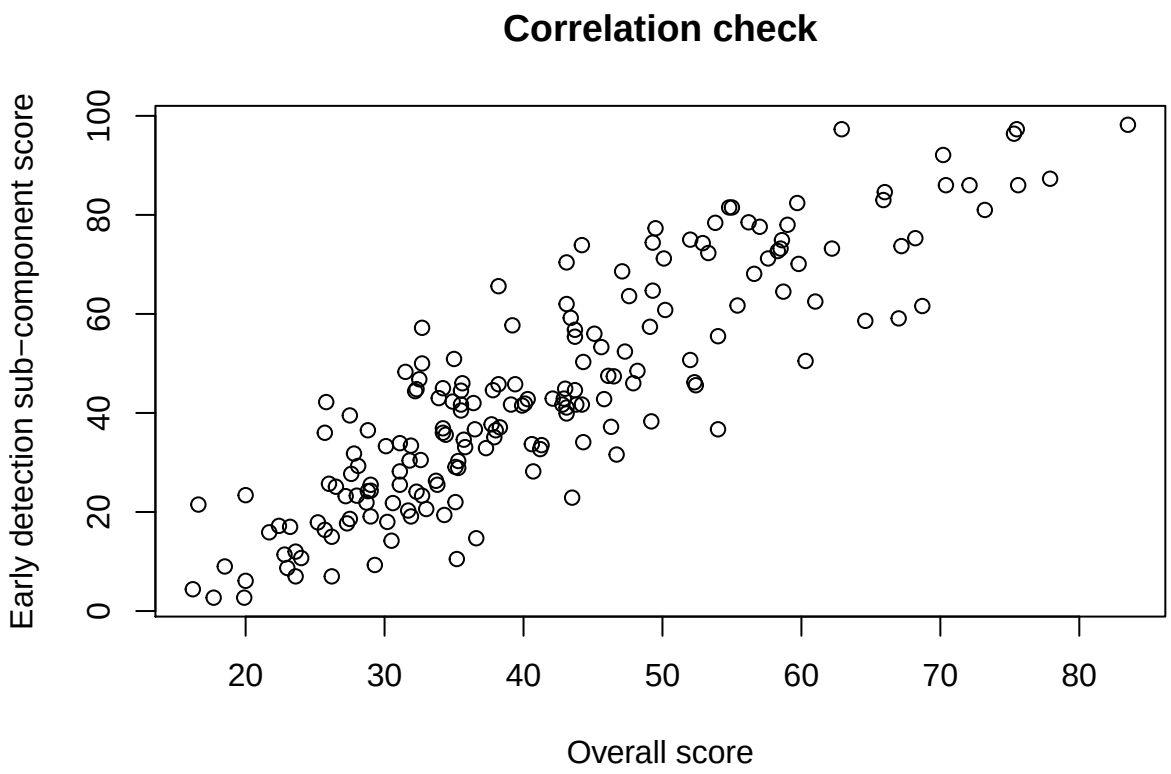
# Reporting systems sub-analysis

Emily Linebarger

2/17/2021

First, run a simple scatter to see what the correlation between the overall score and the "early detection" sub-




component looks like.

Ouch they are really highly correlated. Also check the inverse coefficient matrix.

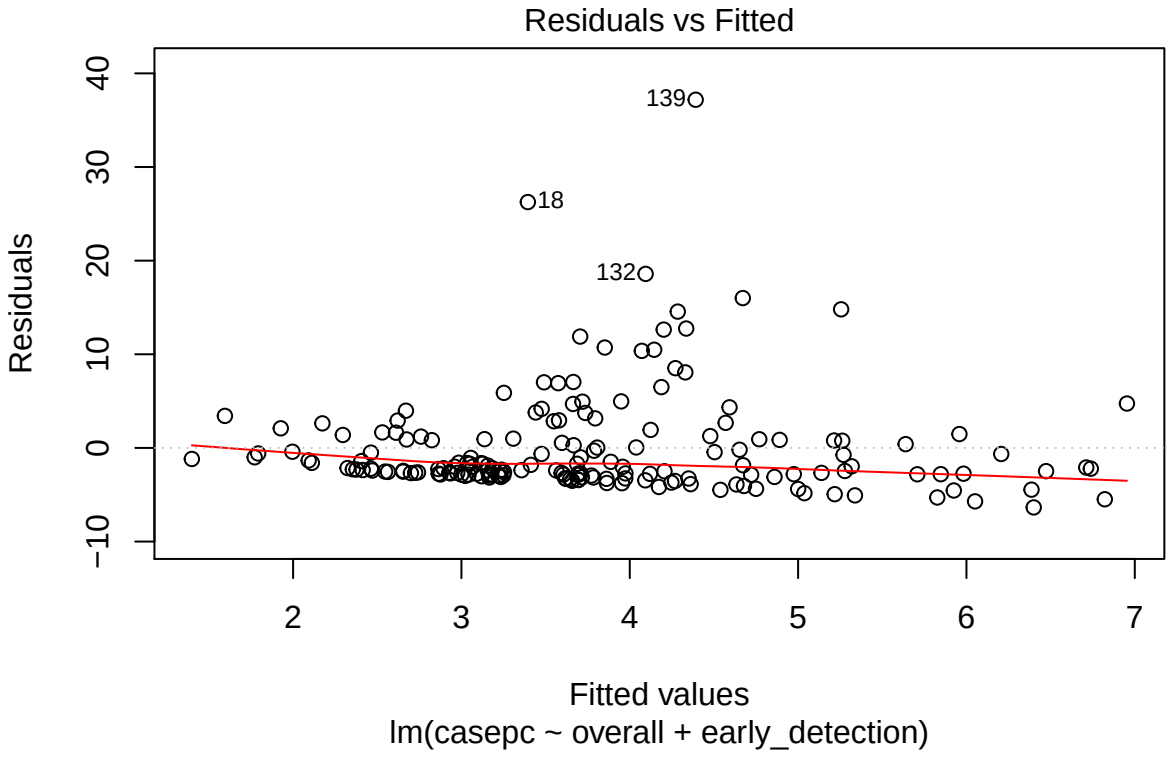
```
##               [,1]          [,2]          [,3]
## [1,]  0.0637661238 -0.0021541733  7.051677e-04
## [2,] -0.0021541733  0.0001235139 -6.721070e-05
## [3,]  0.0007051677 -0.0000672107  4.712524e-05
```

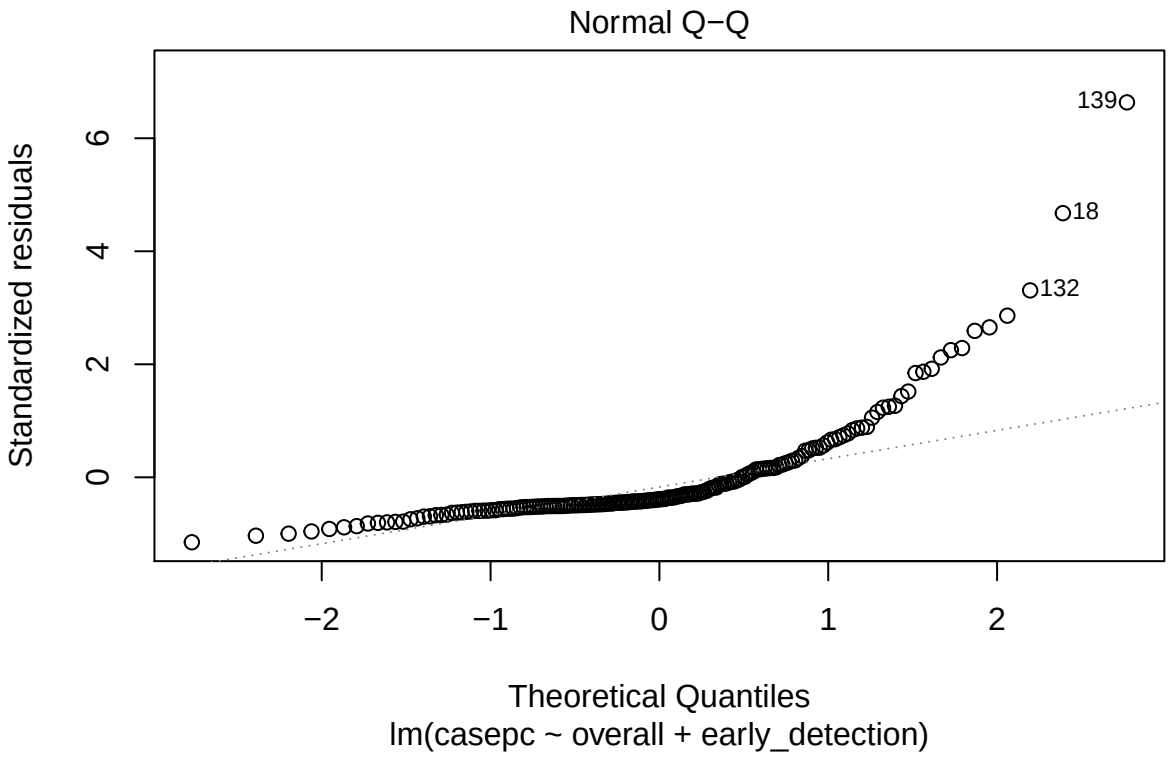
The off-diagonals for "overall" and "early_detection" are not zero, which tells us they are not independent.

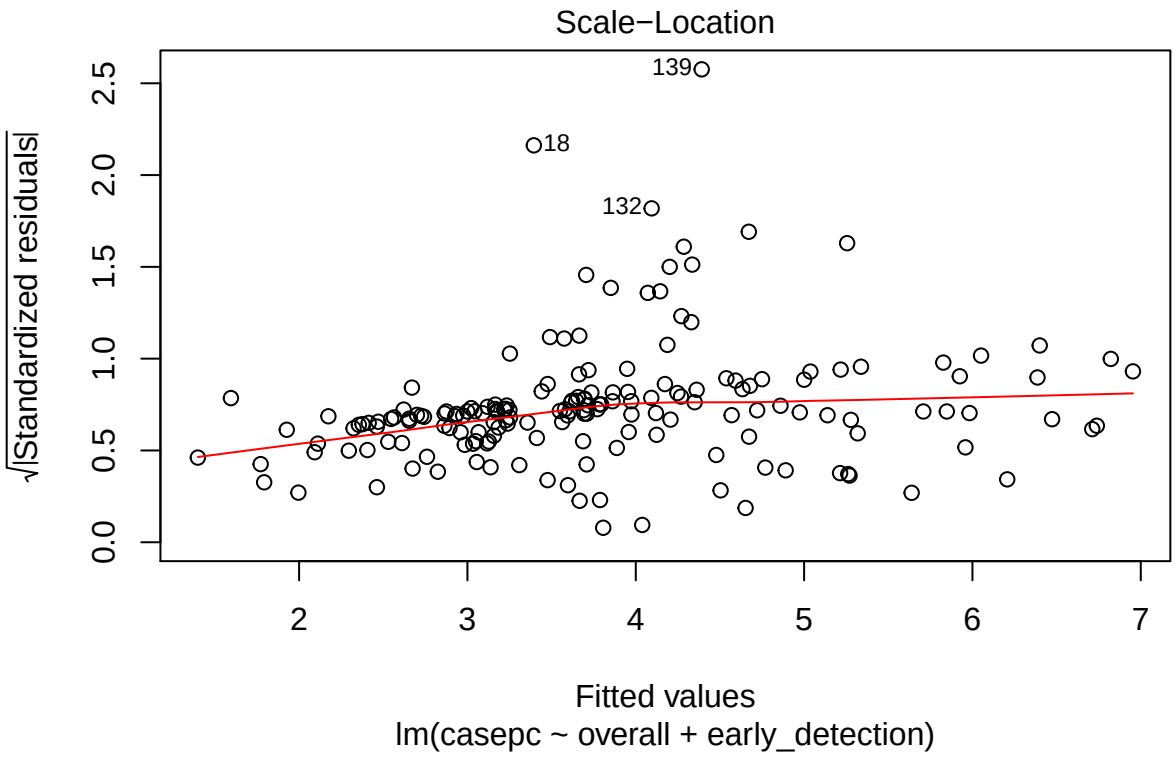
## Outlier screening, with more lenient screen

