R Notebook

Re-analysis of NIEHS Hair Straightener and Uterine Cancer Data

NIEHS issued a press release on Monday October 17, 2022, that stated that their researchers identified an increase risk of uterine cancers in women who used hair relaxers and straighteners. They used a Cox proportional hazards model to make this finding. Alexandra White of NIEHS stated, "we estimated that 1.64% of women who never used hair straighteners would go on to develop uterine cancer by the age of 70; but for frequent users, that risk goes up to 4.05%." This is based on an NIEHS study published in the Journal of the National Cancer Institute by Chang et al (2022).

To my eyes, that's a 2.4x increase – which seems a bit high to me.

What do we know and what is hypothesized by NIEHS:

1. We know that uterine cancers are relatively rare.
2. We know that rates of detected uterine cancers increased at a faster rate in non-hispanic Black women compared to White women from 2011-2015 according to NIH.
3. NIEHS hypothesized that increased uterine cancer rates in Black women could be do to relaxer use.
4. We also know that Black women have been using relaxers and hair products well before 2011. I should know – my parents owned a hair store where the overwhelming majority of our clientelle were Black women. We sold relaxers, gels, mousse, wigs, weaves, wax, shampoos, conditioners, dyes, chalks, you name it and we probably sold it. And we started that store in the 1990s. And stores have been selling relaxers since well before the 1990s. So it doesn't make much sense to blame hair products – the timing just doesn't fit.

What do we know about the NIEHS study?

1. It only had 29 Black women who developed uterine cancer; out of a total of 2,523 that were enrolled. That works out to 1.1%
2. The overwhelming majority of women who developed uterine cancer in the study were White women – 326 of them in fact, out of a total of 29,060. That works out to 1.1%
3. The women in the study had an average age of 58 years old, and most of the women who developed uterine cancer were followed up within 6.5 years.
4. The enrollment period for the study was 2003-2009.

Note that the topline number for uterine cancer rate in this study for both Black and White women was 1.1%. So at the time of the study, we're not seeing any significant difference between groups at the top-line. So now the question becomes – do we see one when we consider hair straightening product use?

Again, from a sociological perspective, the timing just doesn't fit. However, it should be noted that women can and do change up their hair styles, and one may not use relaxers often throughout the year. However, NIEHS noted that 60% of the women in their study that had used a relaxer or straightening product at least once in the prior year were Black.

The other thing to consider is that the number of Black women in the study is really, really small, and not representative of the overall US population. In fact, the number of White women in the study is much

greater than what we have in the population: at 85.6% of the women in the study, or 29,060. Because uterine cancer is such a rare cancer, and there are only 29 Black women who had it in the study, that means that any effect size we see when doing the analysis to identify an effect due to relaxers is likely to be exaggerated and a false positive – a form of Type M (or magnitude) error.

Analysis

My goal is to disprove NIEHS's conclusion that hair straightening product use is associated with an increased risk of uterine cancer (since they say an association exists, I have to test the opposite case under the "falsification" principle of science as articulated by Popper).

I don't have access to the individual data, and cannot get access to it. So I have to use what I have available, which is the occurrence data reported in the paper.

So I know the following:

Question 1:

332 out of 30,329 people who never used straighteners, relaxers or pressing products developed uterine cancer.

38 out of 3,036 people who ever used straighteners, relaxers, or pressing products developed uterine cancer.

So Question 1 is – with uncertainty, are these equivalent rates?

Question 2:

332 out of 30,329 people who never used straighteners, relaxers or pressing products developed uterine cancer.

12 out of 1,464 people who used straighteners, relaxers or pressing products <=4 times in the previous year developed uterine cancer.

