Sanity checks

Sanity check data graphically

```
library(ggplot2)
library(ggrepel)
source("read-data.R")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:testthat':
##
##     matches
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
## In anonymous test function
## This study only has some eye number information missing, fix it:
## Vold et.al (2016)*
```

```
df <- read.data()
```

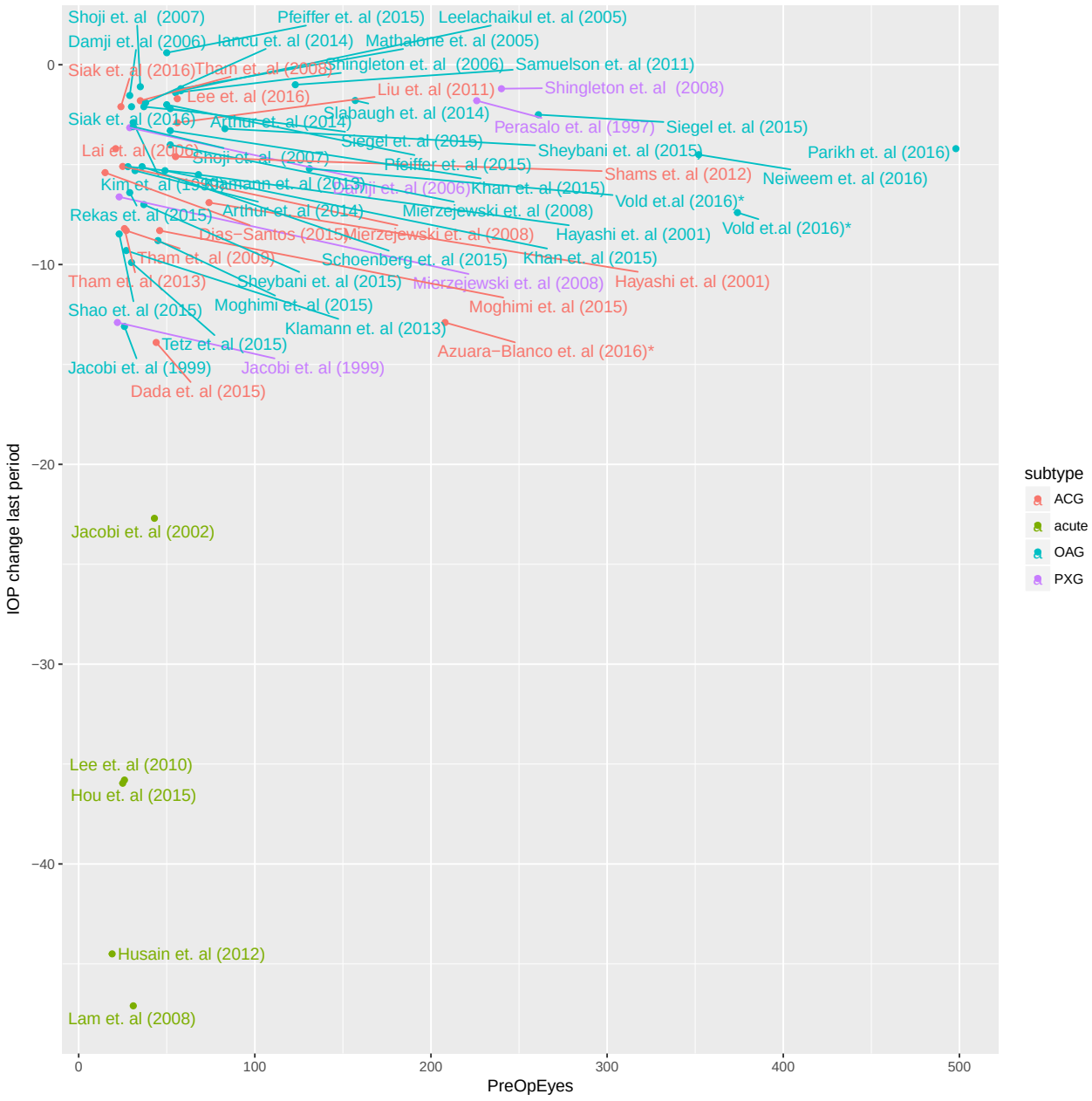
```
## This study only has some eye number information missing, fix it:
## Vold et.al (2016)*
```

Summary plot.

```
ggplot(df, aes(x = PreOpEyes, y =
                 ifelse(!is.na(LastPeriodIOPMean) & !is.na(PreOpIOPMean),
                        LastPeriodIOPMean - PreOpIOPMean,
                        ifelse(!is.na(LastPeriodAbsIOPChangeMean),
                               LastPeriodAbsIOPChangeMean,
                               ifelse(!is.na(OneYAbsIOPChangeMean),
                                      OneYAbsIOPChangeMean,
                                      OneYIOPMean - PreOpIOPMean
                                      )))
                               , label = study.name, color = subtype)) + geom_point() + ylab('IOP chang
```

```
## Warning: Removed 7 rows containing missing values (geom_point).
```

```
## Warning: Removed 7 rows containing missing values (geom_text_repel).
```

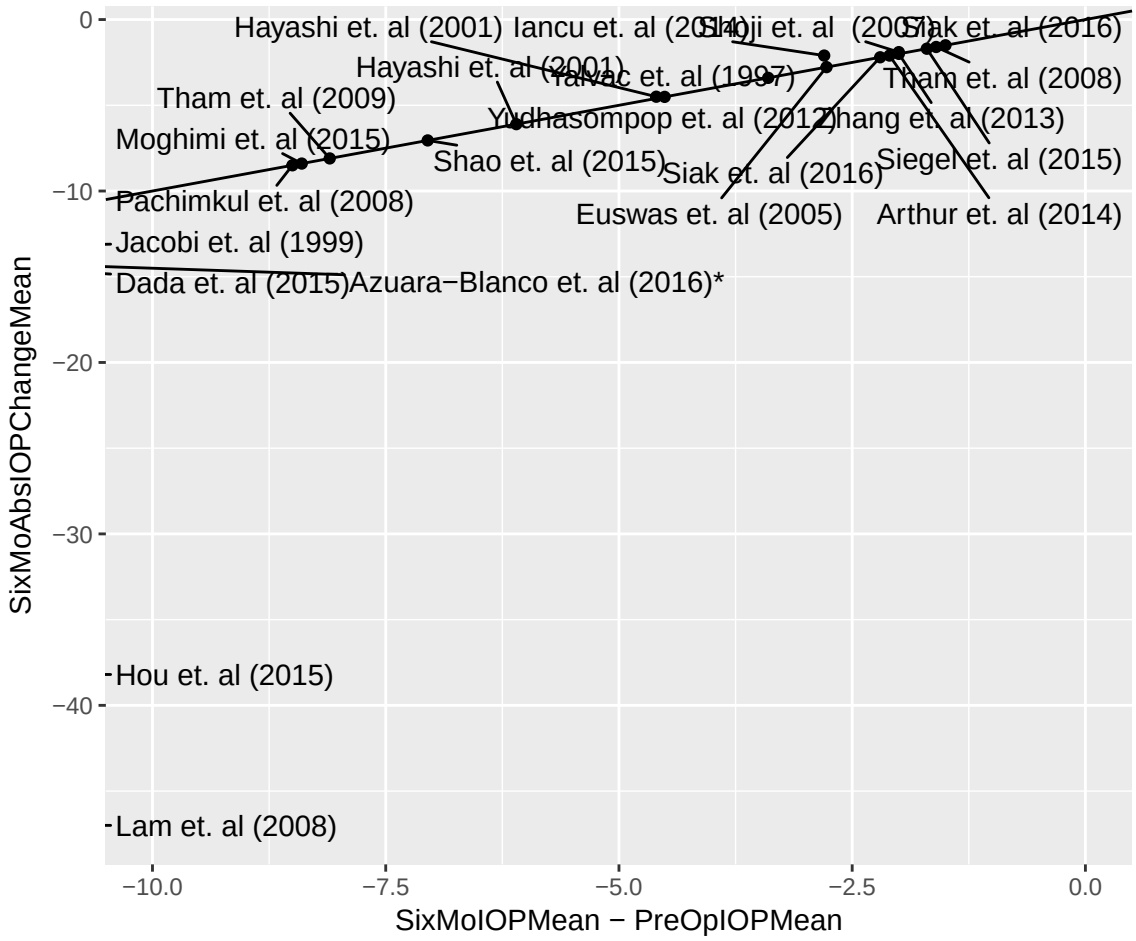


Check that changes add up.

```
ggplot(df, aes(x = SixMoIOPMean - PreOpIOPMean, y = SixMoAbsIOPChangeMean, label = study.name)) +
  geom_point() +
  coord_cartesian(xlim=c(-10, 0)) +
  geom_abline() +
  geom_text_repel()
```

```
## Warning: Removed 44 rows containing missing values (geom_point).
```

```
## Warning: Removed 44 rows containing missing values (geom_text_repel).
```



* Hayashi et al. rejected patients they couldn't follow for 12 months * Yudhasompop was retrospective * Iancu has full follow up * Shoji had low follow up loss * Yalvac was retrospective