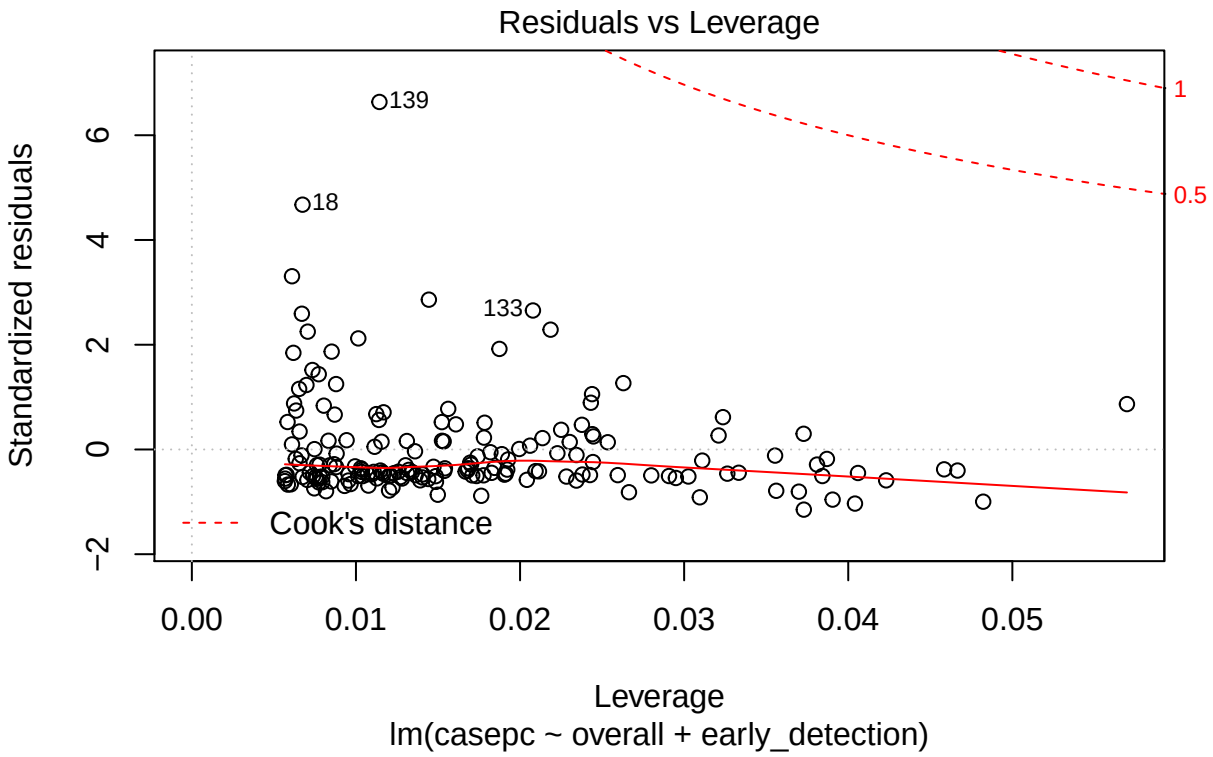
### Cases per-capita

```
##
## Call:
## lm(formula = casepc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.353 -2.852 -2.178  0.933 37.183
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.15810    1.44013   0.110   0.9127
## overall          0.14712    0.06295   2.337   0.0206 *
## early_detection -0.05590    0.03873  -1.443   0.1507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.636 on 175 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.03907,    Adjusted R-squared:  0.02808
## F-statistic: 3.557 on 2 and 175 DF,  p-value: 0.0306
```



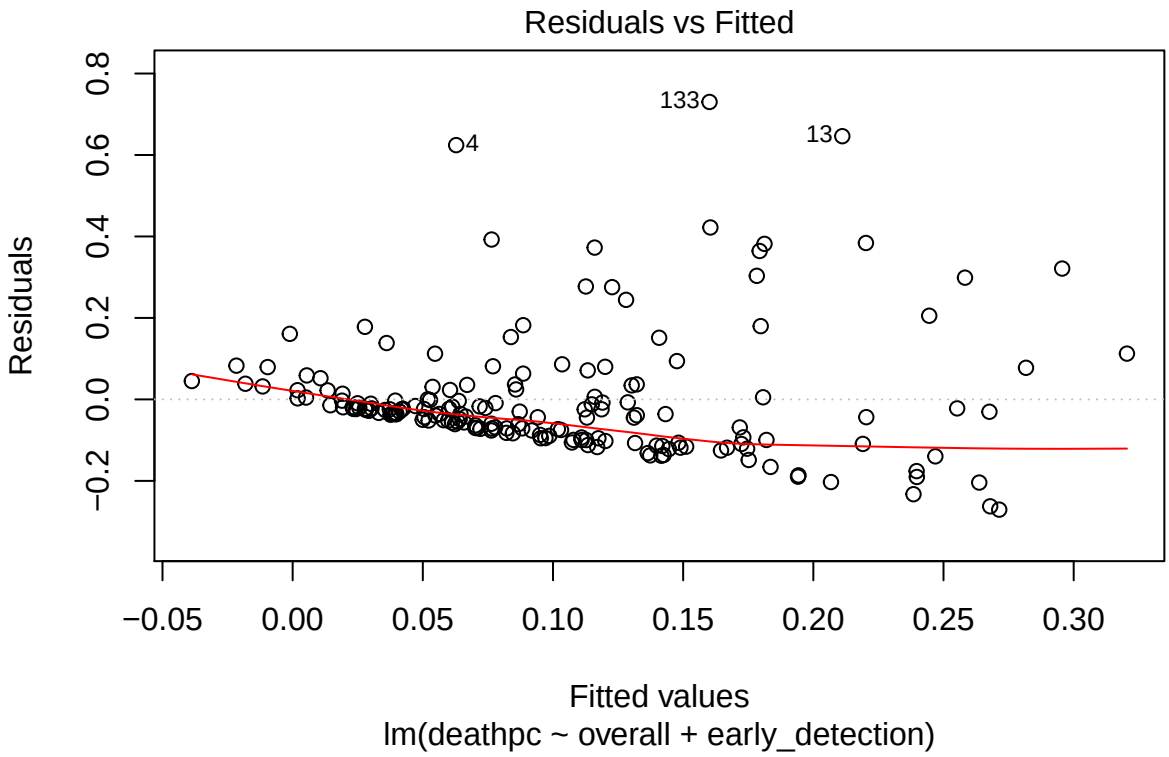

Normal Q−Q
Standardized residuals
0
2
4
6
139
18
132
−2
−1
0
1
2
Theoretical Quantiles
lm(casepc ~ overall + early_detection)

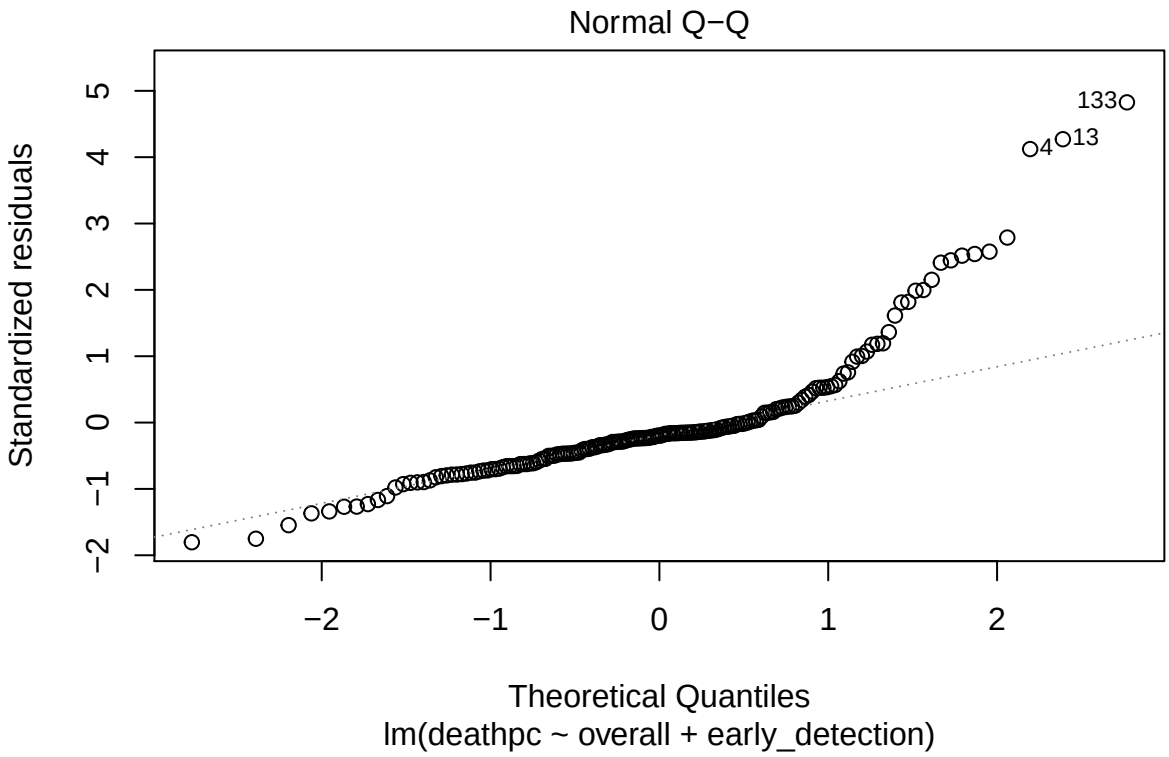

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
139
18
132
2
3
4
5
6
7
Fitted values
lm(casepc ~ overall + early_detection)