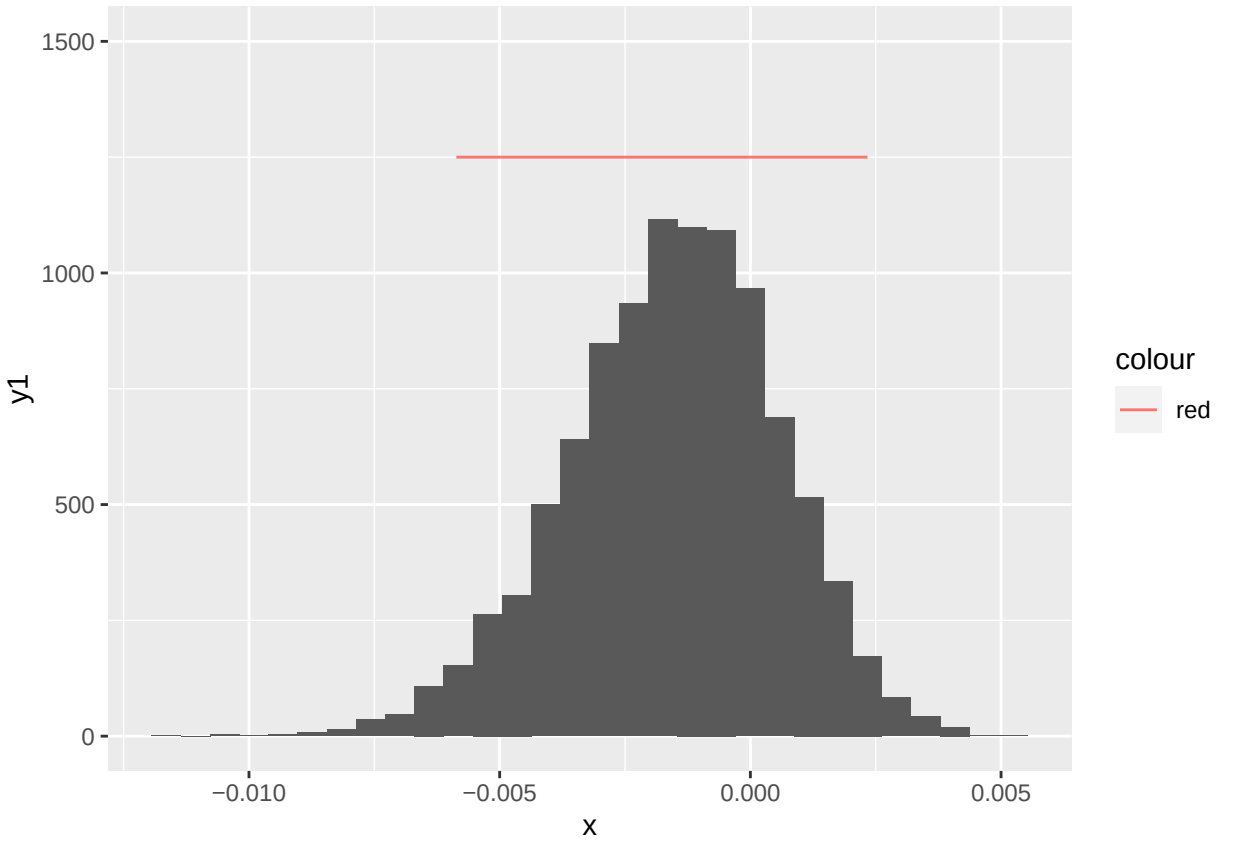
26 out of 1,572 people who used straighteners, relaxers or pressing products >4 times in the previous year developed uterine cancer.

So Question 2 is – with uncertainty, are these equivalent rates?