Working model: if there's a delta, and no annotation for number of eyes, it's retrospective. Sanity check: the delta should be 0

```
df %>% mutate(delta.sixMo = SixMoIOPMean - PreOpIOPMean - SixMoAbsIOPChangeMean) %>%
  filter(!is.na(delta.sixMo)) %>%
  select(study.name, subtype, delta.sixMo, PreOpEyes, SixMoEyes)
```

```
##                      study.name subtype   delta.sixMo PreOpEyes SixMoEyes
## 1          Yalvac et. al (1997)     OAG  1.332268e-15        35        35
## 2          Jacobi et. al (1999)     PXG  0.000000e+00        22        17
## 3        Hayashi et. al (2001)     OAG -1.000000e-01        68        68
## 4        Hayashi et. al (2001)     ACG  1.776357e-15        74        74
## 5          Euswas et. al (2005)     ACG  3.552714e-15        48        48
## 6          Shoji et. al  (2007)     OAG -1.000000e-01        35        35
## 7             Lam et. al (2008)   acute  0.000000e+00        31        31
## 8       Pachimkul et. al (2008)     ACG  0.000000e+00        58        58
## 9            Tham et. al (2008)     ACG  0.000000e+00        35        35
## 10           Tham et. al (2009)     ACG  0.000000e+00        27        27
## 11    Yudhasompop et. al (2012)     ACG -1.776357e-15        60        60
## 12          Zhang et. al (2013)     OAG  0.000000e+00        43        43
## 13        Arthur et. al (2014)      OAG  0.000000e+00        37        36
## 14          Iancu et. al (2014)     OAG -7.000000e-01        38        38
## 15           Dada et. al (2015)     ACG  0.000000e+00        44        44
```

```
## 16            Hou et. al (2015)   acute  0.000000e+00        25        25
## 17        Moghimi et. al (2015)     ACG  0.000000e+00        46        46
## 18           Shao et. al (2015)     ACG  0.000000e+00        20        20
## 19         Siegel et. al (2015)     OAG  0.000000e+00        52        52
## 20 Azuara-Blanco et. al (2016)*     ACG  0.000000e+00       208       195
## 21           Siak et. al (2016)     ACG  0.000000e+00        24        24
## 22           Siak et. al (2016)     OAG  0.000000e+00        30        30
```

Also: the number from Iancy don't match up 100%. Likely source: sloppy math.

Let's see what happens at one year.

```
df %>% mutate(delta.OneY = OneYIOPMean - PreOpIOPMean - OneYAbsIOPChangeMean) %>%
  filter(!is.na(delta.OneY)) %>%
  select(study.name, subtype, delta.OneY, PreOpEyes, OneYEyes)
```

```
##                      study.name subtype    delta.OneY PreOpEyes OneYEyes
## 1        Perasalo et. al (1997)     PXG  0.000000e+00       226      127
## 2          Jacobi et. al (1999)     PXG  0.000000e+00        22       16
## 3         Hayashi et. al (2001)     OAG -1.000000e-01        68       68
## 4         Hayashi et. al (2001)     ACG -1.000000e-01        74       74
## 5    Leelachaikul et. al (2005)     OAG -4.440892e-16        58       58
## 6       Mathalone et. al (2005)     OAG  0.000000e+00        58       34
## 7          Shoji et. al  (2007)     OAG -2.000000e-01        35       32
## 8             Lam et. al (2008)   acute  0.000000e+00        31       31
## 9     Shingleton et. al  (2008)     PXG  0.000000e+00       240      111
## 10           Tham et. al (2008)     ACG  0.000000e+00        35       35
## 11           Tham et. al (2009)     ACG  0.000000e+00        27       27
## 12            Liu et. al (2011)     ACG  0.000000e+00        56       56
## 13         Arthur et. al (2014)     OAG  0.000000e+00        37       34
## 14          Iancu et. al (2014)     OAG -3.000000e-01        38       38
## 15       Slabaugh et. al (2014)     OAG -1.000000e-02       157      157
## 16           Dada et. al (2015)     ACG  0.000000e+00        44       44
## 17            Hou et. al (2015)   acute  0.000000e+00        25       25
## 18        Moghimi et. al (2015)     ACG  0.000000e+00        46       46
## 19       Pfeiffer et. al (2015)     OAG  0.000000e+00        50       44
## 20         Siegel et. al (2015)     OAG  0.000000e+00        52       52
## 21 Azuara-Blanco et. al (2016)*     ACG  0.000000e+00       208      192
## 22           Siak et. al (2016)     ACG  0.000000e+00        24       24
## 23           Siak et. al (2016)     OAG  0.000000e+00        30       30
```

Problem with assumption: Mathalone doesn't fit this pattern.

- Mathalone is retrospective
- The measured IOPs are actually for those in the followed-up group (N = 34)
- TODO(Patrick): fix the coding.

At the last point in the study:

```
df %>% mutate(delta.LastPeriod = LastPeriodIOPMean - PreOpIOPMean - LastPeriodAbsIOPChangeMean) %>%
  filter(!is.na(delta.LastPeriod)) %>%
  select(study.name, subtype, delta.LastPeriod, PreOpEyes, LastPeriodEyes)
```

```
##                      study.name subtype delta.LastPeriod PreOpEyes
## 1          Jacobi et. al (1999)     PXG     0.000000e+00        22
## 2          Jacobi et. al (1999)     OAG     0.000000e+00        26
## 3             Kim et. al (1999)     OAG     0.000000e+00        31
## 4         Hayashi et. al (2001)     OAG    -2.000000e-01        68
```

```
## 5         Hayashi et. al (2001)     ACG     3.000000e-01        74
## 6          Jacobi et. al (2002)   acute     0.000000e+00        43
## 7    Leelachaikul et. al (2005)     OAG     3.000000e-01        58
## 8       Mathalone et. al (2005)     OAG     7.000000e-01        58
## 9             Lai et. al (2006)     ACG     0.000000e+00        21
## 10         Shoji et. al  (2007)     OAG    -1.000000e-01        35
## 11         Shoji et. al  (2007)     OAG     0.000000e+00        31
## 12            Lam et. al (2008)   acute     0.000000e+00        31
## 13   Mierzejewski et. al (2008)     OAG    -1.776357e-15        52
## 14   Mierzejewski et. al (2008)     ACG     0.000000e+00        25
## 15   Mierzejewski et. al (2008)     PXG     0.000000e+00        23
## 16    Shingleton et. al  (2008)     PXG     0.000000e+00       240
## 17           Tham et. al (2008)     ACG     0.000000e+00        35
## 18           Tham et. al (2009)     ACG     0.000000e+00        27
## 19            Lee et. al (2010)   acute     0.000000e+00        26
## 20         Husain et. al (2012)   acute     0.000000e+00        19
## 21          Shams et. al (2012)     ACG    -1.000000e-01        55
## 22        Klamann et. al (2013)     OAG     0.000000e+00        28
## 23        Klamann et. al (2013)     OAG     0.000000e+00        27
## 24           Tham et. al (2013)     ACG     2.000000e-01        26
## 25         Arthur et. al (2014)     OAG     0.000000e+00        37
## 26         Arthur et. al (2014)     OAG     0.000000e+00        32
## 27           Dias-Santos (2015)     ACG     0.000000e+00        15
## 28           Khan et. al (2015)     OAG     0.000000e+00        49
## 29           Khan et. al (2015)     OAG     0.000000e+00        52
## 30        Moghimi et. al (2015)     OAG     0.000000e+00        45
## 31       Pfeiffer et. al (2015)     OAG     0.000000e+00        50
## 32       Pfeiffer et. al (2015)     OAG     0.000000e+00        50
## 33          Rekas et. al (2015)     OAG     0.000000e+00        29
## 34     Schoenberg et. al (2015)     OAG     0.000000e+00        36
## 35           Shao et. al (2015)     OAG     0.000000e+00        23
## 36       Sheybani et. al (2015)     OAG     0.000000e+00        83
## 37       Sheybani et. al (2015)     OAG     0.000000e+00        37
## 38         Siegel et. al (2015)     OAG     0.000000e+00        52
## 39         Siegel et. al (2015)     OAG     0.000000e+00       261
## 40           Tetz et. al (2015)     OAG     0.000000e+00        30
## 41 Azuara-Blanco et. al (2016)*     ACG     0.000000e+00       208
## 42            Lee et. al (2016)     ACG    -2.220446e-16        56
## 43        Neiweem et. al (2016)     OAG     0.000000e+00       352
## 44         Parikh et. al (2016)     OAG     0.000000e+00       498
## 45           Vold et.al (2016)*     OAG     2.000000e-01       131
## 46           Vold et.al (2016)*     OAG     1.776357e-15       374
##    LastPeriodEyes
## 1              13
## 2              13
## 3              31
## 4              68
## 5              74
## 6              43
## 7              54
## 8              24
## 9              21
## 10             20
## 11             16
```

```
## 12             31
## 13             52
## 14             25
## 15             23
## 16             22
## 17             35
## 18             27
## 19             26
## 20             18
## 21             55
## 22             28
## 23             27
## 24             26
## 25             17
## 26             11
## 27             15
## 28             49
## 29             52
## 30             45
## 31             43
## 32             47
## 33             29
## 34             36
## 35             23
## 36             83
## 37             34
## 38             52
## 39            261
## 40             27
## 41            182
## 42             56
## 43            352
## 44            498
## 45            131
## 46            374
```

That fits into this general pattern. Now let's look at those with no delta (only absolutes).

```
df %>% filter(is.na(SixMoAbsIOPChangeMean) & !is.na(SixMoIOPMean)) %>%
  select(study.name, subtype, PreOpIOPMean, SixMoIOPMean, PreOpEyes, SixMoEyes)
```

```
## [1] study.name   subtype      PreOpIOPMean SixMoIOPMean PreOpEyes
## [6] SixMoEyes
## <0 rows> (or 0-length row.names)
```

- Euswas is retrospective
- Lam had full follow up except one eye that died after 17 months (lung cancer)
- Pachimkul was prospective and had full follow up

Generally however, the SixMoEyes annotation is much more likely in prospective studies than in retrospective studies.