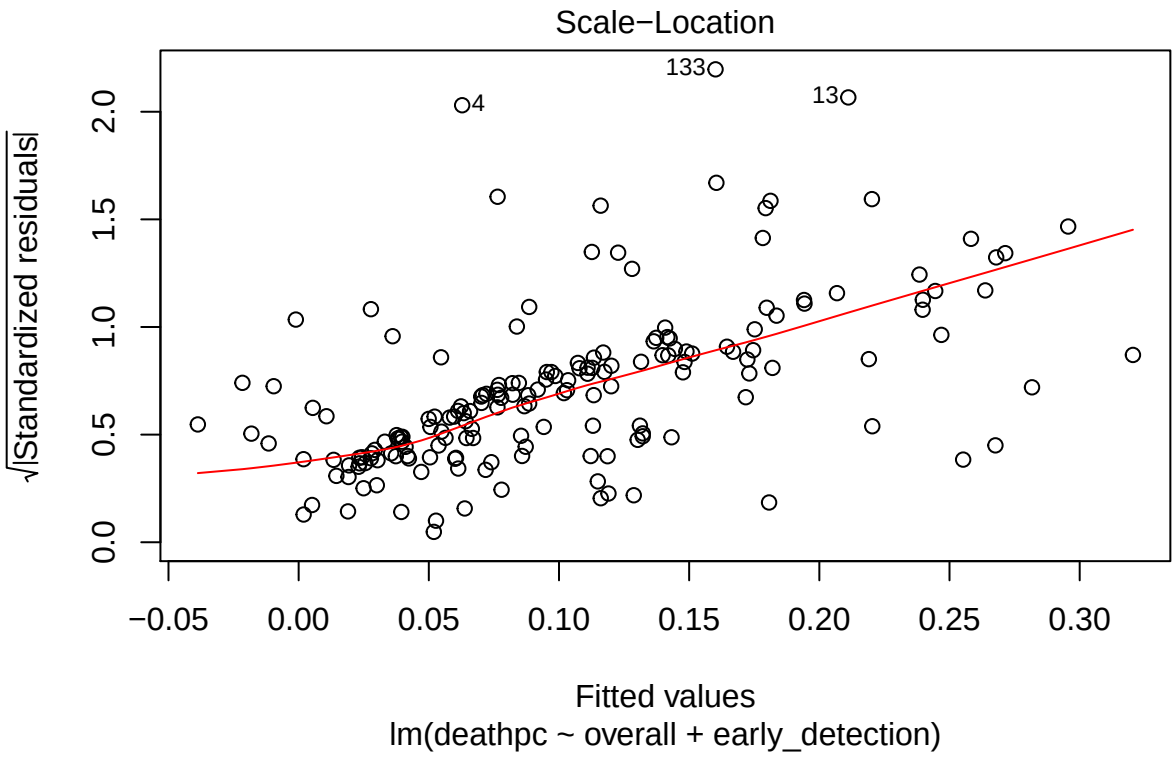


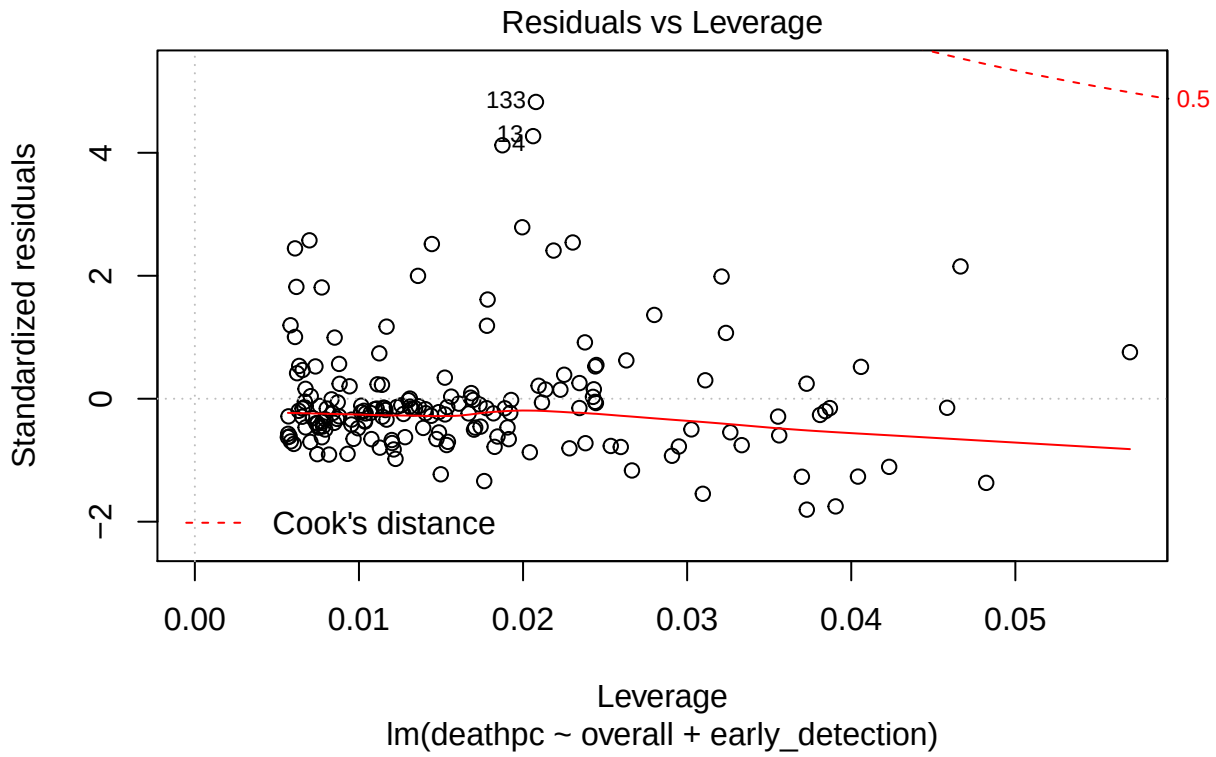

### Deaths per-capita

```
##
## Call:
## lm(formula = deathpc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27057 -0.08066 -0.02993  0.02415  0.73017
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.124988   0.039062  -3.200 0.001633 **
## overall          0.006697   0.001708   3.922 0.000126 ***
## early_detection -0.001158   0.001050  -1.102 0.271822
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1529 on 175 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1849, Adjusted R-squared:  0.1756
## F-statistic: 19.85 on 2 and 175 DF,  p-value: 1.705e-08
```


Residuals vs Fitted
Residuals
Fitted values
lm(deathpc ~ overall + early_detection)
133
4
13
0.8
0.6
0.4
0.2
0.0
−0.2
−0.05
0.00
0.05
0.10
0.15
0.20
0.25
0.30

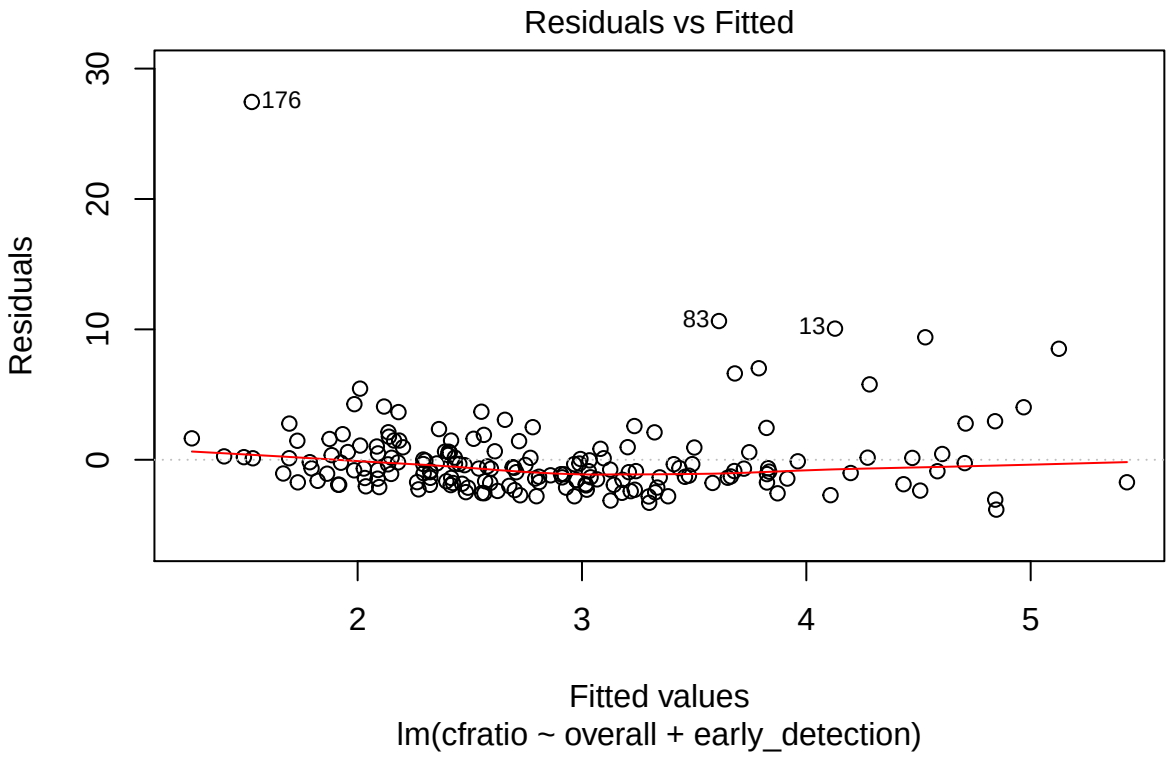

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(deathpc ~ overall + early_detection)
133
4
13

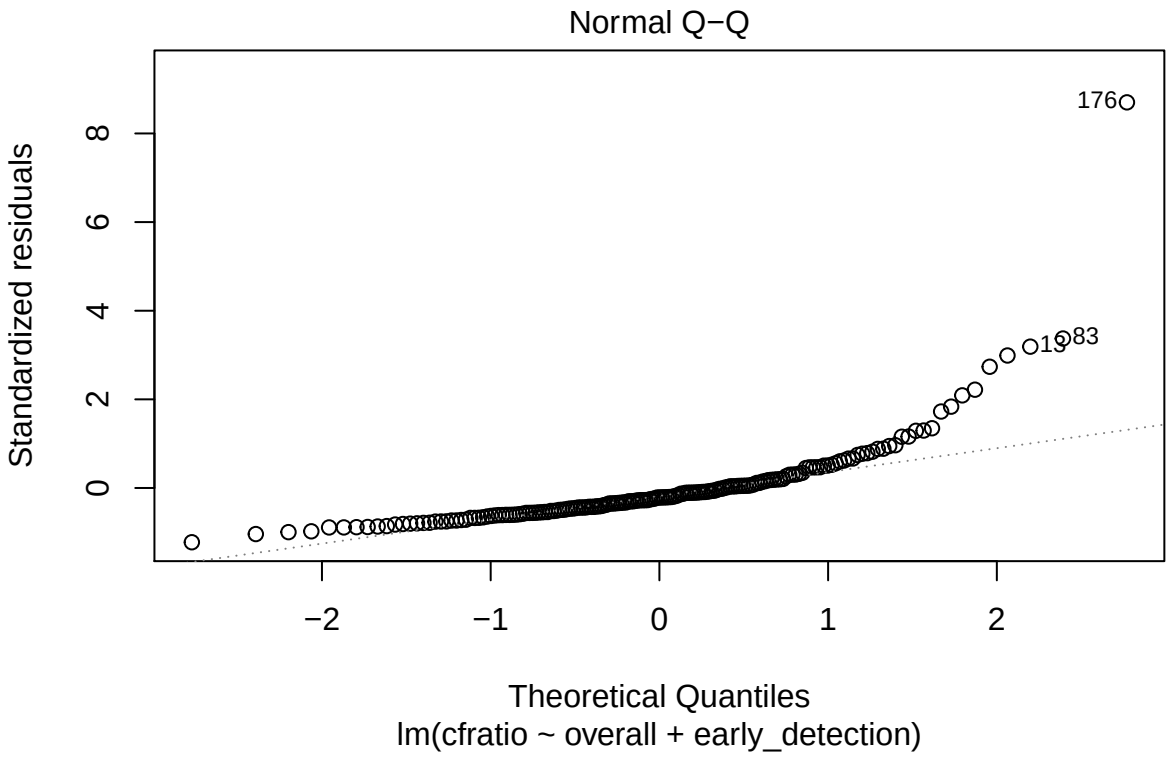

Scale−Location
133
13
4
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
−0.05
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Fitted values
lm(deathpc ~ overall + early_detection)

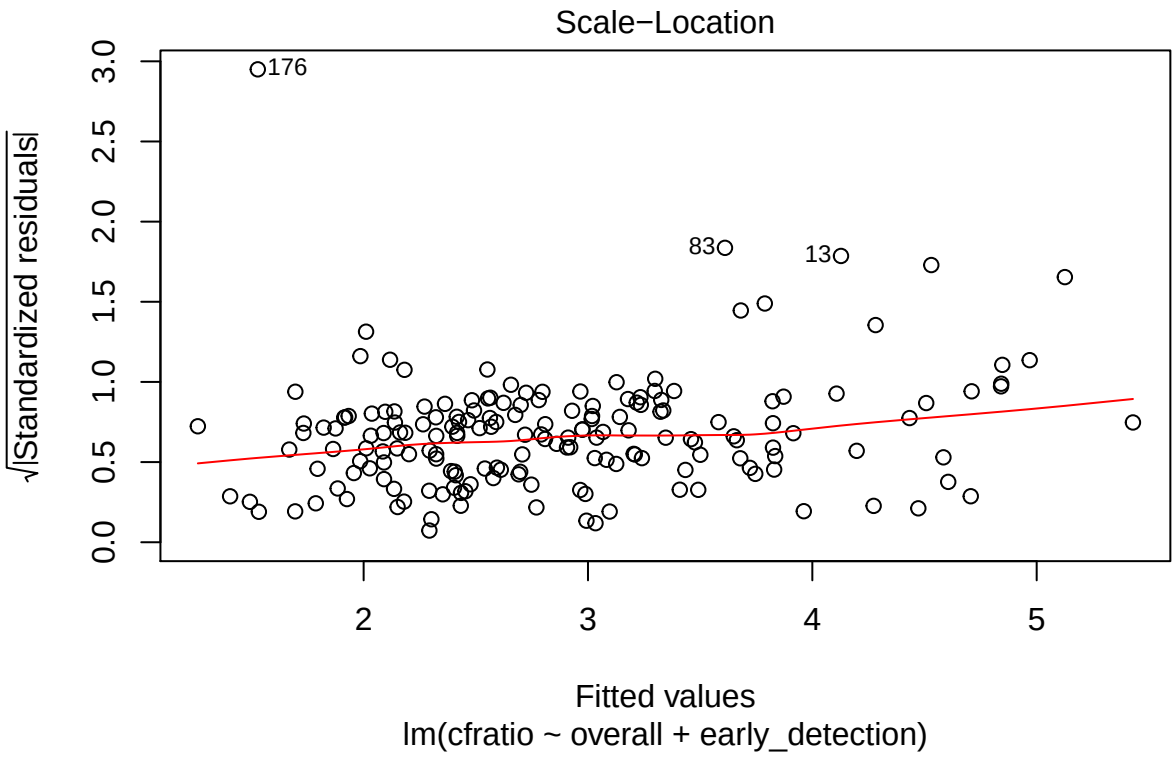


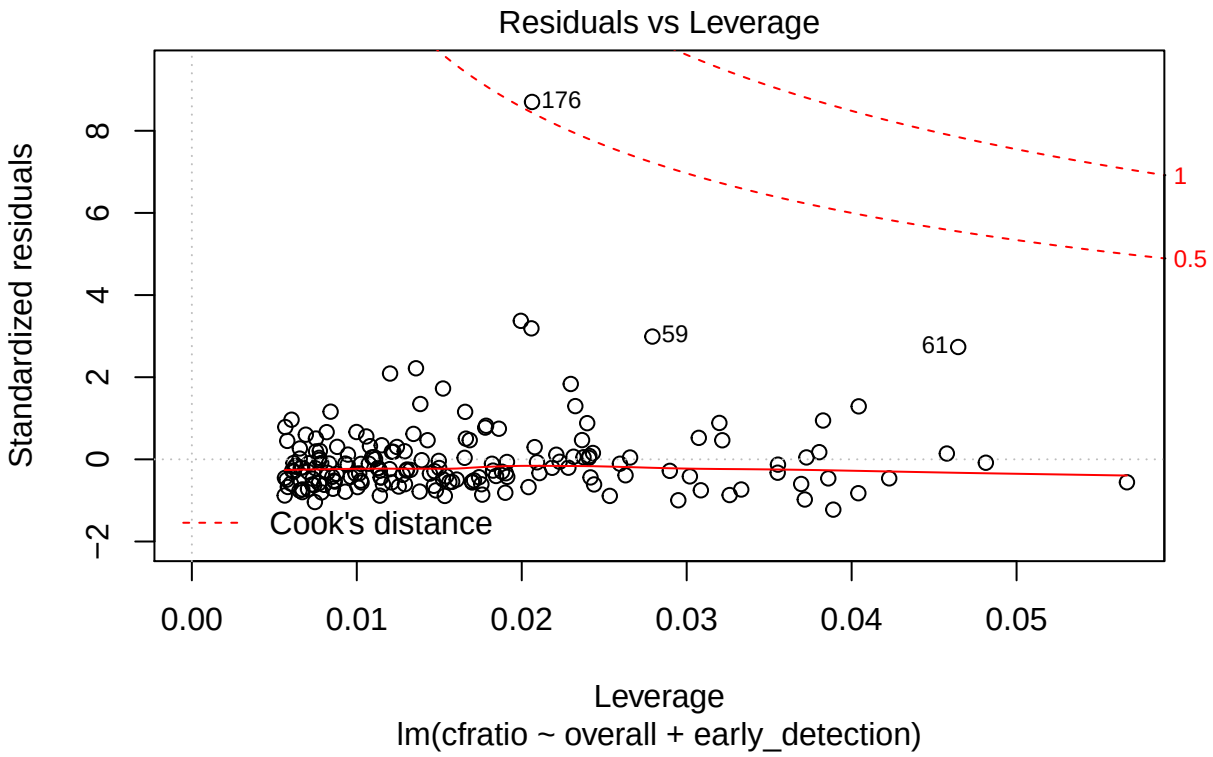

### Case-fatality ratio

```
##
## Call:
## lm(formula = cfratio ~ overall + early_detection, data = regression_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8244 -1.7151 -0.6816  0.5830 27.4410
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.25183    0.80926   0.311    0.756
## overall          0.07393    0.03555   2.080    0.039 *
## early_detection -0.01014    0.02189  -0.463    0.644
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.187 on 176 degrees of freedom
## Multiple R-squared:  0.06706,    Adjusted R-squared:  0.05646
## F-statistic: 6.326 on 2 and 176 DF,  p-value: 0.002223
```