Question 1

```
library(ggplot2)
library(HDInterval)
never_used <- rbeta(10000, 332, 30329-332)
ever_used <- rbeta(10000, 38, 3036-38)
diff_never_ever <- never_used - ever_used
diff_never_ever <- data.frame(x <- diff_never_ever)
dne_hdi <- hdi(diff_never_ever$x)
dne_hdi_df <- data.frame(x1 = dne_hdi[1], x2 = dne_hdi[2], y1 = 1250, y2=1250)
ggplot(diff_never_ever, aes(x=x)) +
  ylim(0, 1500) +
  geom_histogram() +
  geom_segment(data=dne_hdi_df, aes(x=x1, y=y1, xend=x2, yend=y2, color = "red"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
median(diff_never_ever$x)
```

```
## [1] -0.001442742
```

Looking at the graph above, we can see that the probability of someone who has never used straighteners, relaxers or pressing products developed uterine cancer at approximately the same rate as someone who has. This is based on the fact that the 95% Highest Density Interval (the red line) includes a difference of 0.0, and has a median difference of -0.00147 – which is practically 0.

So yes, the uterine cancer rates are essentially equivalent.

Question 2

In this case we're trying to see if the uterine cancer rates are different if someone used a straighteners, relaxers or pressing products less than 4 times or more than 4 times.

This next block will process the data.

```
used_4_less <- rbeta(10000, 12, 1464-12)
used_4_more <- rbeta(10000, 26, 1572-26)
diff_used_less <- data.frame(x = never_used - used_4_less)
diff_used_more <- data.frame(x = never_used - used_4_more)

diff_used_less_hdi <- hdi(diff_used_less$x)
diff_used_more_hdi <- hdi(diff_used_more$x)
```

```
diff_used_less_hdi_df <- data.frame(x1 = diff_used_less_hdi[1],
                                    x2 = diff_used_less_hdi[2],
                                    y1 = 1250, y2=1250)

diff_used_more_hdi_df <- data.frame(x1 = diff_used_more_hdi[1],
                                    x2 = diff_used_more_hdi[2],
                                    y1 = 1250, y2=1250)

median(diff_used_less$x)
```

```
## [1] 0.002967449
```

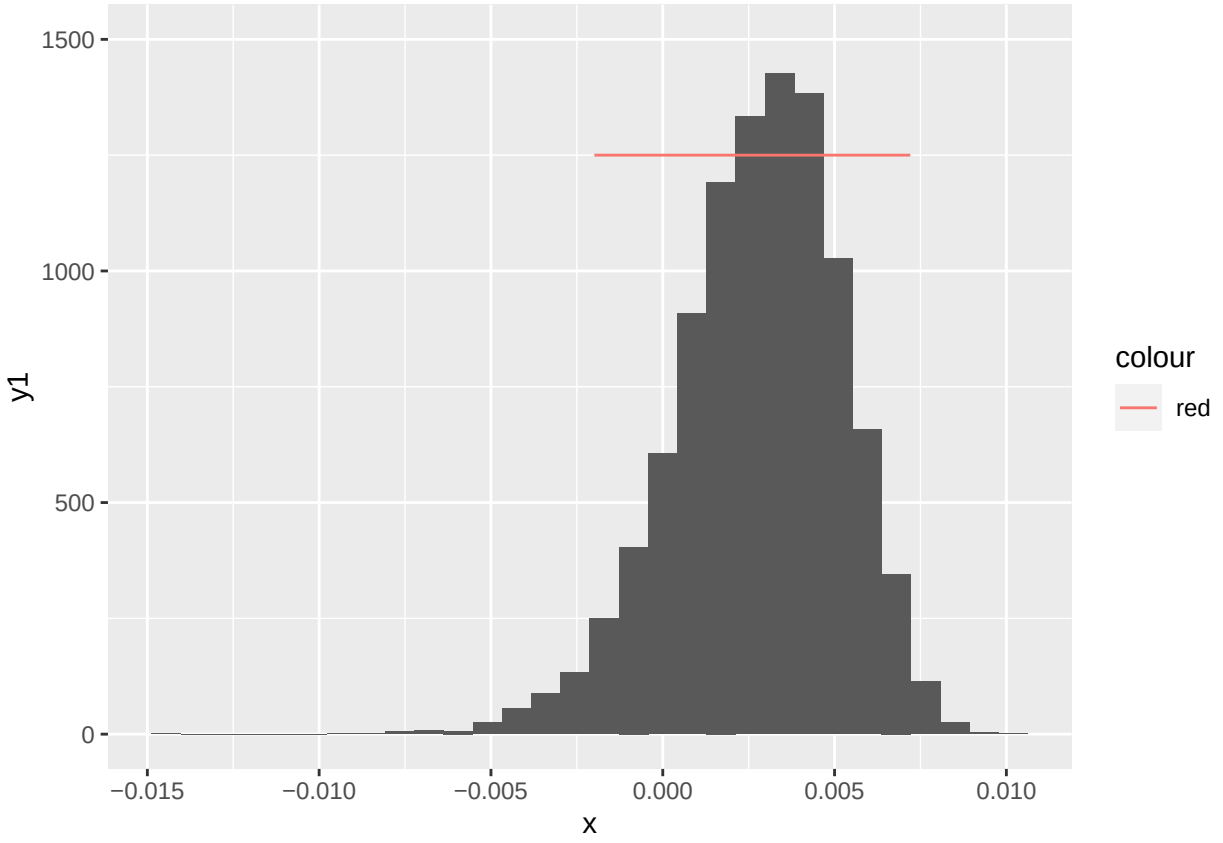
```
median(diff_used_more$x)
```

```
## [1] -0.005439432
```

