1.

The variable 'treatment' is a dummy variable with a value of 1 if the individual won the lottery and thus had the option to apply to Medicaid, and 0 if they did not win the lottery. The variable 'ohp_all_ever_survey' is a dummy variable that takes the value of 1 if the individual enrolled in Medicaid and 0 if they did not. The important variation here is that even if the individual wins the lottery (treatment value of 1), they could potentially choose to not apply to Medicaid or have their application rejected (ohp_all_ever_survey value of 0). Likewise just because the individual did not win the lottery (treatment value of 0), they could still potentially have gotten into Medicaid through another process (ohp_all_ever_survey value of 1). The thing that was randomized in this experiment was who won the lottery, not who gets Medicaid or not. Since someone could win the lottery and still not get on Medicaid, or not win the lottery and still get on Medicaid through other means, the variable showing if they were on Medicaid or not cannot be used as the treatment variable since that was not randomized. Randomization is needed in an experiment to get rid of the selection bias. Those who get on Medicaid could be a specific group of people which would lower the external validity of the results if being on Medicaid was used as the treatment since it was not randomized. Since the only thing that was randomized was who won the lottery, the treatment variable needs to be if someone won the lottery or not.

2 and 3.

We choose 6 variables that are set prior to the experiment and so should be randomized across groups and run a regression on them with the treatment variable to show a balance test between groups.

Balance Test between Treatment and Control Groups

	Dependent variable:					
	Age	Education	Gender	Depression prior to Lottery	Diabetes prior to Lottery	Hypertension prior to Lottery
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Effect	0.380*	0.022	-0.006	-0.018**	-0.001	-0.001
	(0.212)	(0.016)	(0.009)	(0.009)	(0.005)	(0.007)
Control Mean	40.606***	2.238***	0.569***	0.350***	0.072***	0.183***
	(0.153)	(0.012)	(0.006)	(0.006)	(0.003)	(0.005)
Observations	12,228	12,218	12,229	12,229	12,229	12,229
R^2	0.0003	0.0001	0.00004	0.0004	0.00000	0.00000
Adjusted R^2	0.0002	0.0001	-0.00004	0.0003	-0.0001	-0.0001
Residual Std. Error	11.697 (df = 12226)	0.908 (df = 12216)	0.496 (df = 12227)	0.474 (df = 12227)	0.257 (df = 12227)	0.386 (df = 12227)
F Statistic	3.225* (df = 1; 12226)	1.737 (df = 1; 12216)	0.463 (df = 1; 12227)	4.549** (df = 1; 12227)	0.029 (df = 1; 12227)	0.037 (df = 1; 12227)
Note:						*p<0.1;* ***p<0.05;*** p<0.01

4.

The table is made up of variables that should not be affected by being assigned to the treatment or control group (variables that are predetermined prior to the experiment and whose values will not change after the experiment). By looking at these variables, such as health standards before the experiment began, we can show that the groups were randomly assigned as long as there is not much difference in assignment rate for any of the variables between groups. The table shows the means of the control group in row 2 with the standard deviations underneath. Row 1 shows the difference in the treatment group means from the control groups means, with the standard deviations beneath in parenthesis. The ideal situation to prove that the individuals have been perfectly randomly assigned would be to have the difference between the treatment and control groups be zero. Looking at the value of the differences, we can see that all the values are very close to zero. There are two columns that show a star for statistical significance. Looking at these two columns (age, and depression prior to the lottery), we see that the 95% confidence intervals for these variables are (-0.044, 0.804) and (-0.036, 0) respectively. In both of these confidence intervals 0 is included showing that the value we got in the table could happen even with the true value being 0, 95% of the time. Other than the confidence interval including zero, we can see that the treatment effect that is significant for age is .38 years which is not a tangible difference in the real world and the coefficient for depression of .018 is also a small amount of less than one person being affected. This shows that even if they show significance, these columns are still reasonably close to zero. Looking at all the columns, we can see that compared to the mean value of the control group, the confidence intervals of the differences between the treatment and control groups are reasonably small. Since the standard deviations and means for the differences between treatment and control groups are small when compared to the mean value of the control group, we can say that the balance table is consistent with individuals being randomly assigned to the two groups.

5.

To determine the compliance rate we run a regression of the treatment variable of winning the lottery on the variable showing if they ever signed up for Medicaid. This gives us the amount of people who won the lottery and signed up for Medicaid. The table (below) shows that the compliance rate was 25.4%.

Estimation of the Compliance Rate

	Dependent variable:
	Ever on Medicaid
Treatment Effect	0.254***
	(0.008)
Control Mean	0.158***
	(0.005)
Observations	12,229
R^2	0.078
Adjusted R^2	0.078
Residual Std. Error	0.436 (df = 12227)
F Statistic	1,031.581*** (df = 1; 12227)
Note:	*p<0.1; **p<0.05;** p<0.01

6.

Intent to Treat Effect

	Dependent variable:					
	Post Lottery Blood Pressure (1)	Post Lottery Depression (2)	Post Lottery Diabetes (3)	Post Lottery Hypertension (4)	Post Lottery Number of Medications (5)	Post Lottery Number of Doctor Visits (6)
Treatment Effect	-0.058	0.005	0.009***	0.002	0.128**	0.396*
	(0.300)	(0.004)	(0.002)	(0.004)	(0.053)	(0.216)
Control Group	119.130***	0.049***	0.012***	0.057***	1.838***	5.746***
	(0.219)	(0.003)	(0.001)	(0.003)	(0.037)	(0.155)
Observations	12,188	12,095	12,186	11,945	11,912	12,158
R^2	0.00000	0.0001	0.001	0.00003	0.0005	0.0003
Adjusted R^2	-0.0001	0.00003	0.001	-0.0001	0.0004	0.0002
Residual Std. Error	16.550 (df = 12186)	0.221 (df = 12093)	0.127 (df = 12184)	0.234 (df = 11943)	2.891 (df = 11910)	11.895 (df = 12156)
F Statistic	0.038 (df = 1; 12186)	1.306 (df = 1; 12093)	13.906*** (df = 1; 12184)	0.315 (df = 1; 11943)	5.855** (df = 1; 11910)	3.357* (df = 1; 12156)
Note:						*p<0.1; **p<0.05;** p<0.01