Plan:

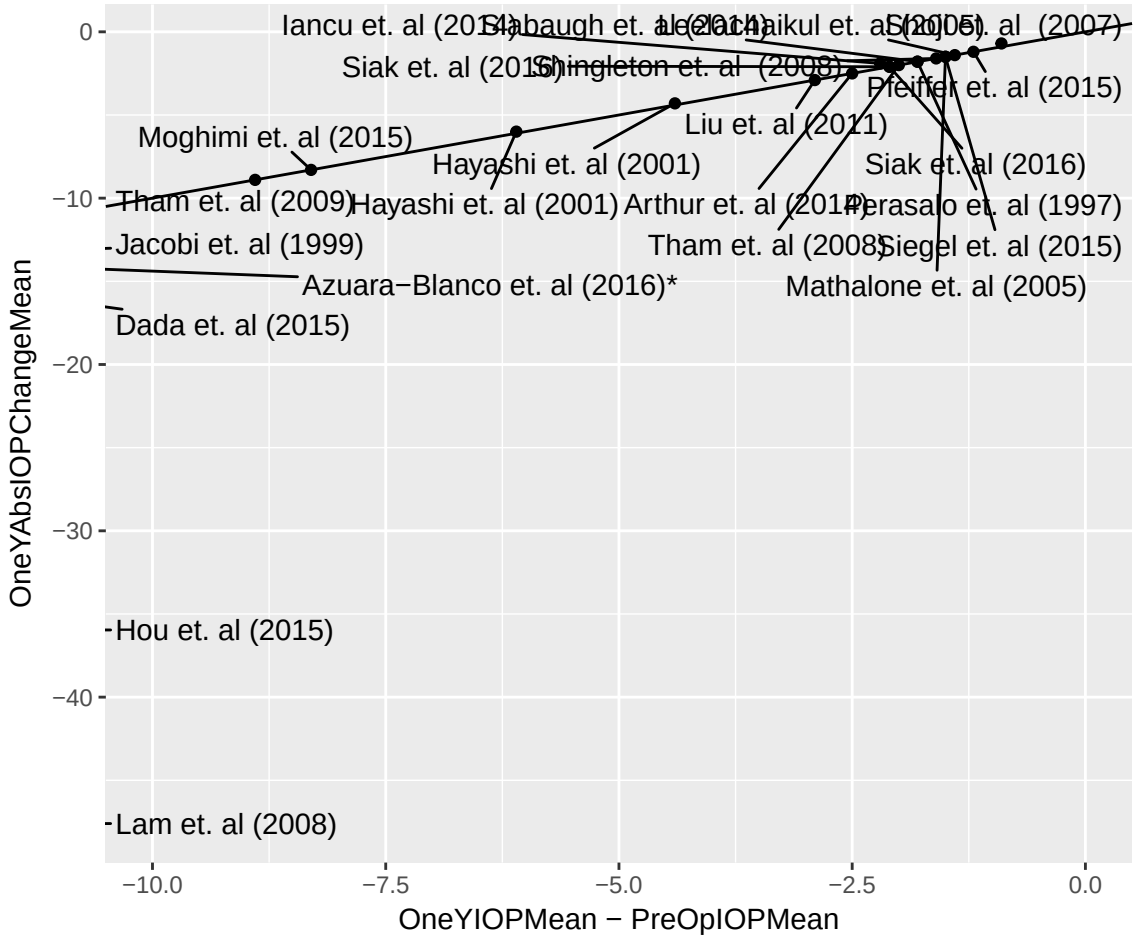
- Impute PutativePropective from pattern of eye studies
- Redo the analysis based on this distinction to see whether it matters much
- Annotate pro or retro in the dataset

- Hound the N eyes values
- Annotate mechanism of exclusion (only included follow through. . . , or whatever)

```
ggplot(df, aes(x = OneYIOPMean - PreOpIOPMean, y = OneYAbsIOPChangeMean, label = study.name)) +
  geom_point() +
  coord_cartesian(xlim=c(-10, 0)) +
  geom_abline() +
  geom_text_repel()
```

```
## Warning: Removed 43 rows containing missing values (geom_point).
```

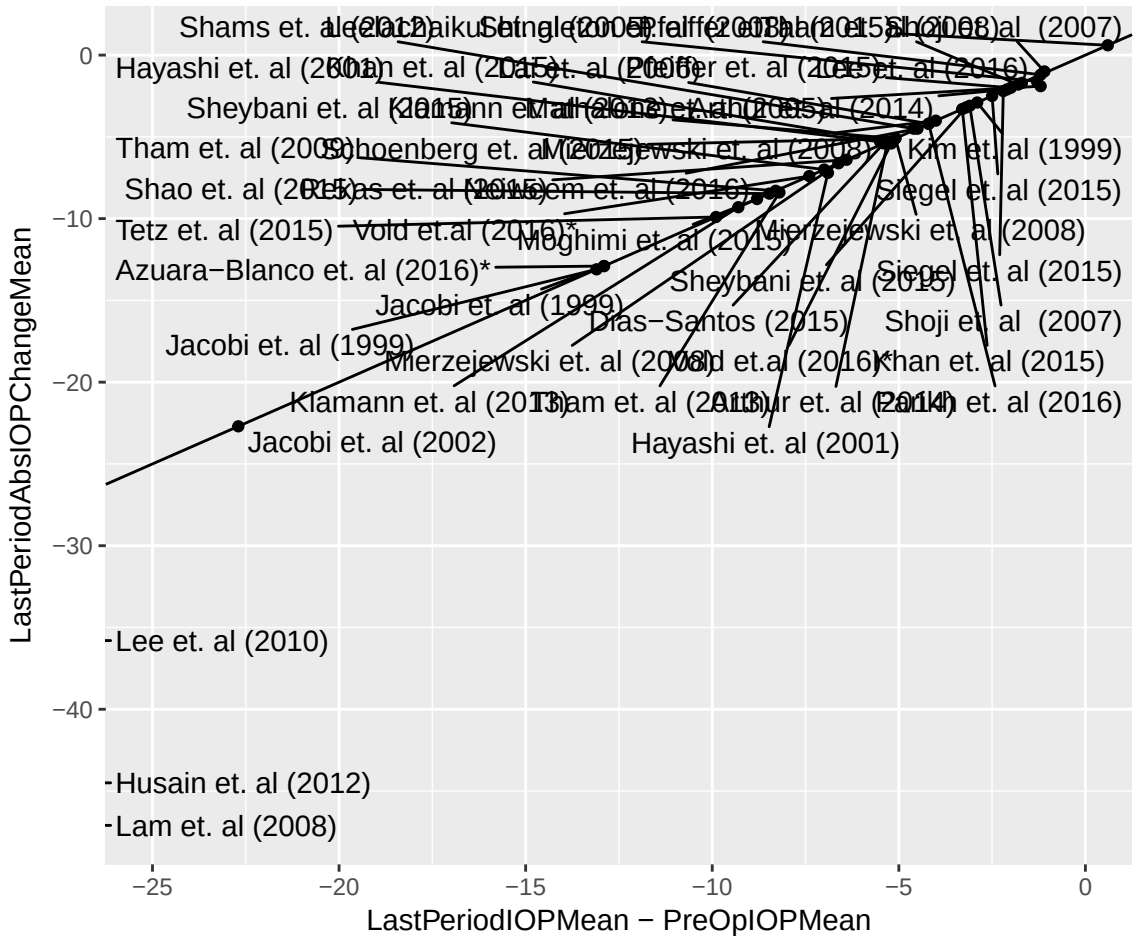
```
## Warning: Removed 43 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x = LastPeriodIOPMean - PreOpIOPMean, y = LastPeriodAbsIOPChangeMean, label = study.name
  geom_point() +
  coord_cartesian(xlim=c(-25, 0)) +
  geom_abline() +
  geom_text_repel()
```

```
## Warning: Removed 20 rows containing missing values (geom_point).
```