This next block will graph the difference distribution for used <= 4 times:

```
ggplot(diff_used_less, aes(x=x)) +
  ylim(0, 1500) +
  geom_histogram() +
  geom_segment(data=diff_used_less_hdi_df, aes(x=x1, y=y1, xend=x2, yend=y2, color = "red"))
```

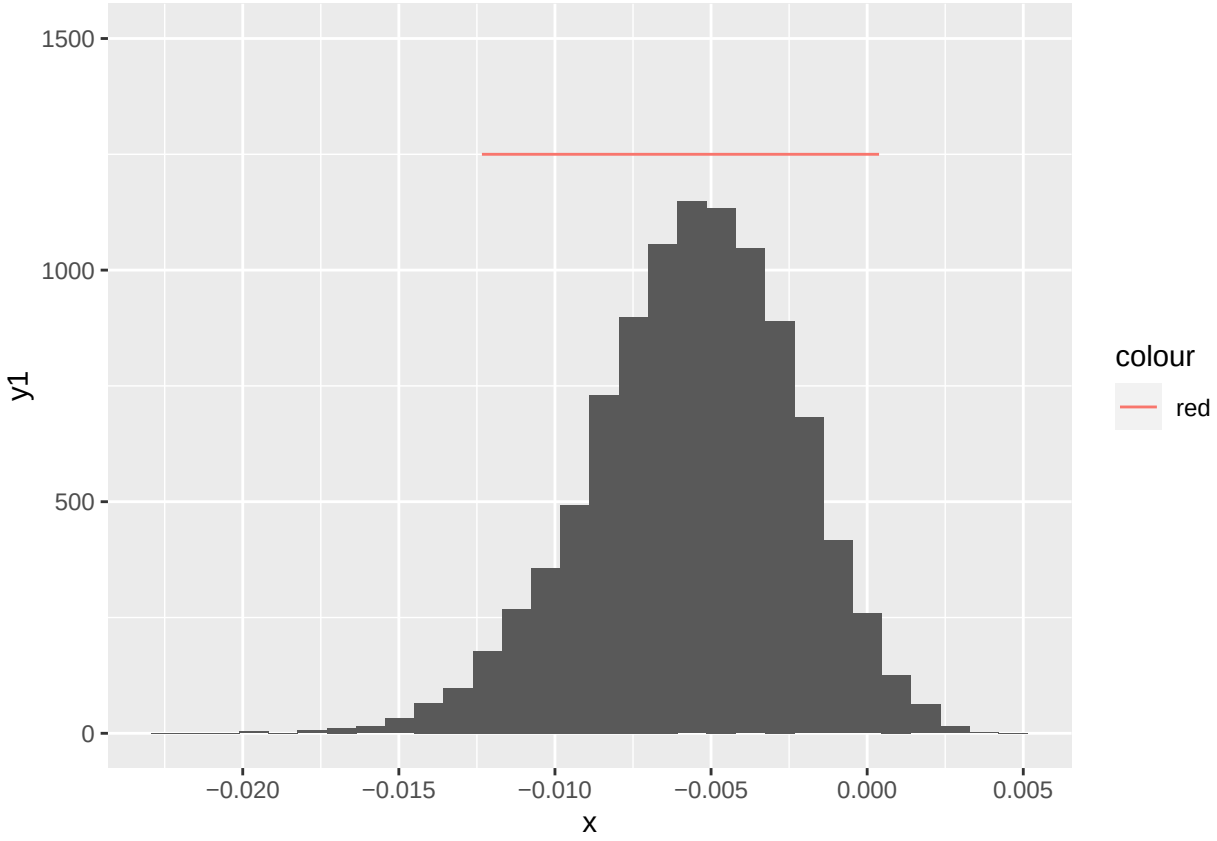
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