Residuals vs Fitted
Residuals
30
20
10
0
176
83
13
2
3
4
5
Fitted values
lm(cfratio ~ overall + early_detection)


Normal Q−Q
Standardized residuals
0
2
4
6
8
176
183
−2
−1
0
1
2
Theoretical Quantiles
lm(cfratio ~ overall + early_detection)

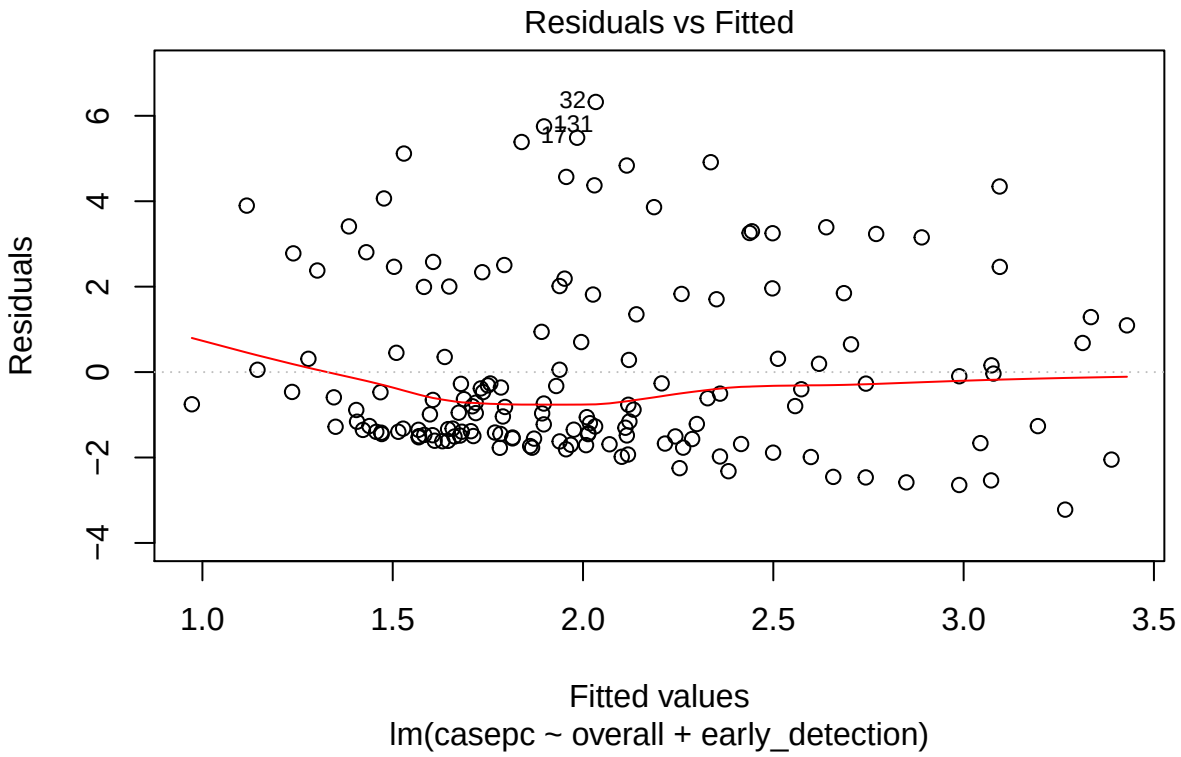

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
3.0
176
83
13
2
3
4
5
Fitted values
lm(cfratio ~ overall + early_detection)

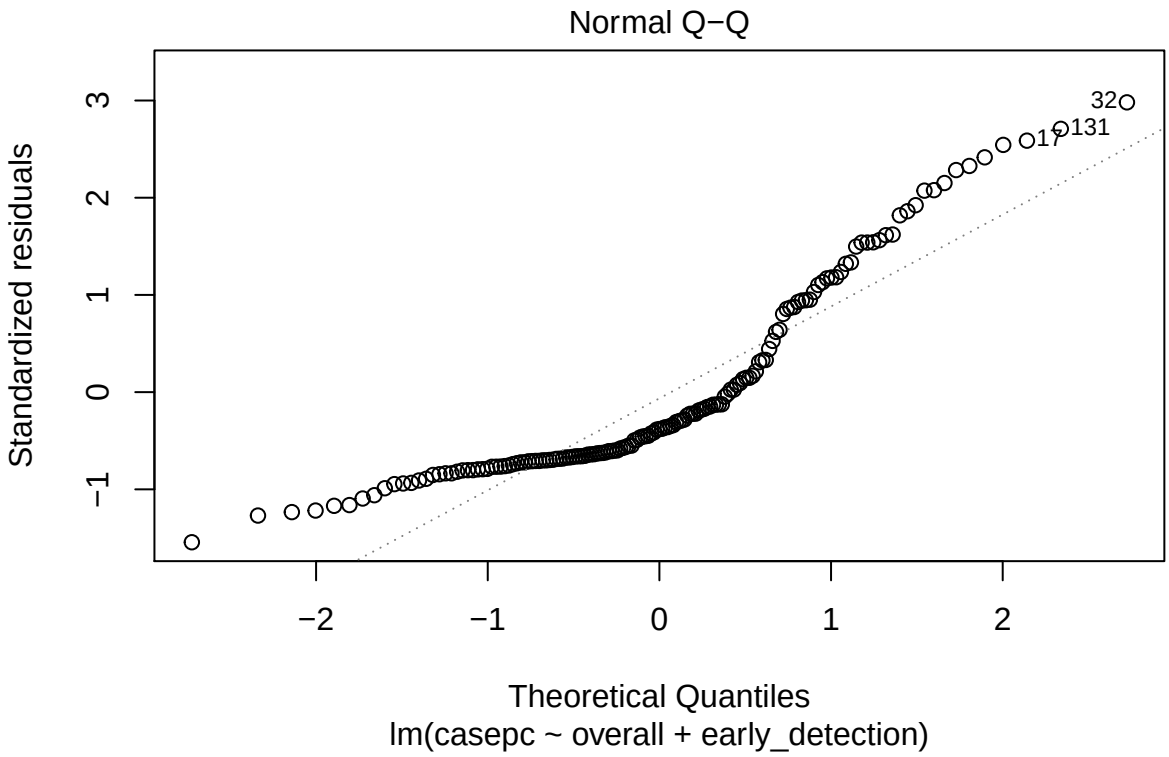


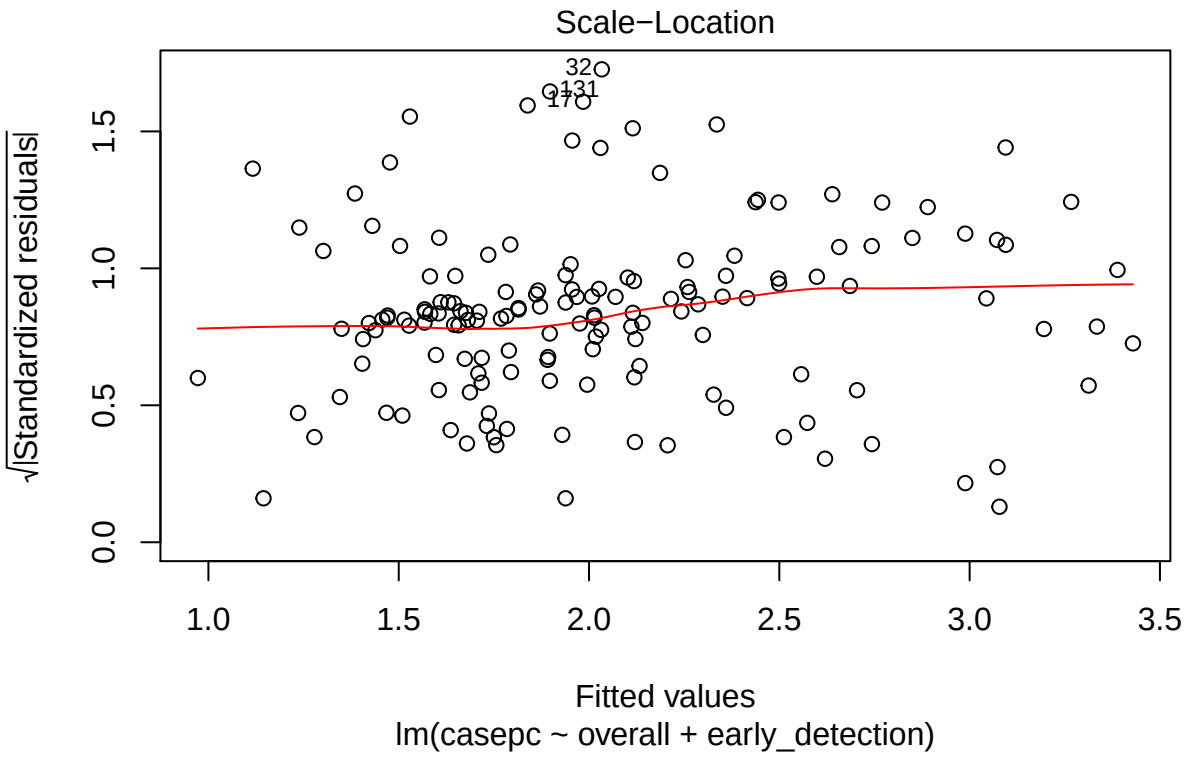

## More strict outlier screening

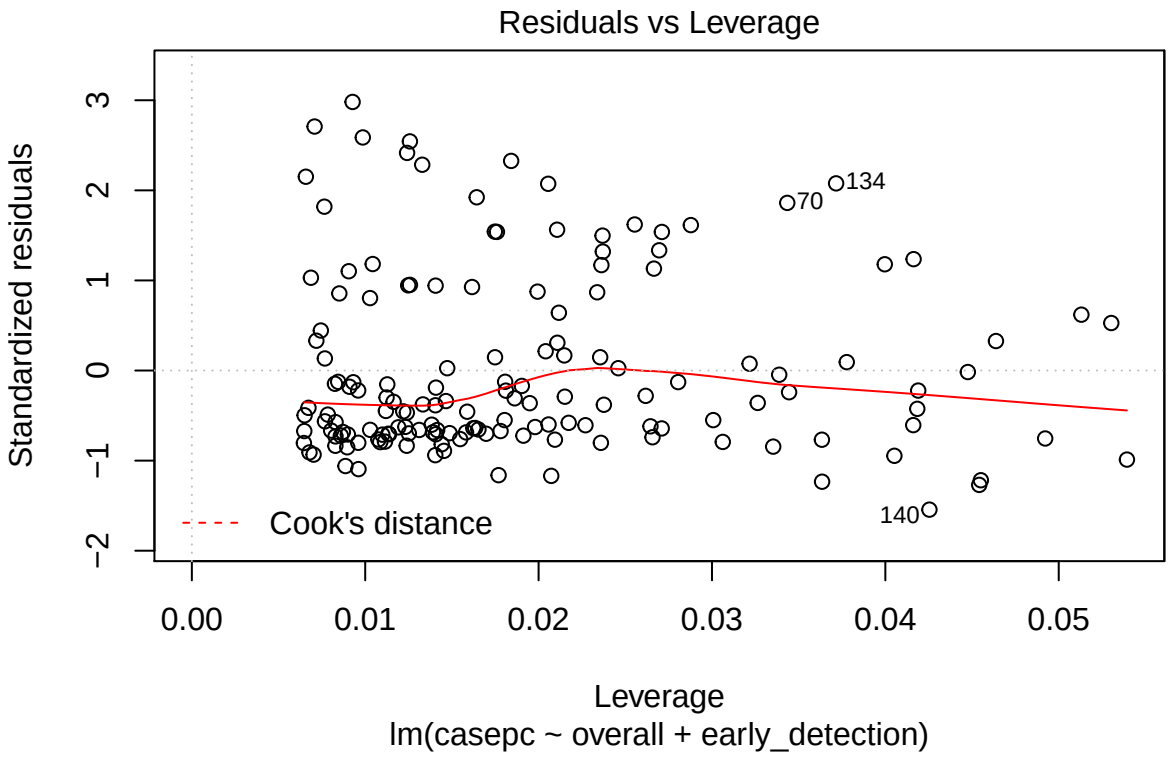
### Cases per-capita

```
##
## Call:
## lm(formula = casepc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2201 -1.4847 -0.8002  1.1902  6.3238
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.38322    0.56018   0.684   0.4950
## overall          0.06234    0.02478   2.515   0.0129 *
## early_detection -0.02074    0.01547  -1.341   0.1821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.131 on 152 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.05903,    Adjusted R-squared:  0.04665
## F-statistic: 4.768 on 2 and 152 DF,  p-value: 0.009812
```