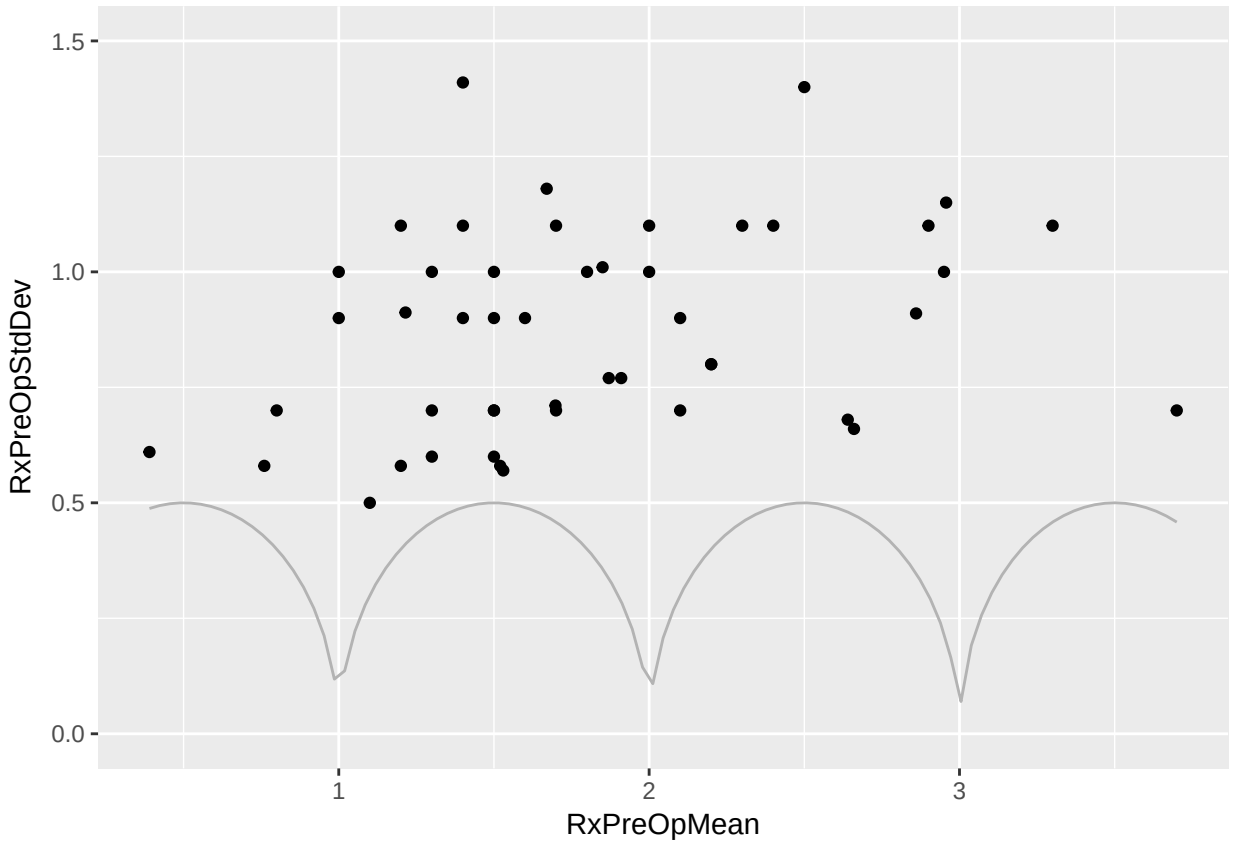
```
## Warning: Removed 20 rows containing missing values (geom_text_repel).
```



Check the relationship between RxPreOpMean and s.d.

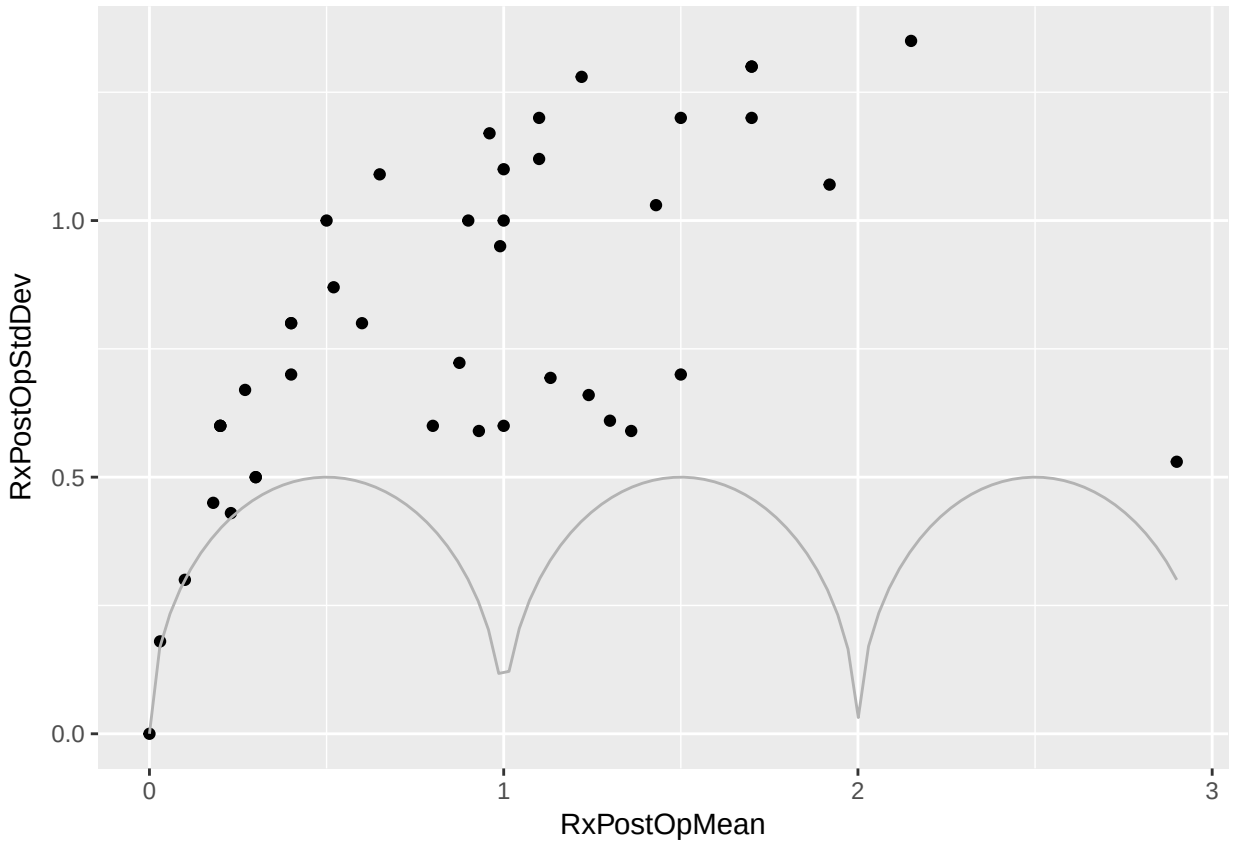
```
ggplot(df, aes(x = RxPreOpMean, y = RxPreOpStdDev)) +
  geom_point() +
  coord_cartesian(y=c(0, 1.5)) +
  stat_function(fun = function(x) sqrt((x - floor(x)) * (1 - (x - floor(x)))), color="gray70")
```

```
## Warning: Removed 18 rows containing missing values (geom_point).
```




```
ggplot(df, aes(x = RxPostOpMean, y = RxPostOpStdDev)) + geom_point() +
  stat_function(fun = function(x) sqrt((x - floor(x)) * (1 - (x - floor(x)))), color="gray70")
```

```
## Warning: Removed 23 rows containing missing values (geom_point).
```

