (50.2)	18 (6.1)	75 (25.4)	54 (18.3)	295 (100)

Note: Percentage of total sample is given in parentheses

Table 11: Type of PE by Age Group

		AGE			
		<18	18-35	>35	
Strain	No	2 (0.7)	18 (6.1)	166 (56.3)	186 (63.1)
	Yes	2 (0.7)	18 (6.1)	89 (30.2)	109 (36.9)
		4 (1.4)	36 (12.2)	256 (86.8)	295 (100)

Note: Percentage of total sample is given in parentheses

Next Steps

All of these tables demonstrate a summary of how some of our variables of interest interact. It is important to note that no conclusions as to the significance of variable effects or differences can be draw here because no actual statistical tests have been carried out (apart from the two-sample proportions test for positivity in chest only and CAP scans).

The next steps here will be to fit a logistic regression model for positivity using the key variables of interest listed above. Additionally, models for mean radiation and mean constrast will be fit to see which variables have significant effects.