Residuals vs Fitted
Residuals
Fitted values
lm(casepc ~ overall + early_detection)
32
131
17
−4
−2
0
2
4
6
1.0
1.5
2.0
2.5
3.0
3.5

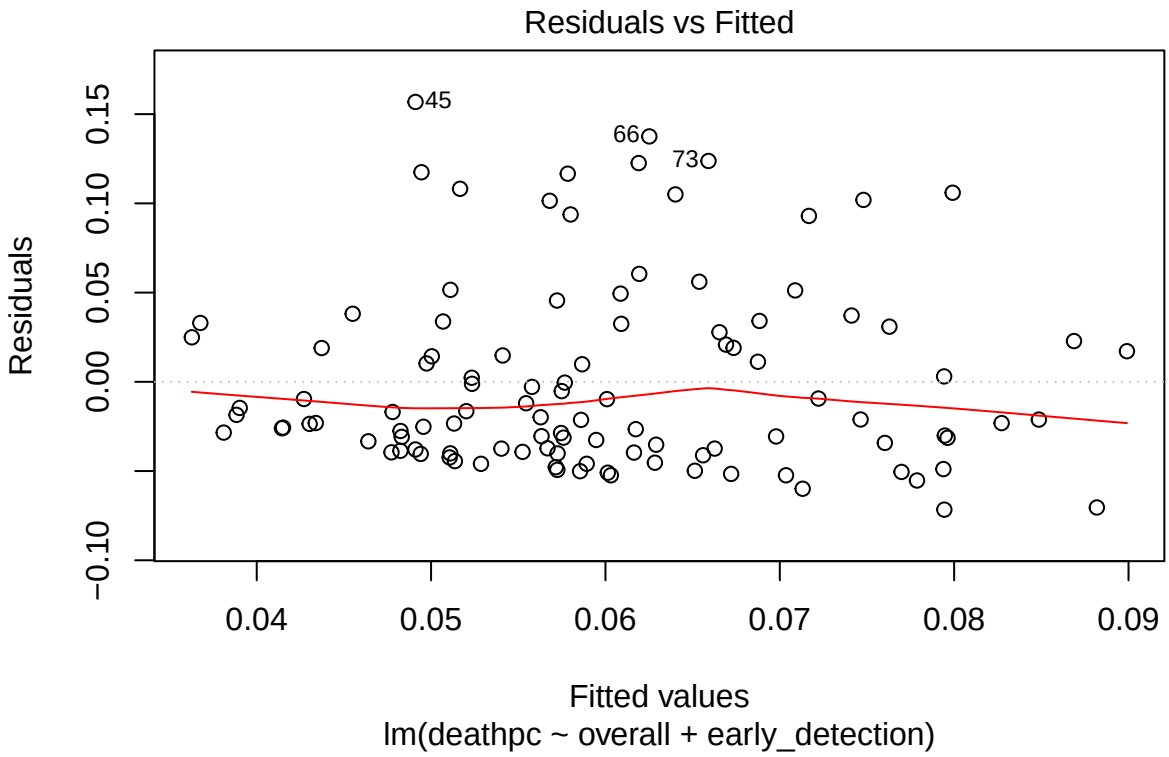

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(casepc ~ overall + early_detection)
32
131
17
−2
−1
0
1
2
−1
0
1
2
3

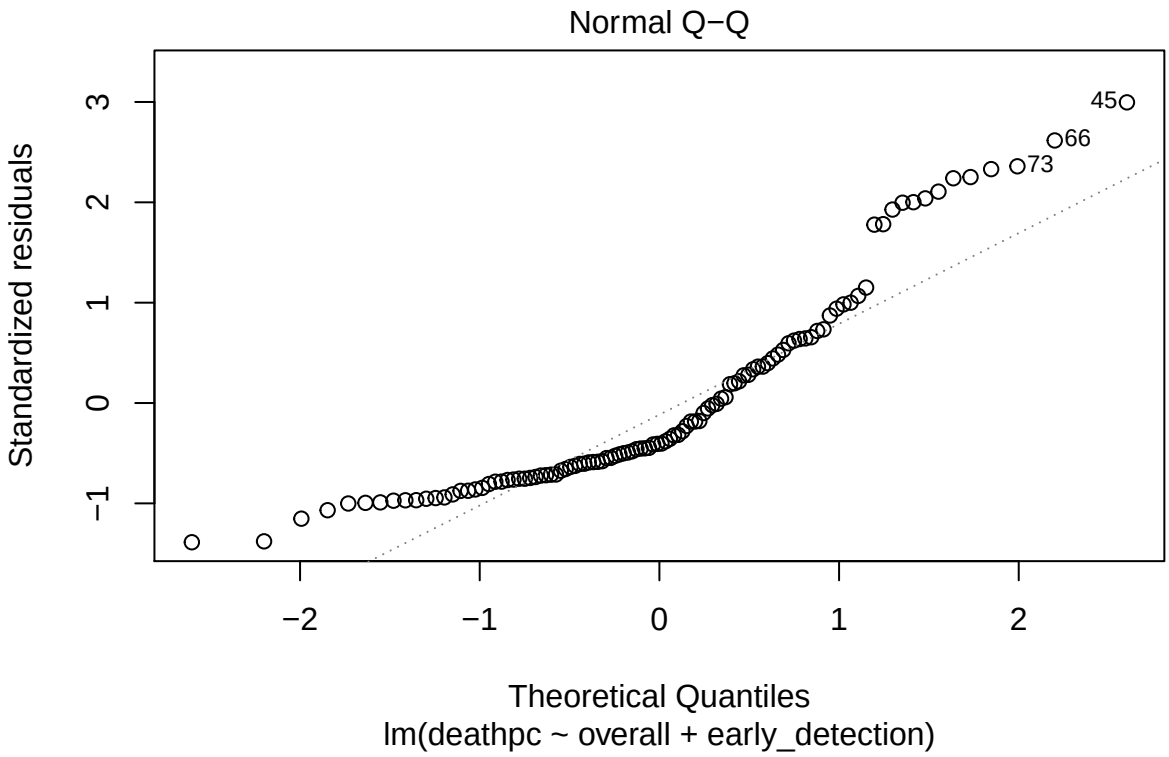

Scale−Location
32
131
17
√|Standardized residuals|
0.0
0.5
1.0
1.5
1.0
1.5
2.0
2.5
3.0
3.5
Fitted values
lm(casepc ~ overall + early_detection)

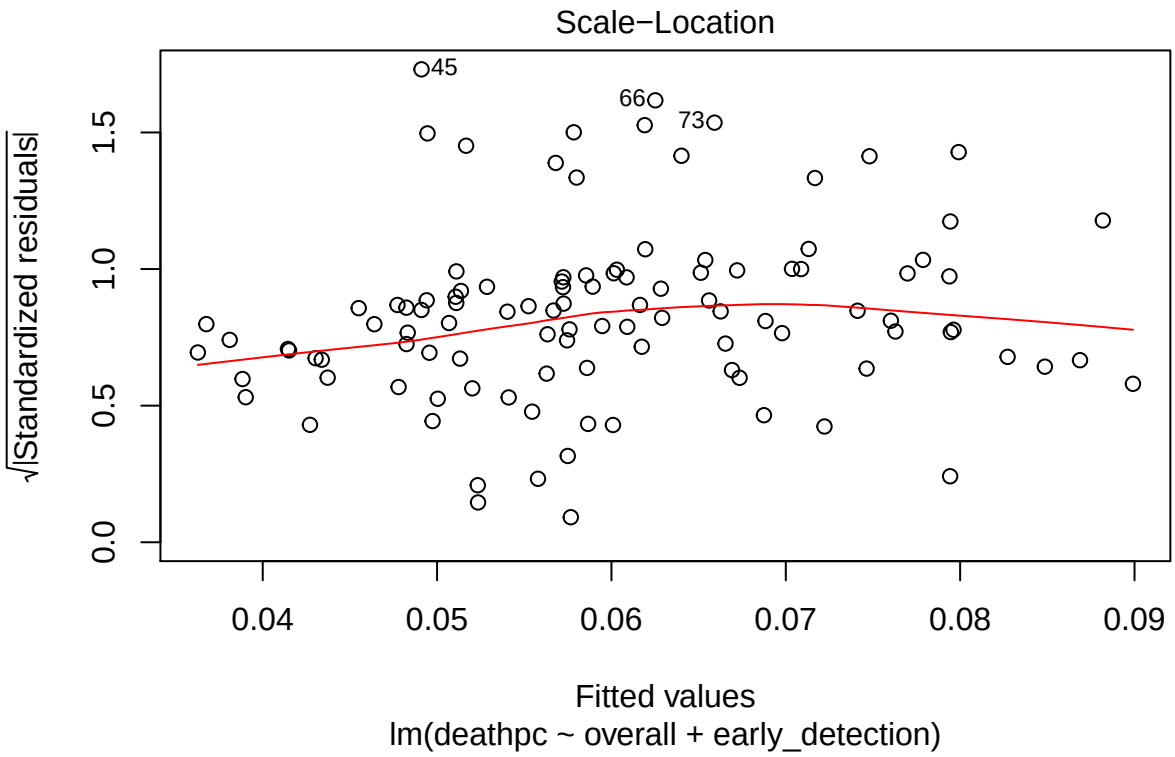


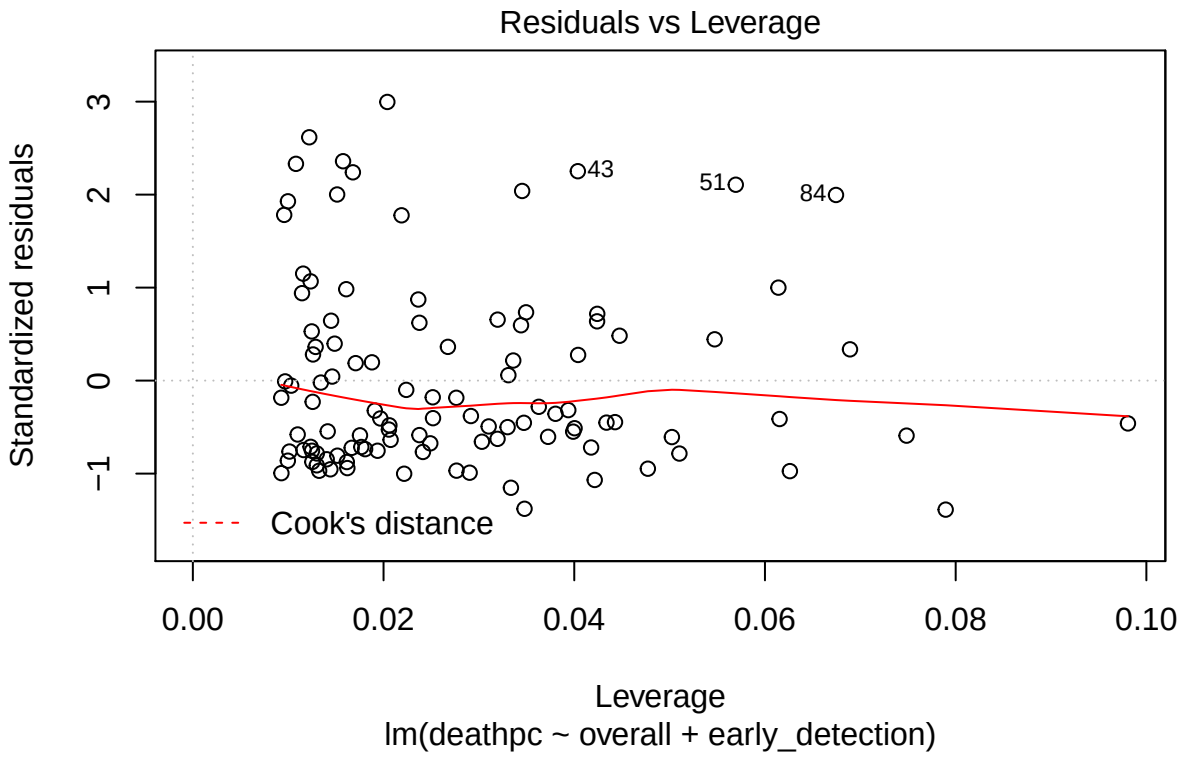

### Deaths per-capita

```
##
## Call:
## lm(formula = deathpc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.07161 -0.03809 -0.02113  0.02565  0.15682
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.0252879  0.0193233   1.309    0.194
## overall         0.0006099  0.0007898   0.772    0.442
## early_detection 0.0002521  0.0004507   0.559    0.577
##
## Residual standard error: 0.05288 on 105 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.05195,    Adjusted R-squared:  0.03389
## F-statistic: 2.877 on 2 and 105 DF,  p-value: 0.06075
```