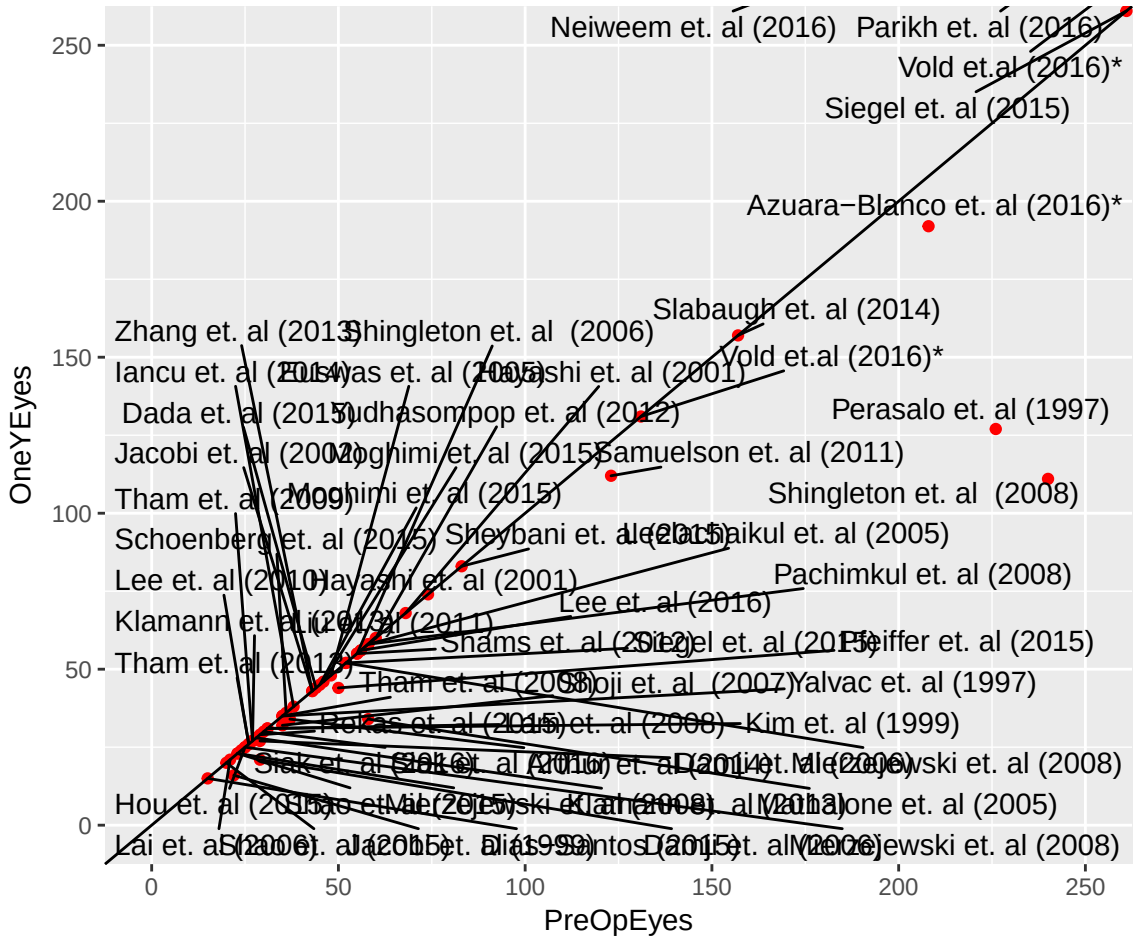


Examine loss at one year.

```
ggplot(df, aes(x=PreOpEyes, y=OneYEyes, label=study.name)) +
  geom_point(color="red") +
  geom_abline() +
  geom_text_repel() + coord_cartesian(xlim=c(0, 250), ylim=c(0, 250))
```

```
## Warning: Removed 10 rows containing missing values (geom_point).
```

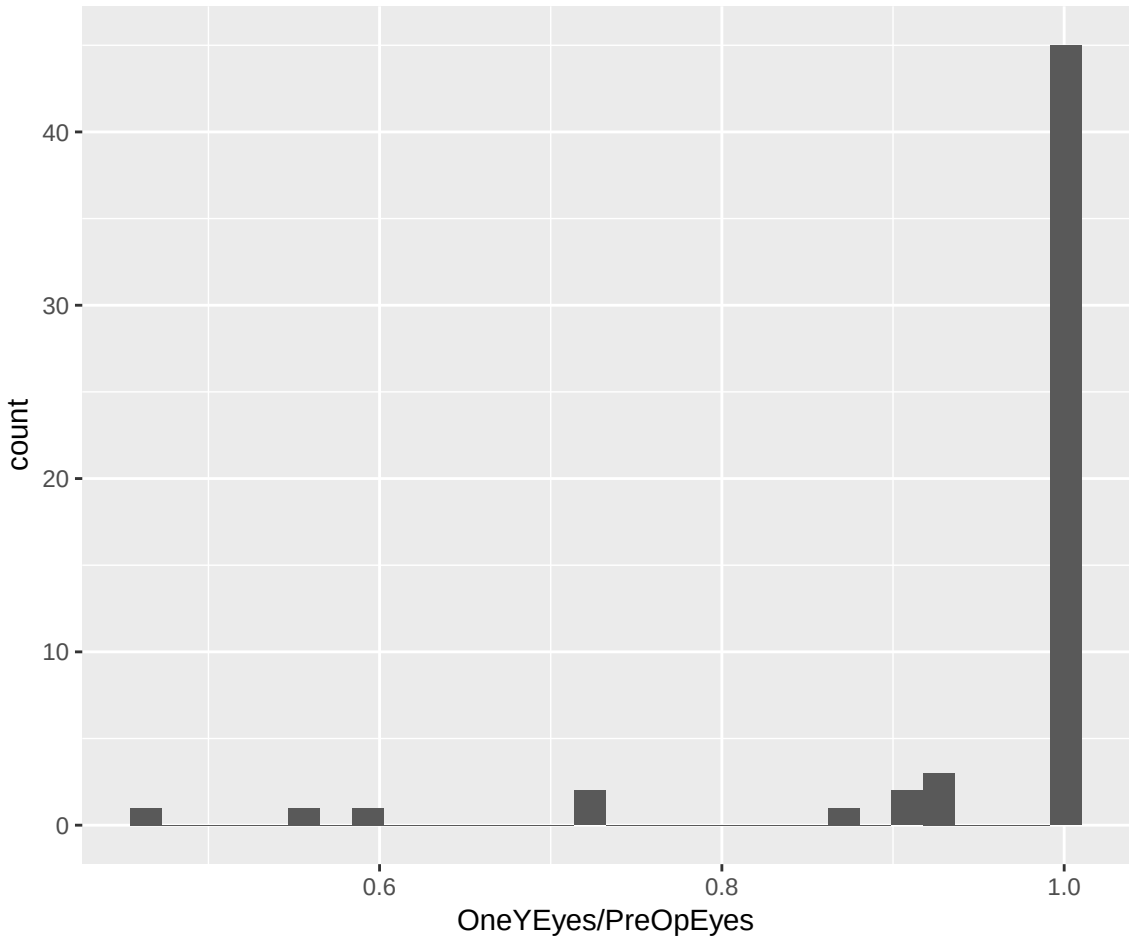
```
## Warning: Removed 10 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=OneYEyes / PreOpEyes, label=study.name)) +
  geom_histogram() +
  coord_cartesian()
```