This next block will graph the difference distribution for used > 4 times:

```
ggplot(diff_used_more, aes(x=x)) +
  ylim(0, 1500) +
  geom_histogram() +
  geom_segment(data=diff_used_more_hdi_df, aes(x=x1, y=y1, xend=x2, yend=y2, color = "red"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Again, we can see that there's no difference; 0 is a plausible value, and the size of the differences in the distribution are all extremely small (most are less than 1%).

So what does this mean? Well, it means that based on these analyses there really isn't much going on.

Now, what if I were to treat these as hazard ratios? Well, that's going to be Question 3.

Question 3

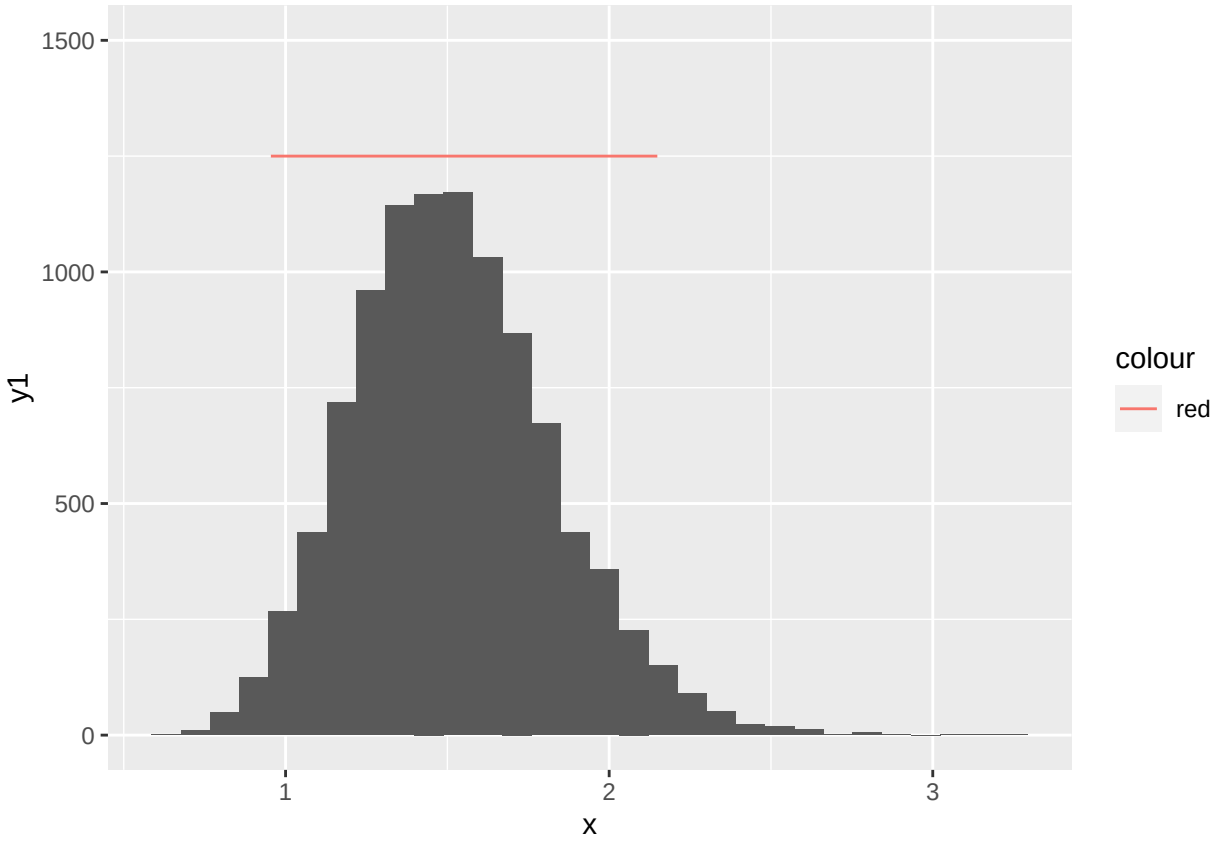
Here, I'm going to use just the cases where someone used straighteners, relaxers or pressing products more than 4 times, and the control. The control group will be our denominator in the ratio, and the more than 4 times group will be the numerator.