Residuals vs Fitted
45
66
73
Residuals
−0.10
0.00
0.05
0.10
0.15
0.04
0.05
0.06
0.07
0.08
0.09
Fitted values
lm(deathpc ~ overall + early_detection)

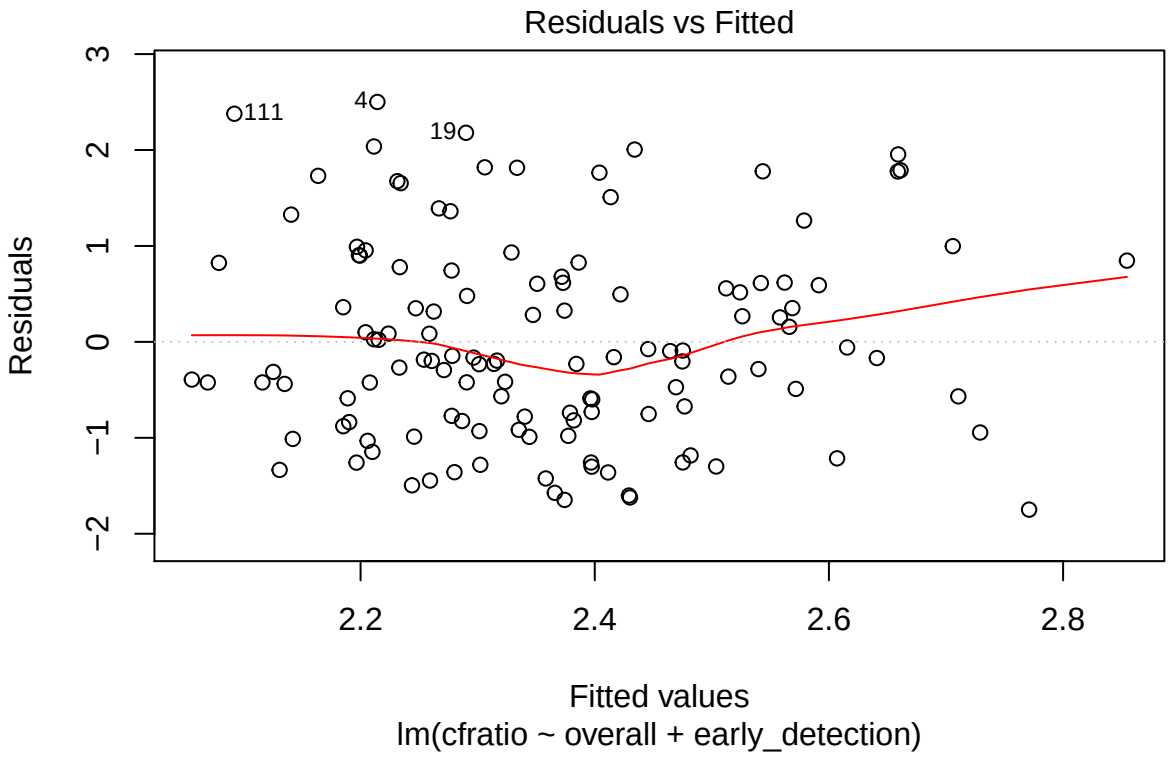

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(deathpc ~ overall + early_detection)
45
66
73
−2
−1
0
1
2
−1
0
1
2
3

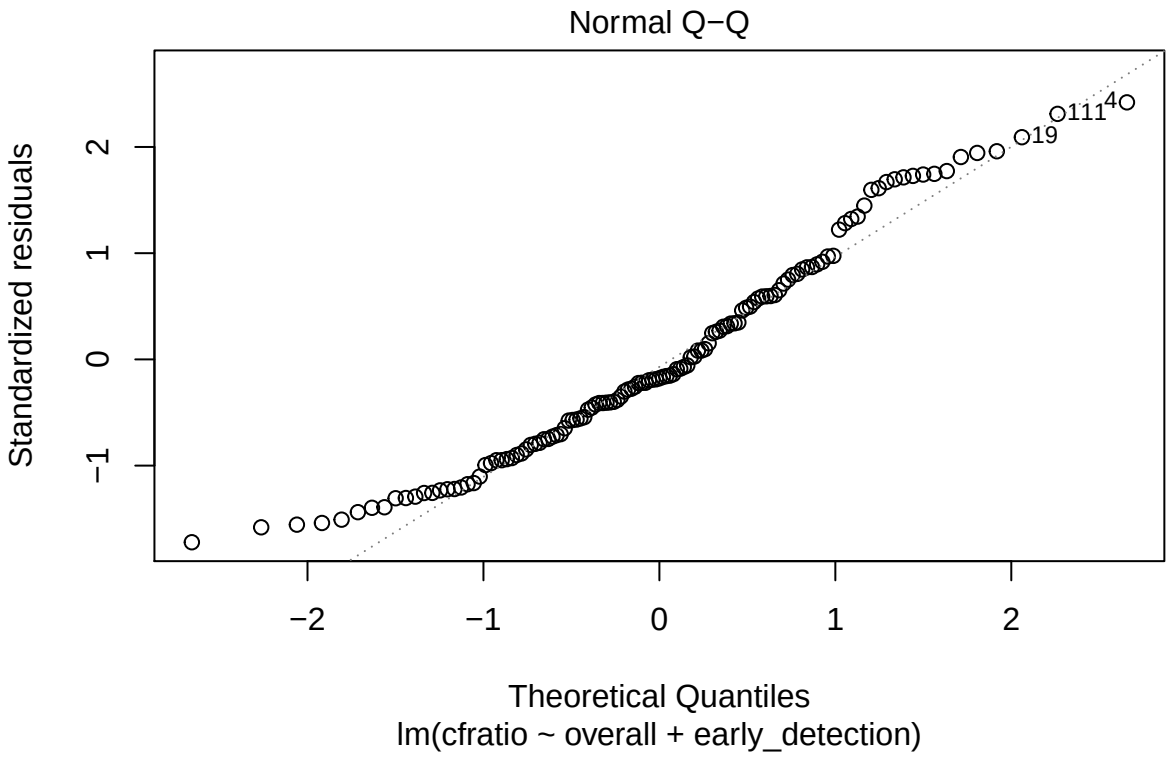

Scale−Location
45
66
73
√|Standardized residuals|
0.0
0.5
1.0
1.5
0.04
0.05
0.06
0.07
0.08
0.09
Fitted values
lm(deathpc ~ overall + early_detection)

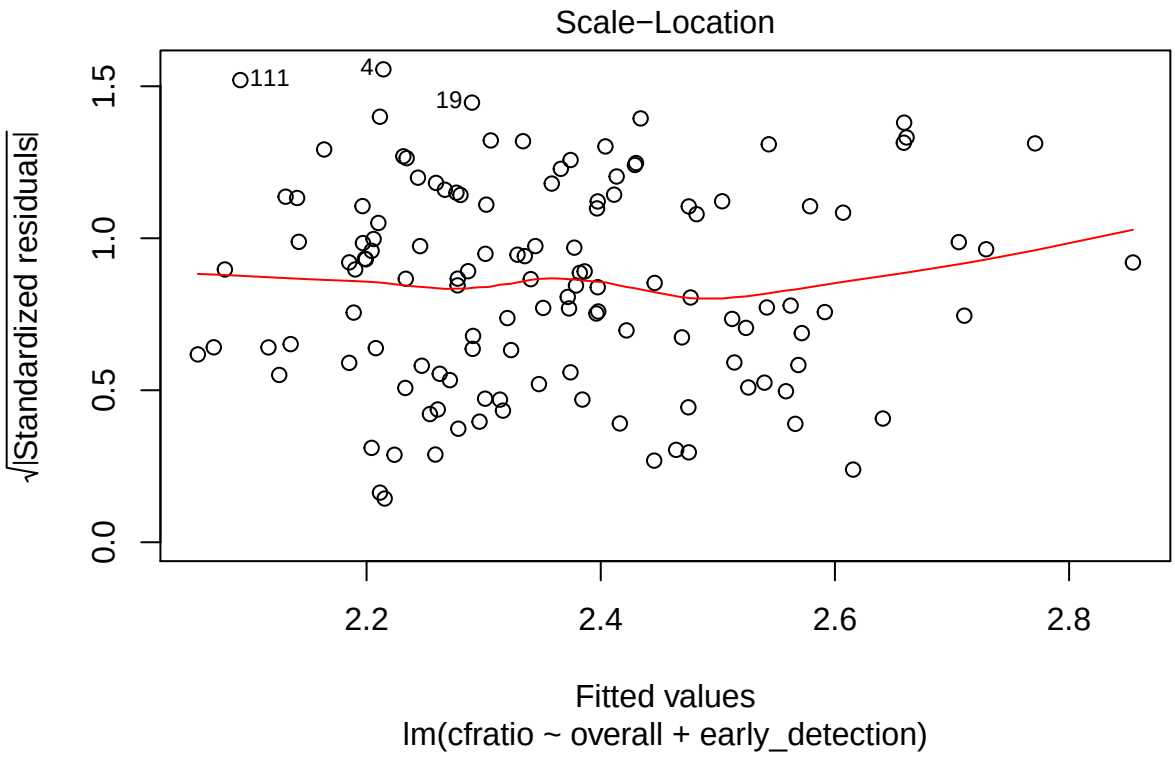


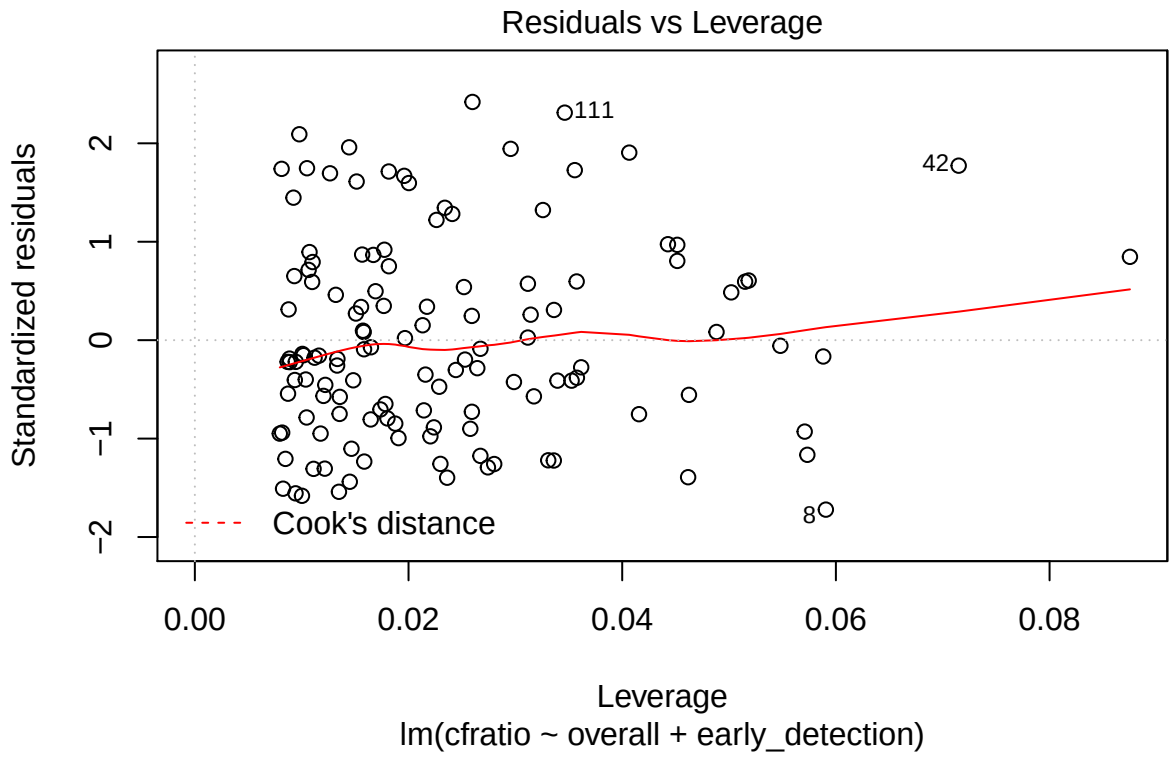

### Case-fatality ratio

```
##
## Call:
## lm(formula = cfratio ~ overall + early_detection, data = regression_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7485 -0.7985 -0.1851  0.6483  2.5010
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     1.883823   0.325962   5.779 5.71e-08 ***
## overall         0.010320   0.013927   0.741    0.460
## early_detection 0.001109   0.008379   0.132    0.895
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.047 on 124 degrees of freedom
## Multiple R-squared:  0.02387,    Adjusted R-squared:  0.008123
## F-statistic: 1.516 on 2 and 124 DF,  p-value: 0.2236
```