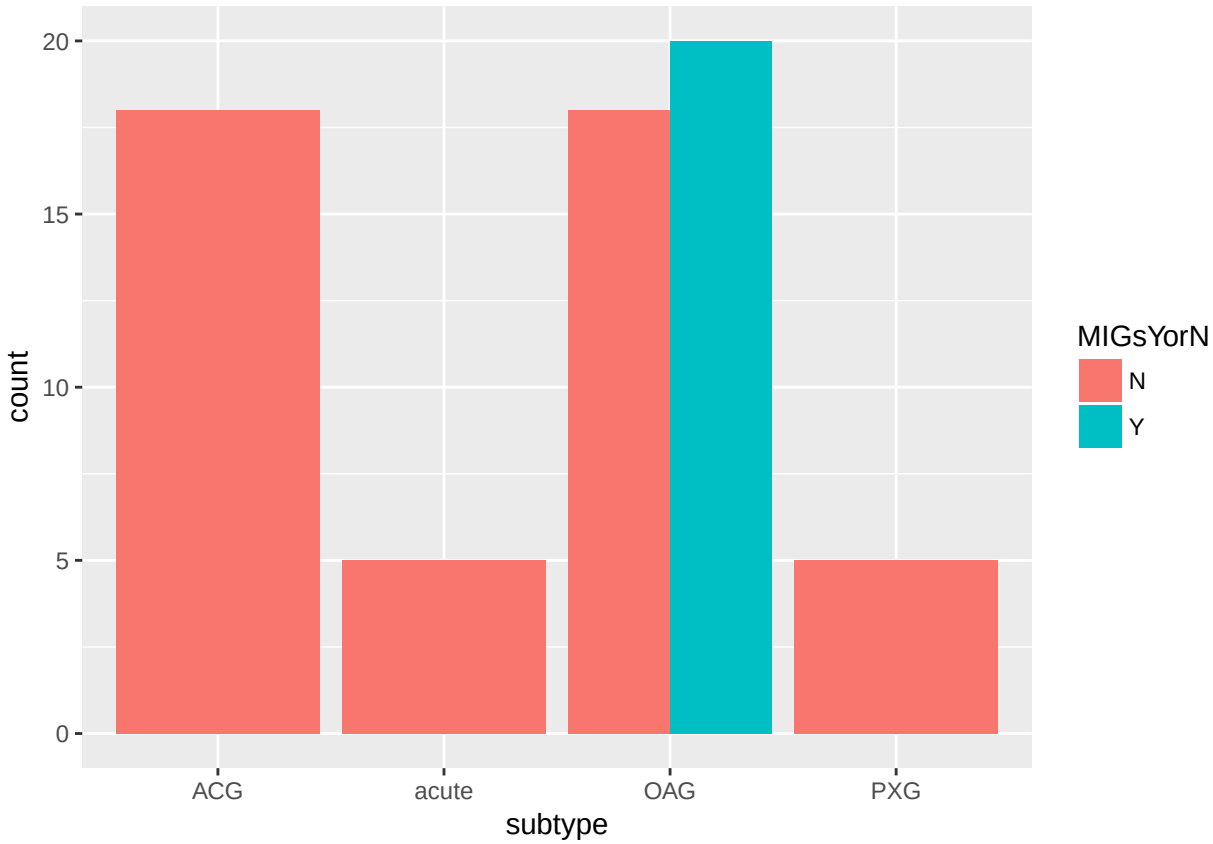
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 10 rows containing non-finite values (stat_bin).
```



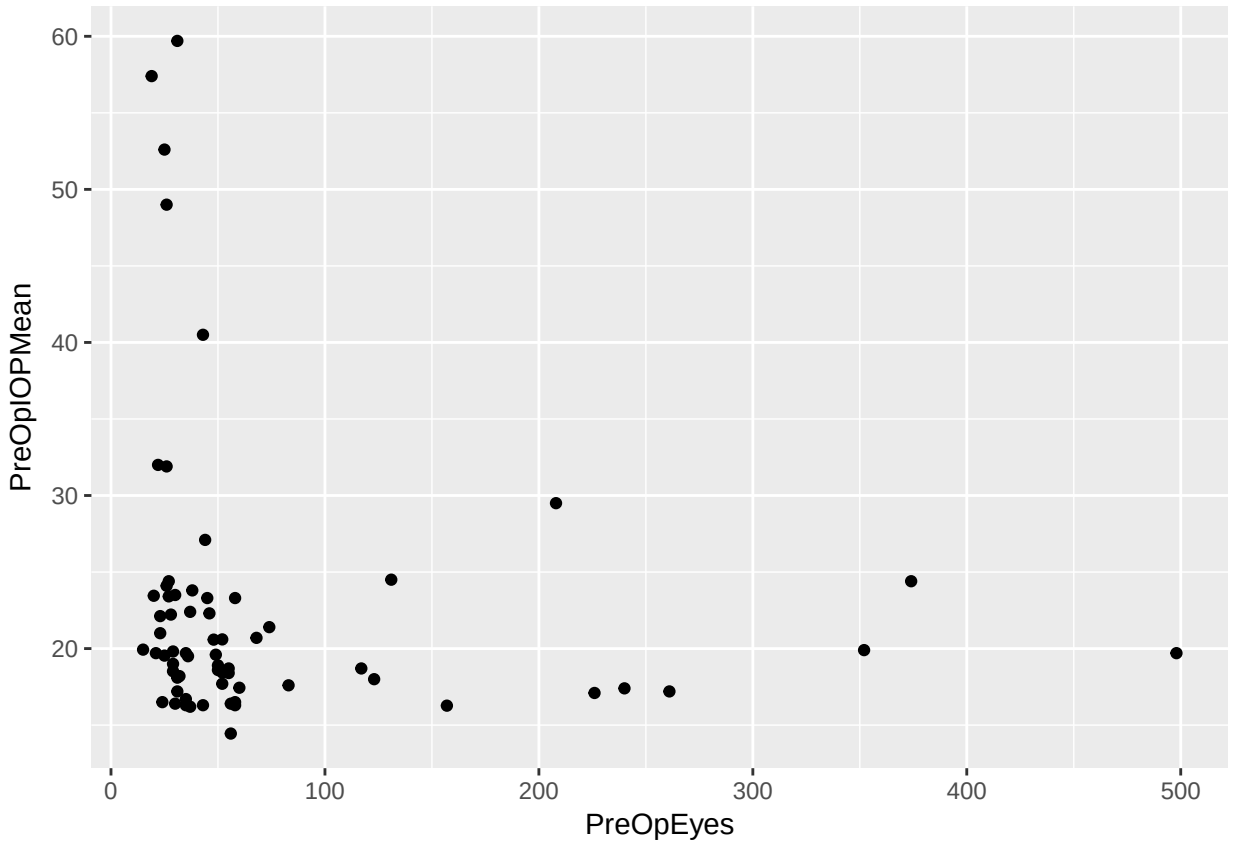

```
ggplot(df, aes(x=subtype, fill=MIGsYorN)) + geom_histogram(stat="count", position = 'dodge')
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



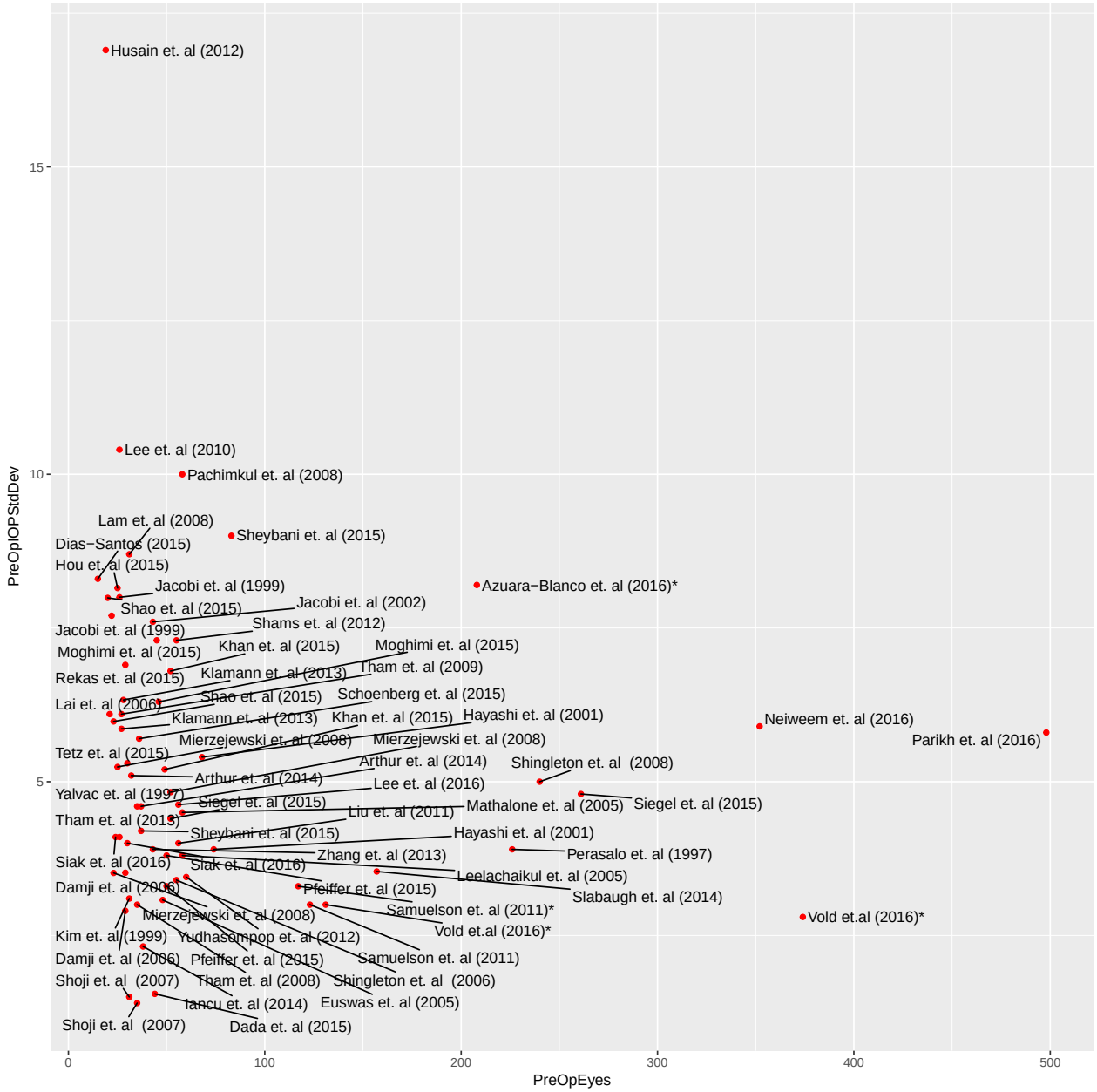
Look at the distribution of eyes and IOP means.

```
ggplot(df, aes(x=PreOpEyes, y=PreOpIOPMean)) + geom_point()
```



Look at number of eyes and standard deviation.