```
diff_used_more_ratio <- data.frame(x = used_4_more / never_used)
diff_used_more_ratio_hdi <- hdi(diff_used_more_ratio$x)
diff_used_more_ratio_hdi_df <- data.frame(x1 = diff_used_more_ratio_hdi[1],
```

```
                                    x2 = diff_used_more_ratio_hdi[2],
                                    y1 = 1250, y2=1250)

ggplot(diff_used_more_ratio, aes(x=x)) +
  ylim(0, 1500) +
  geom_histogram() +
  geom_segment(data=diff_used_more_ratio_hdi_df, aes(x=x1, y=y1, xend=x2, yend=y2, color = "red"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
mean(diff_used_more_ratio$x)
```

```
## [1] 1.518263
```

Question 4

We can see that in the never-ever use comparison that there's clearly no difference. But NIEHS found a difference when looking at >4 times use, but not at <= 4 times use. I also don't see a difference if I look at the >4 times use.

But something is odd here: The number of people who had uterine cancer in the <= 4x use of straighteners, relaxers or pressing products group is a lot smaller than we'd expect when we compare it to the never use group. Also, it appears that >4 times use is 2.17x larger than the proportion of people who use relaxers <=4 times (that is, 26 / 12 = 2.17x).

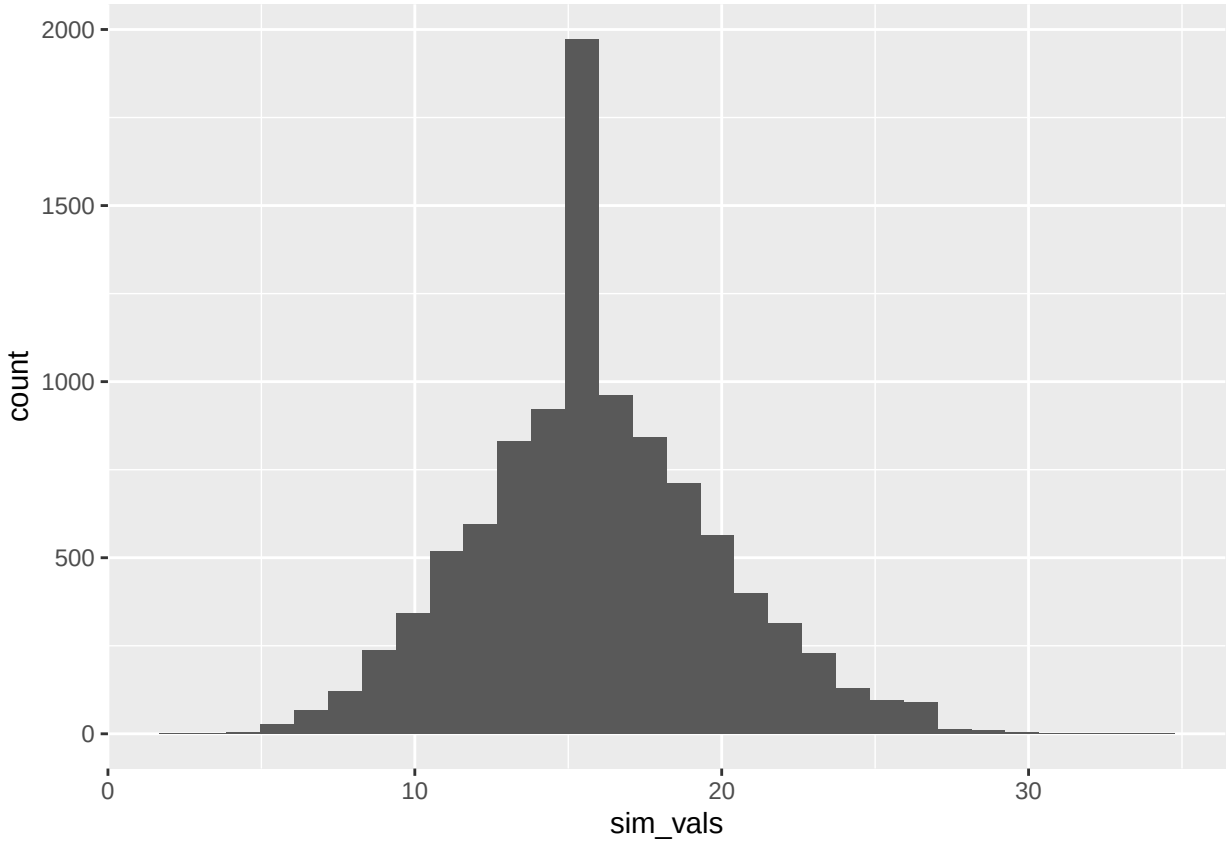
So, the rate in the <= 4 times group is 12/1464 = 0.8%. Compare this to the never use group where the uterine cancer rate is 1.09%. Isn't that interesting? Wouldn't we expect the <= 4 times group to be closer to the never use group?

So, let's see if we can get these types of numbers for <= 4 times and > 4 times groups using the median of the uterine cancer rate in the never use group.

```
x <- median(never_used)
sim_less_4 <- data.frame(sim_vals = rbinom(10000, 1464, x))
sim_more_4 <- data.frame(sim_vals = rbinom(10000, 1572, x))

ggplot(sim_less_4, aes(x=sim_vals)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
median(sim_less_4$sim_vals)
```