Residuals vs Fitted
Residuals
Fitted values
lm(cfratio ~ overall + early_detection)
111
4
19
3
2
1
0
−1
−2
2.2
2.4
2.6
2.8


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(cfratio ~ overall + early_detection)
111
19
4
−2
−1
0
1
2

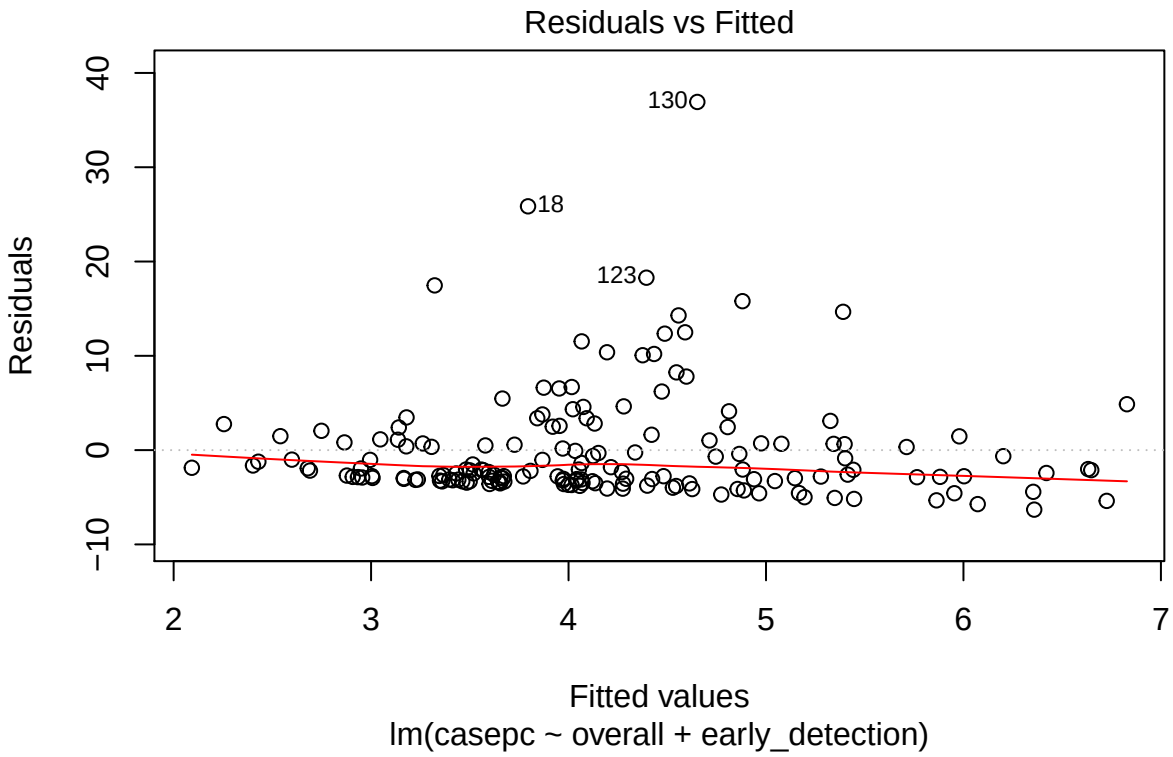

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
111
4
19
2.2
2.4
2.6
2.8
Fitted values
lm(cfratio ~ overall + early_detection)

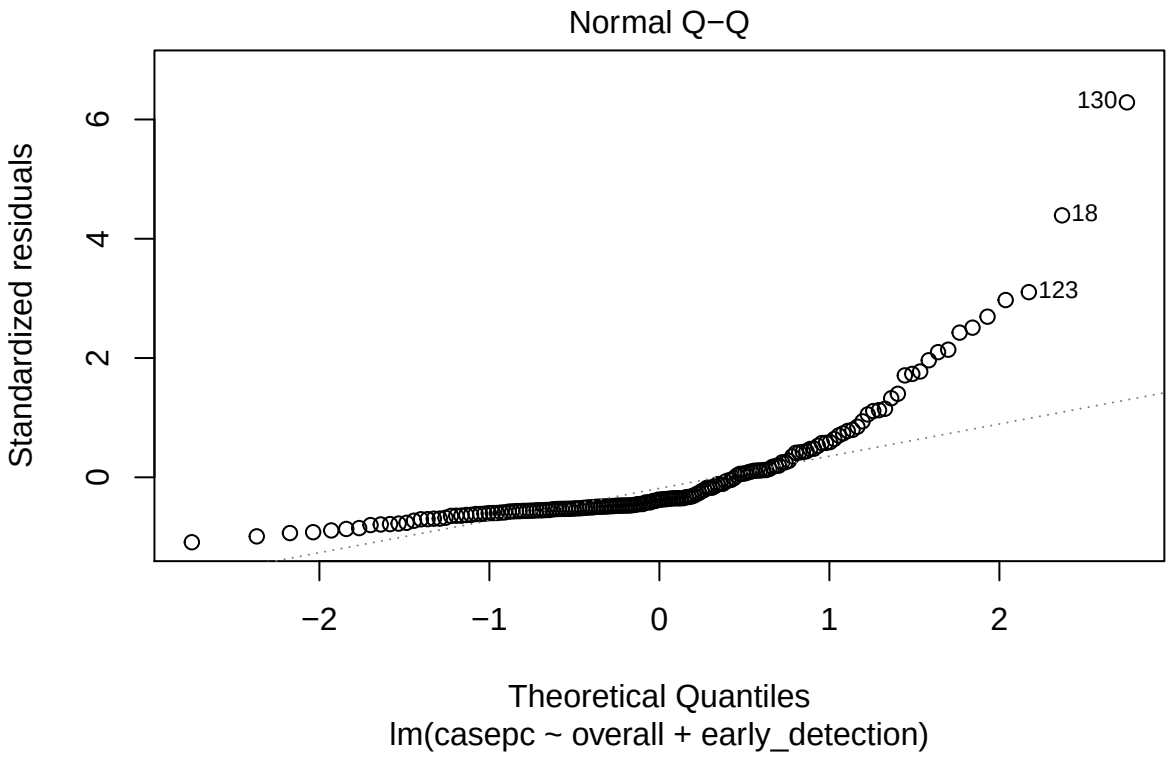


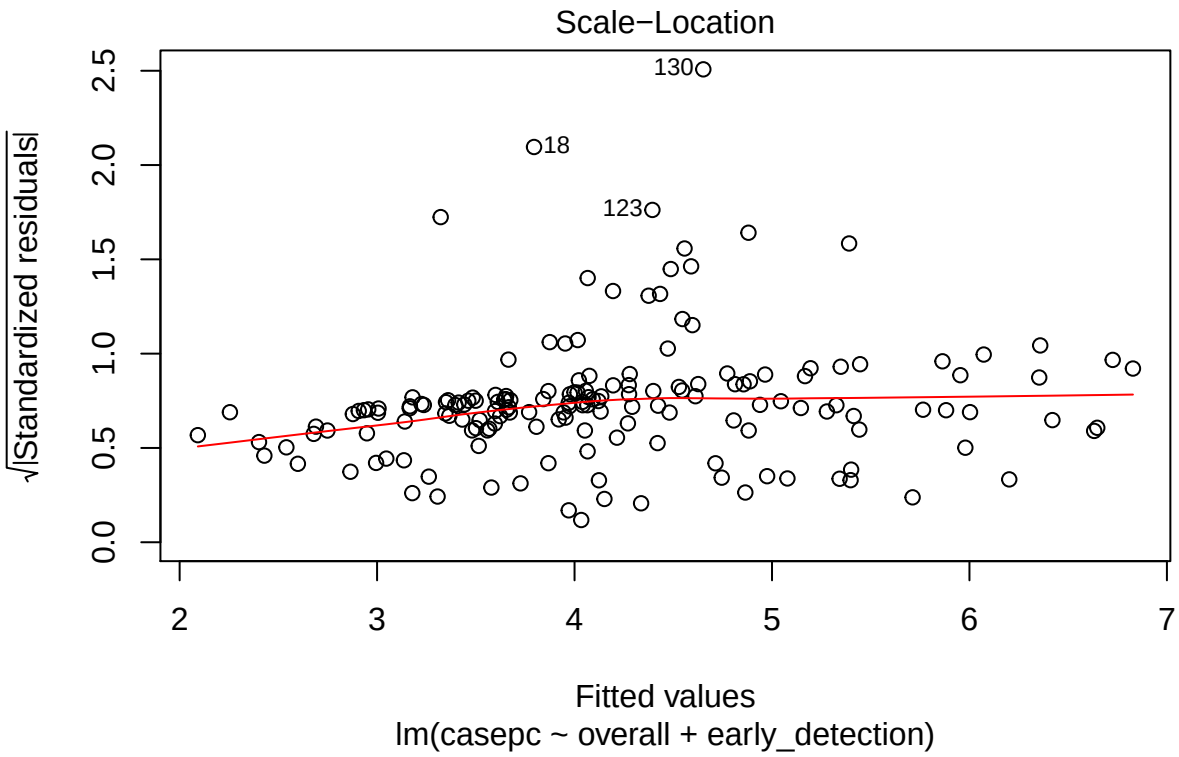

## One more attempt, just dropping any locations where deaths = 0 (N=14)

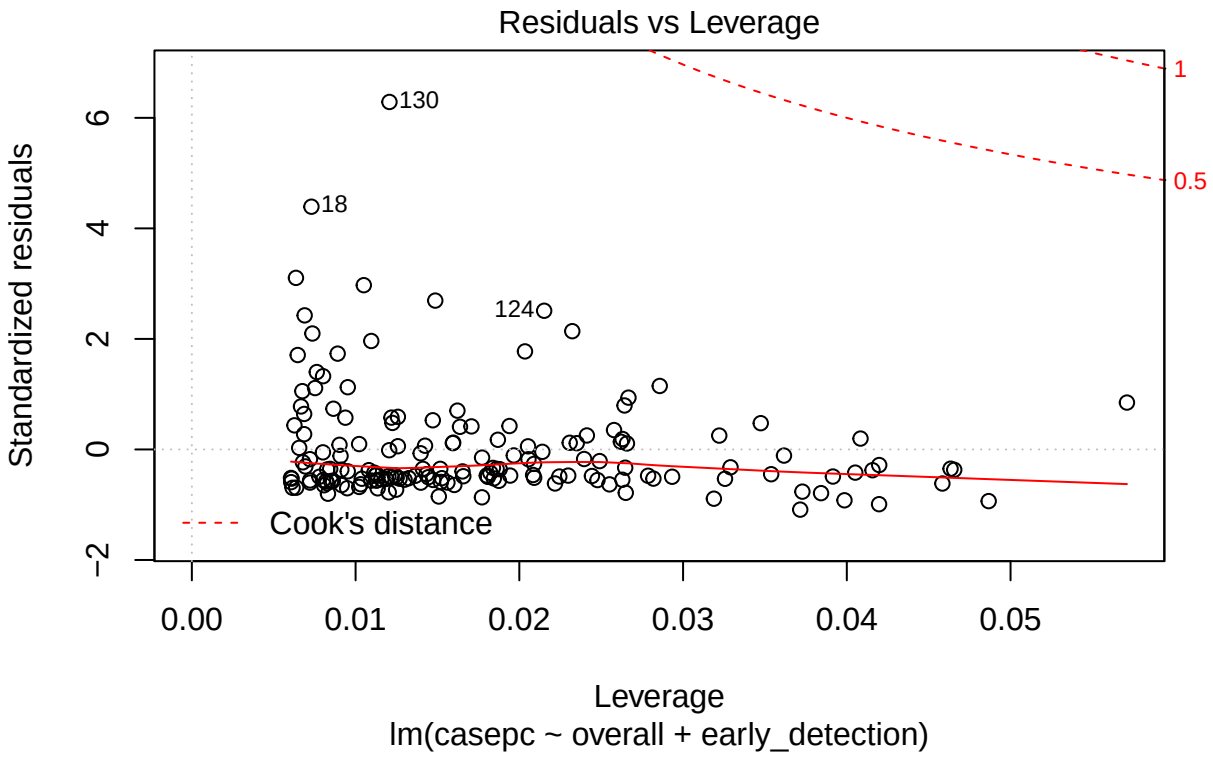
### Cases per capita

```
##
## Call:
## lm(formula = casepc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.312 -3.217 -2.188  1.045 36.922
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.03594    1.54190   0.672    0.503
## overall          0.12598    0.06757   1.864    0.064 .
## early_detection -0.04814    0.04227  -1.139    0.256
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.908 on 165 degrees of freedom
## Multiple R-squared:  0.02727,    Adjusted R-squared:  0.01548
## F-statistic: 2.313 on 2 and 165 DF,  p-value: 0.1021
```