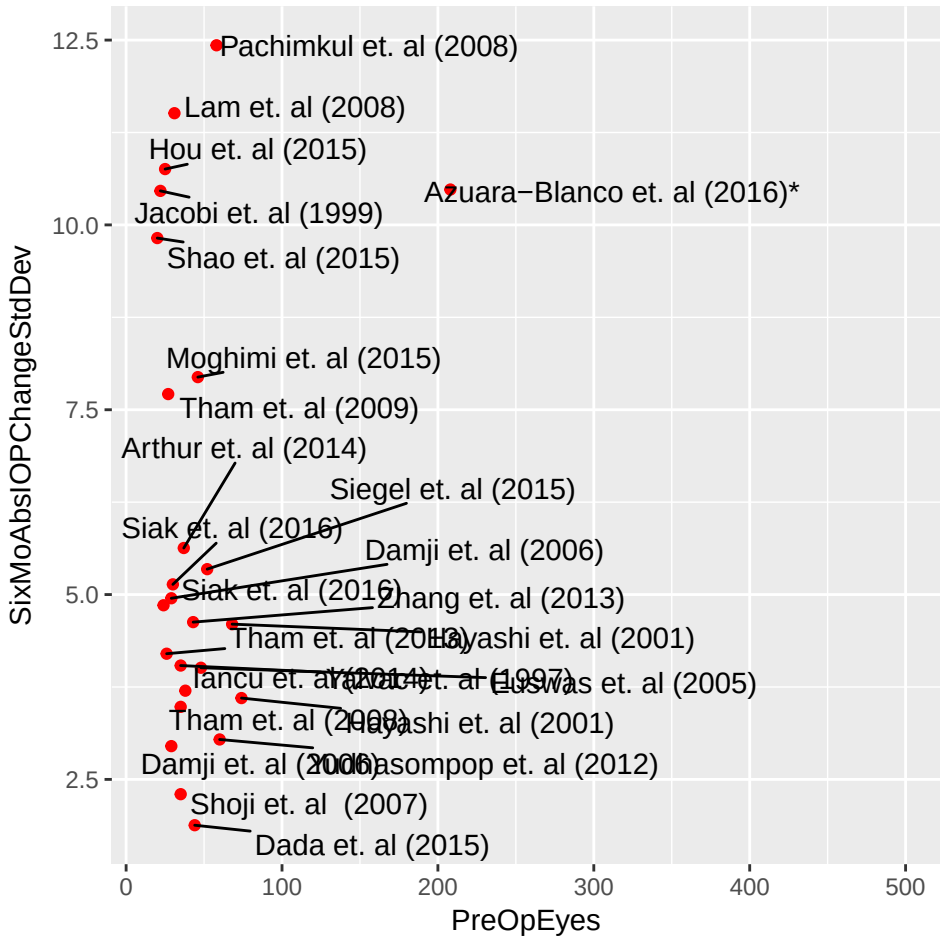
```
ggplot(df, aes(x=PreOpEyes, y=PreOpIOPStdDev, label=study.name)) + geom_point(color="red") + geom_text_
```




```
ggplot(df, aes(x=PreOpEyes, y=SixMoAbsIOPChangeStdDev, label=study.name)) + geom_point(color="red") + g
```

```
## Warning: Removed 41 rows containing missing values (geom_point).
```

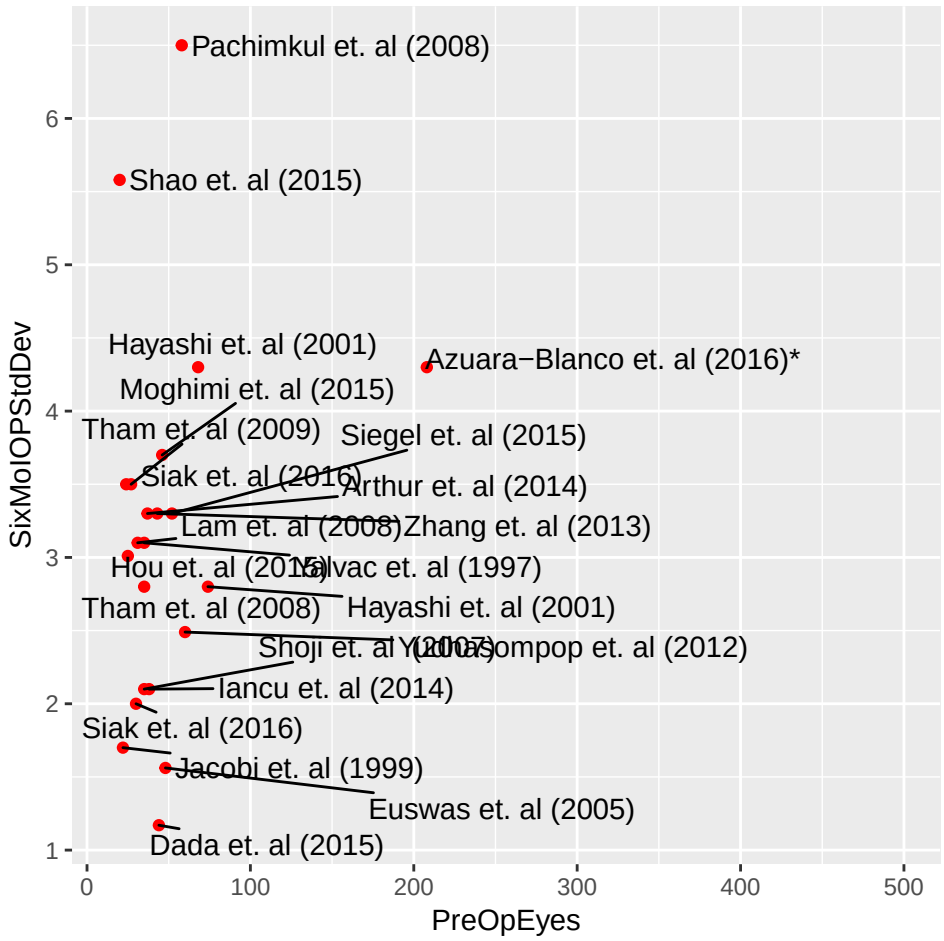
```
## Warning: Removed 41 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=PreOpEyes, y=SixMoIOPStdDev, label=study.name)) + geom_point(color="red") + geom_text_
```

```
## Warning: Removed 44 rows containing missing values (geom_point).
```

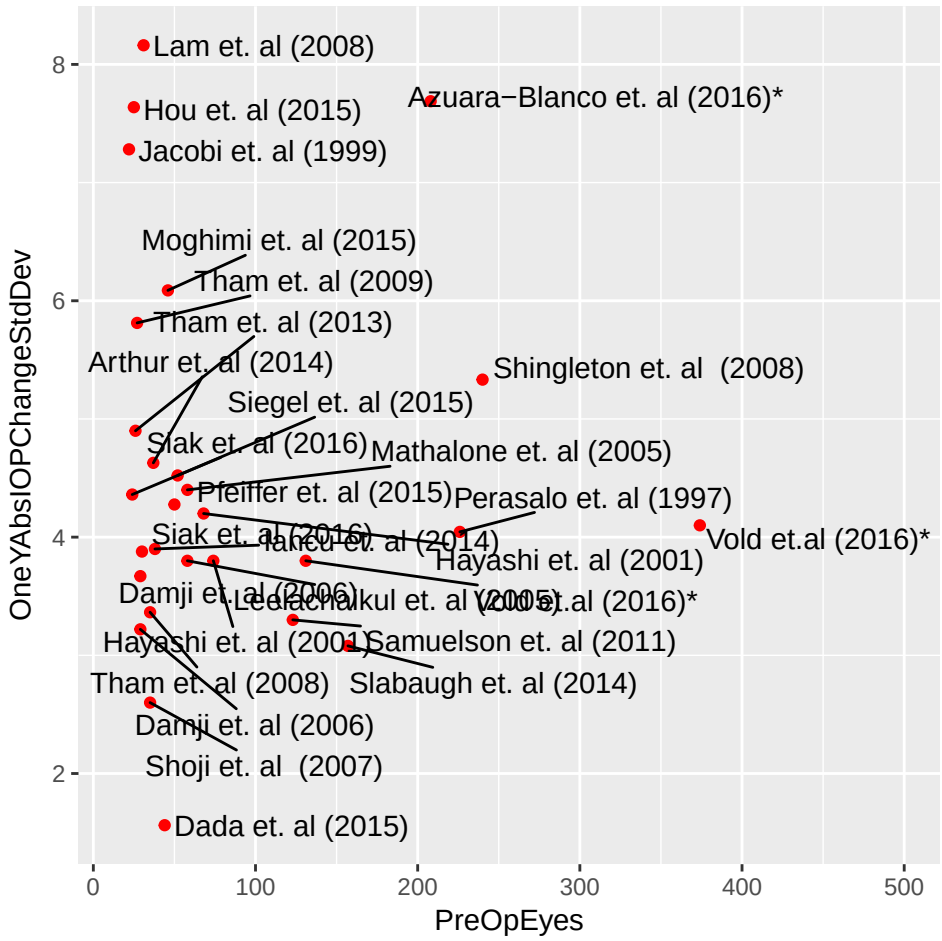
```
## Warning: Removed 44 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=PreOpEyes, y=OneYAbsIOPChangeStdDev, label=study.name)) + geom_point(color="red") + ge
```

```
## Warning: Removed 38 rows containing missing values (geom_point).
```