The ITT effect can be seen on the Treatment Effect row. We can see from this table that the effect of winning the lottery (the ITT effect) is quite close to zero for many of the health outcomes we used as variables. The effects seen on this table for the most part seem to reflect near zero ITT effects, as the coefficient for the treatment effect and the standard deviation would make a confidence interval that includes zero for nearly every variable. The variables that seem like they might have some effect of being treated are the number of medications taken and the number of doctor visits, since the treatment effect coefficient combined with the standard deviation is somewhat closer magnitude wise to the control group coefficient. What this could show is that the treatment did not have a significantly large effect on physical health characteristics but it did have an effect on the amount of health care that those in the treatment group made use of.

7.

Variable (Post Experiment)	Blood Pressure	Depression	Diabetes	Hypertension	Number of Medication Taken	Number of Doctor Visits
ATET	-0.228	0.020	.035	.008	.504	1.560

The treatment on the treated effect (ATET) point estimate can be found with the formula of the intent to treat effect (ITT)/the compliance rate. We saw in the previous questions that the compliance rate was quite a bit lower than 100%. The ITT and ATET would be the same number if there were 100% compliance rate but in this case the ATET must be calculated separately. The ITT shows the effect of actually being selected for the treatment, in this case winning the lottery. The ATET in this case shows the effect of actually applying for Medicaid after winning the lottery on the various variables, which is why the compliance rate is an issue since many of those who won the lottery did not actually end up on Medicaid. By dividing the ITT by the compliance rate we can see how being selected for treatment affects the variable of interest but only for those that actually complied which is what we want for interpretation of effects. The ATET, since it includes the ITT and compliance rate, can be looked at as a much more accurate way to measure the causal effect of Medicaid than if just the ITT was looked at. In this case the ATET effect is quite small for many of the variables with the most interesting again being number of medication and number of doctor visits, as these variables have a slightly larger magnitude than the other health variables looked at. As mentioned previously this could be a result of the Medicaid having close to zero or zero effect on health circumstances but having access to healthcare causes individuals with Medicaid to seek out doctors and medications more frequently.

8.

Attrition bias occurs since experiments can only get results from those individuals that complete the study. This means that those who leave the experiment or die part way through will not be included in the final results. Bias can occur because those that end up leaving the study in either the control or treatment group could be biased towards a certain group of people and so the validity or statistical power of the study suffers unless attrition is done randomly. Our balance test shows that for the most part the treatment and control groups are well randomized but it does not show anything about whether the randomization remains at the end of the experiment. Attrition bias seems like it is something to potentially be worried about in this situation, due in part to the low compliance rate. Since only about 25% of those in the treatment group actually complied there could be an issue where certain people are more likely to comply. This could be seen for example if those in the treatment group see they won the lottery and the people with decent health apply for Medicaid because it's simple to do after winning the lottery or they feel some pressure to actually apply after winning. If those of similar health standards in the control group do not feel that same pressure since they are fairly healthy, then they would not apply to Medicaid at the same rate and perhaps change the results. In another way attrition bias could occur if those in one of the groups leave the experiment at a higher rate than those in the other group; perhaps those in the control group who did not win the lottery move away at a higher rate since they do not have the opportunity to apply to Medicaid while those who won, stay in Oregon so they can use the Medicaid they applied for. These are examples that show potential attrition bias and why it is important to check for it, but it does not seem likely that this study had unbalanced leaving of groups. Since the people who would sign up for a lottery to get Medicaid would almost solely be poor or disenfranchised individuals, it seems unlikely that they would have the ability to get up and leave the experiment regardless of which group they were assigned. To determine if the potential attrition bias is significant, the researchers can assume a worst case scenario for the outcomes that left the experiment and if the results of the experiment stay essentially the same, then the attrition bias was likely not significant. In the case of this experiment, we could say that the worst case scenario is if the treatment effect showed 0 difference between the control and treatment groups. This could be the worst case scenario because you would expect that having the treatment to be allowed to apply to Medicaid would make the person who is treated be healthier, and if the difference between groups is 0 that shows that the treatment did not affect the individual's health. Since the results of this experiment already show an ATET effect of near 0 for most of the health characteristics post treatment, we would say that the attrition bias is likely not a big problem since even putting the worst case scenario of 0 effect for any missing data would not change the results which say that the effect of treatment is near 0. In all from what we can gather from the provided data and information it seems like attrition bias is not something to be too worried about in this study as it seems unlikely that a certain type of people would disproportionately leave the study based on the group they were assigned to and the results of the study do not change much even assuming the worst.

9.

The Oregon Medicaid lottery allowed for a research designed randomized experiment to be done where individuals were randomly chosen to be given the opportunity to apply to newly opened Medicaid slots. Various health characteristics were studied by those who won and lost the lottery as a way to judge the effectiveness of Medicaid on healthcare received. By looking at the treatment on the treated effect (ATET), the causal effect of Medicaid on various health statistics was visible. We found that the effect of having Medicaid coverage resulted in a .5% increase in the number of medications taken and a 1.5% increase in the amount of doctor visits that the individual attended. The effects of Medicaid on actual health characteristics were a bit more limited, closer to zero, showing that Medicaid perhaps causes individuals to pay more attention to their health but has less of an impact on their actual health.