```
## [1] 16
```

```
median(sim_less_4$sim_vals) / 1464
```

```
## [1] 0.01092896
```

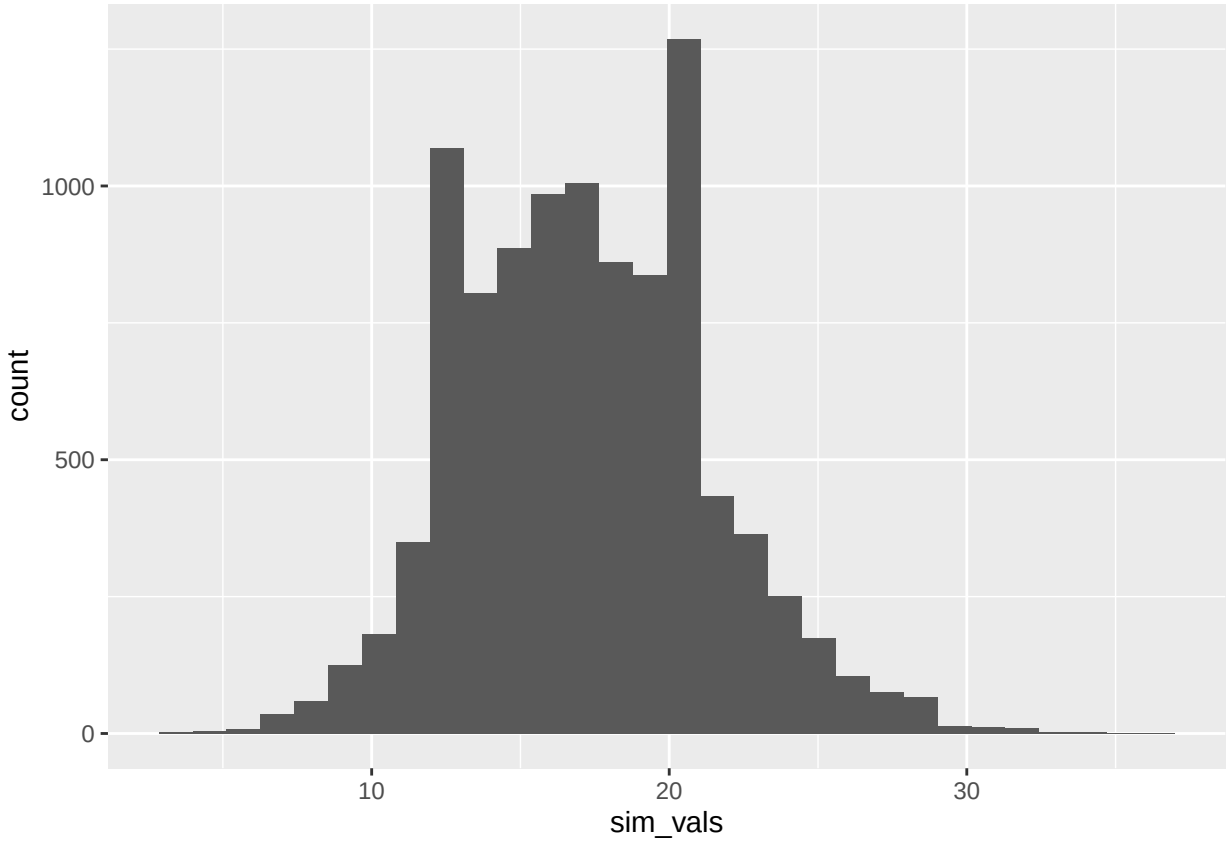
```
hdi(sim_less_4$sim_vals, 0.99)
```

```
## lower upper
##     7    26
## attr(,"credMass")
## [1] 0.99
```

So, we can see here that 12 cases, using the uterine cancer rate from the never used group, is entirely plausible for the less than 4 times use group. In fact, we can see that the median is closer to 16. And more importantly, we can see that the rate at the median is 1.1%, which is the same rate as for the never used group.

```
ggplot(sim_more_4, aes(x=sim_vals)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
median(sim_more_4$sim_vals)
```

```
## [1] 17
```

```
median(sim_more_4$sim_vals) / 1572
```

```
## [1] 0.01081425
```

```
pbinom(26, 1572, x)
```

```
## [1] 0.9834993
```

```
hdi(sim_more_4$sim_vals, 0.99)
```

```
## lower upper
##     7    28
## attr(,"credMass")
## [1] 0.99
```

So, we can see here that 26 cases, using the uterine cancer rate from the never used group, is entirely plausible for the less than 4 times use group (it is within the 99% Highest Density Interval). In fact, we can see that the median is closer to 17. And more importantly, we can see that the rate at the median is 1.1%, which is the same rate as for the never used group.

All this to say, these values are entirely plausible if we assume the cancer rate in the never used group. Thus, in other words, NIEHS' results are actually entirely plausible under the null hypothesis that there is no difference due to the amount of times that someone uses straighteners, relaxers or pressing products.

In other words, NIEHS' results are likely false positives.