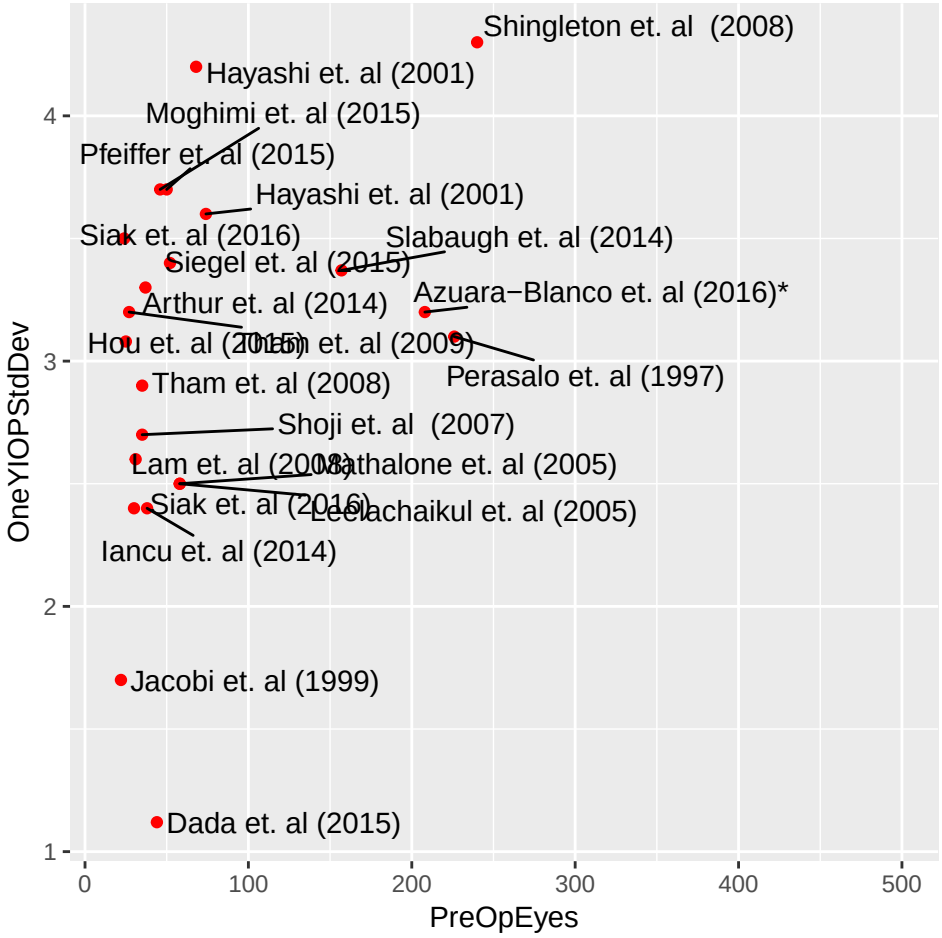
```
## Warning: Removed 38 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=PreOpEyes, y=OneYIOPStdDev, label=study.name)) + geom_point(color="red") + geom_text_r
```

```
## Warning: Removed 44 rows containing missing values (geom_point).
```

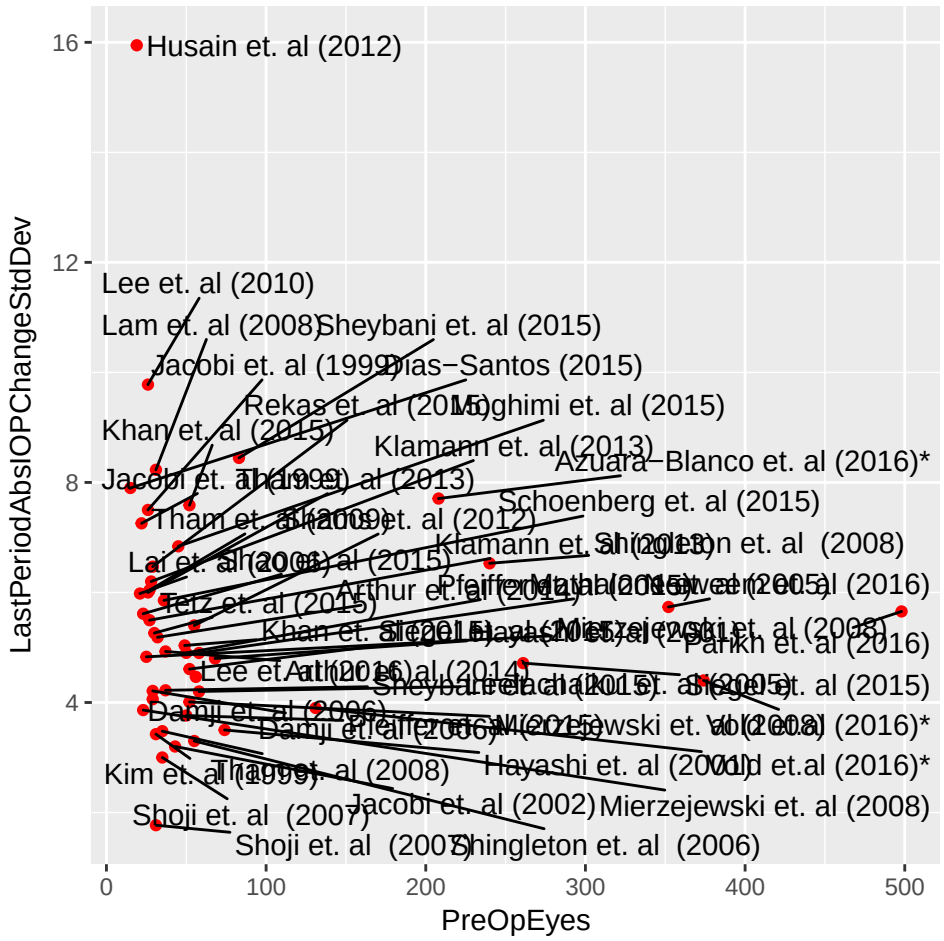
```
## Warning: Removed 44 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=PreOpEyes, y=LastPeriodAbsIOPChangeStdDev, label=study.name)) + geom_point(color="red"
```

```
## Warning: Removed 17 rows containing missing values (geom_point).
```

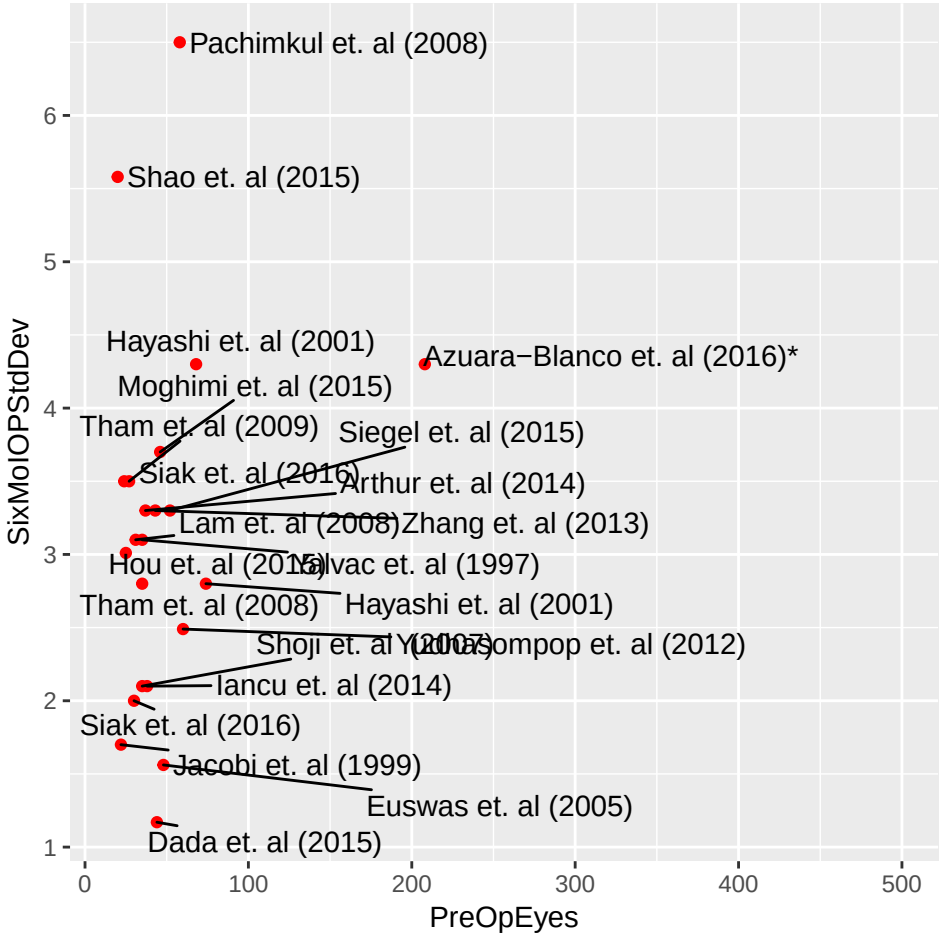
```
## Warning: Removed 17 rows containing missing values (geom_text_repel).
```




```
ggplot(df, aes(x=PreOpEyes, y=SixMoIOPStdDev, label=study.name)) + geom_point(color="red") + geom_text_
```

```
## Warning: Removed 44 rows containing missing values (geom_point).
```

```
## Warning: Removed 44 rows containing missing values (geom_text_repel).
```


Pachimkul et. al (2008)
Shao et. al (2015)
Hayashi et. al (2001)
Azuara−Blanco et. al (2016)*
Moghimi et. al (2015)
Tham et. al (2009)
Siegel et. al (2015)
Siak et. al (2016)
Arthur et. al (2014)
Lam et. al (2008)
Zhang et. al (2013)
Hou et. al (2016)
Salvac et. al (1997)
Tham et. al (2008)
Hayashi et. al (2001)
Shoji et. al (2010)
Iancu et. al (2014)
Siak et. al (2016)
Jacobi et. al (1999)
Euswas et. al (2005)
Dada et. al (2015)
SixMoIOPStdDev
PreOpEyes
1
2
3
4
5
6
0
100
200
300
400
500