Residuals vs Fitted
Residuals
Fitted values
lm(casepc ~ overall + early_detection)
130
18
123
−10
0
10
20
30
40
2
3
4
5
6
7

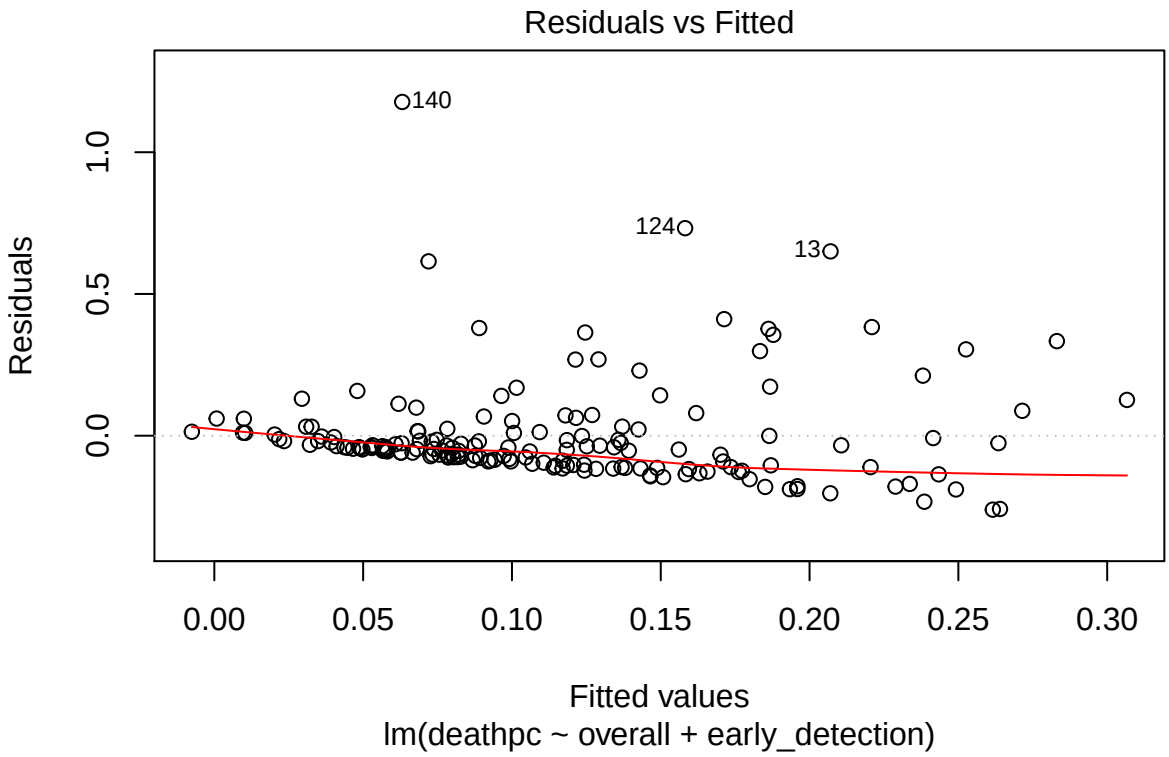

Normal Q−Q
Standardized residuals
0
2
4
6
130
18
123
−2
−1
0
1
2
Theoretical Quantiles
lm(casepc ~ overall + early_detection)

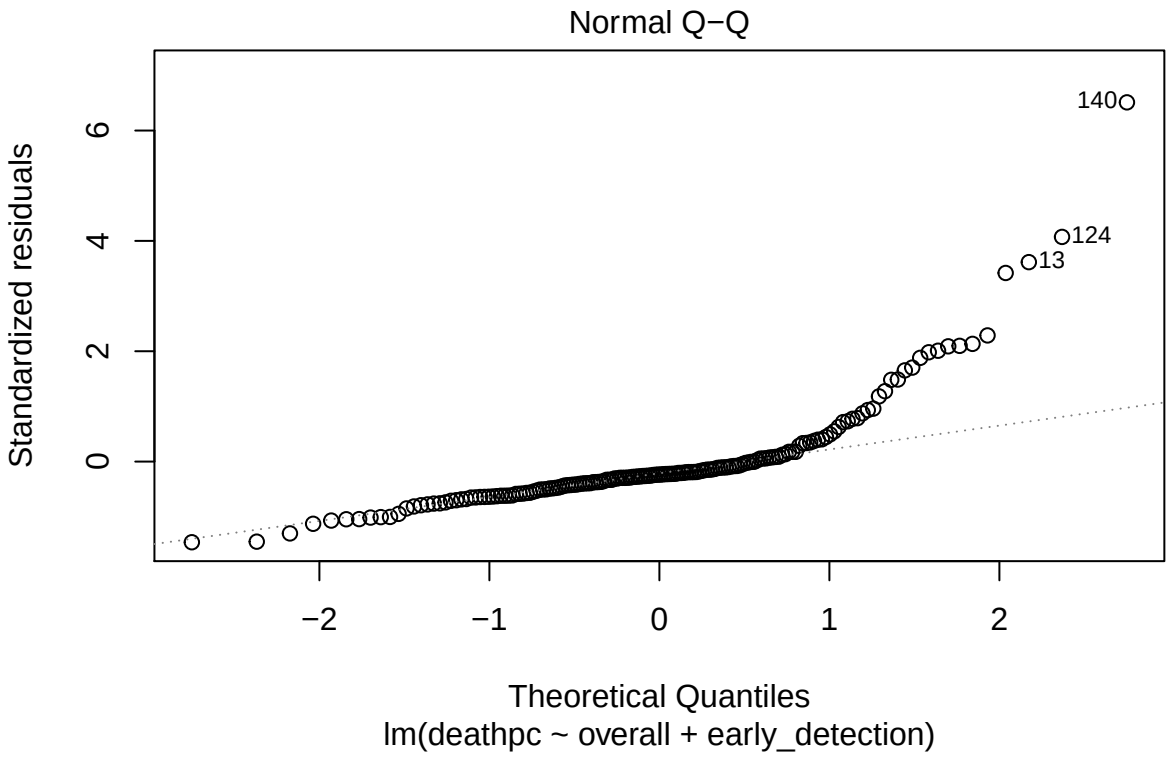

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
130
18
123
2
3
4
5
6
7
Fitted values
lm(casepc ~ overall + early_detection)

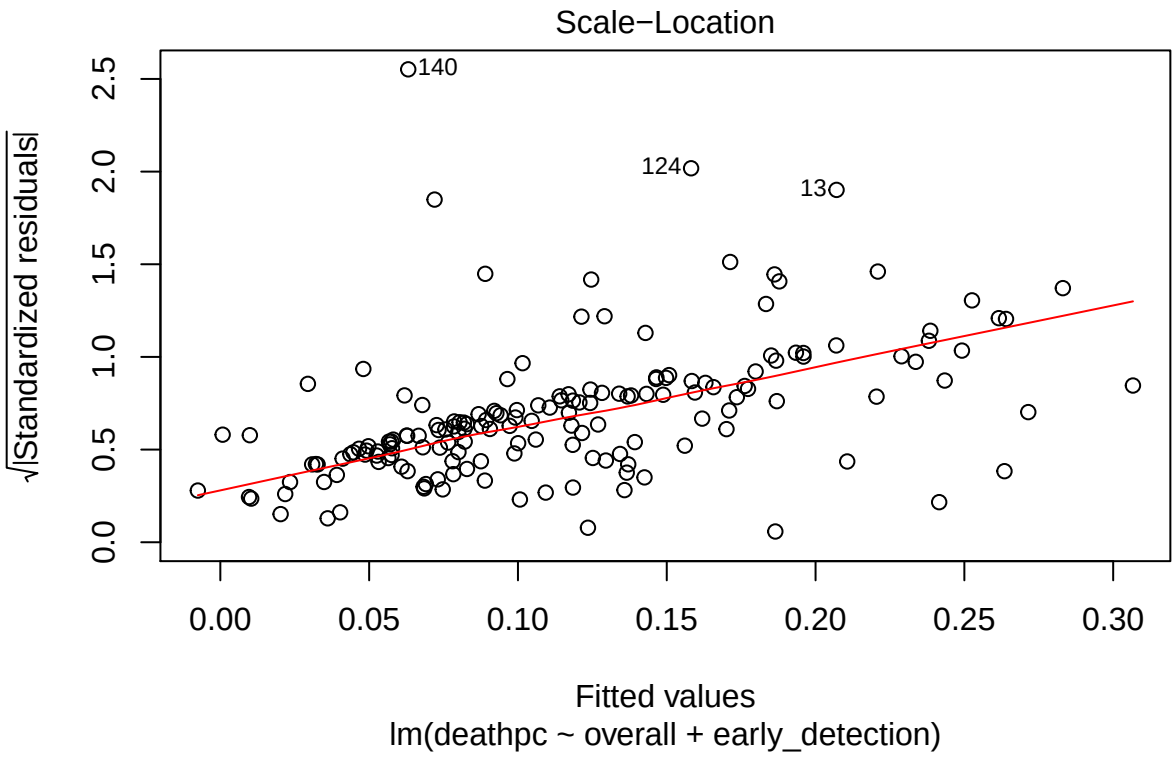


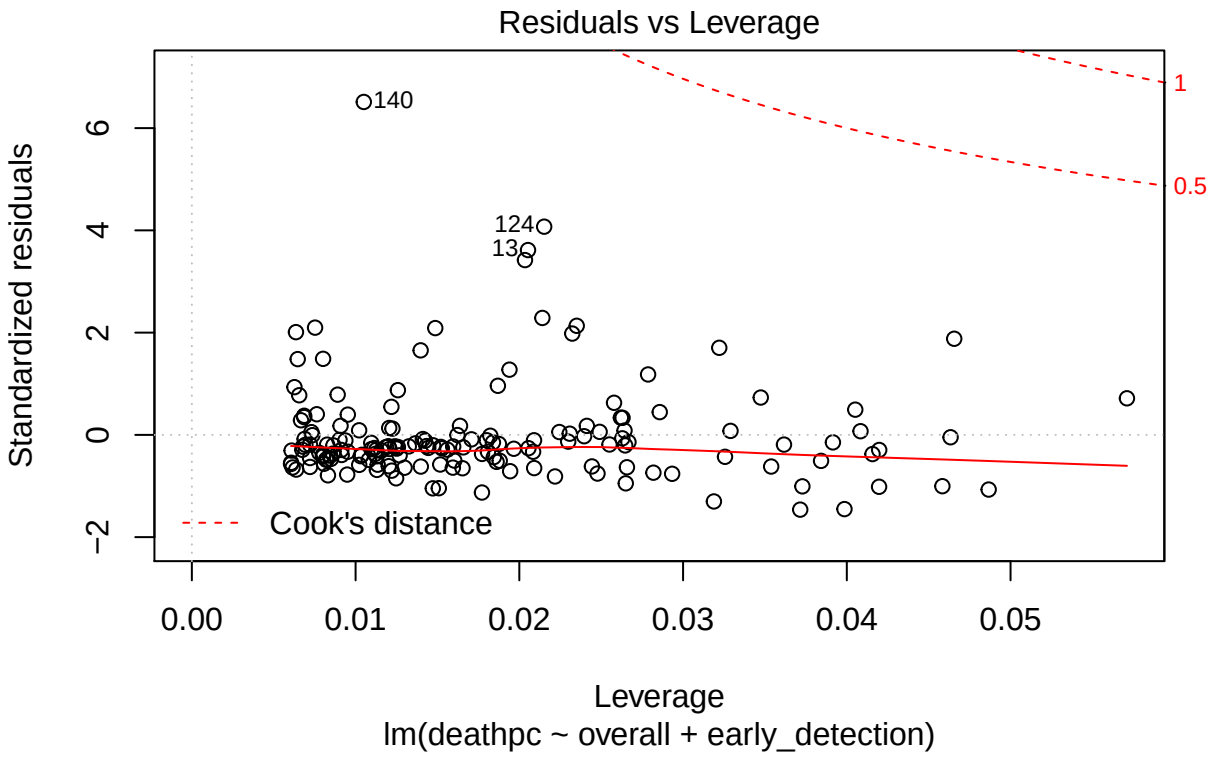

### Deaths per-capita

```
##
## Call:
## lm(formula = deathpc ~ overall + early_detection, data = regression_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26075 -0.09059 -0.04260  0.01421  1.17727
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.0840089  0.0474359  -1.771   0.0784 .
## overall          0.0054003  0.0020789   2.598   0.0102 *
## early_detection -0.0006138  0.0013003  -0.472   0.6376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1818 on 165 degrees of freedom
## Multiple R-squared:  0.116,  Adjusted R-squared:  0.1053
## F-statistic: 10.83 on 2 and 165 DF,  p-value: 3.82e-05
```


Residuals vs Fitted
140
124
13
Residuals
1.0
0.5
0.0
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Fitted values
lm(deathpc ~ overall + early_detection)

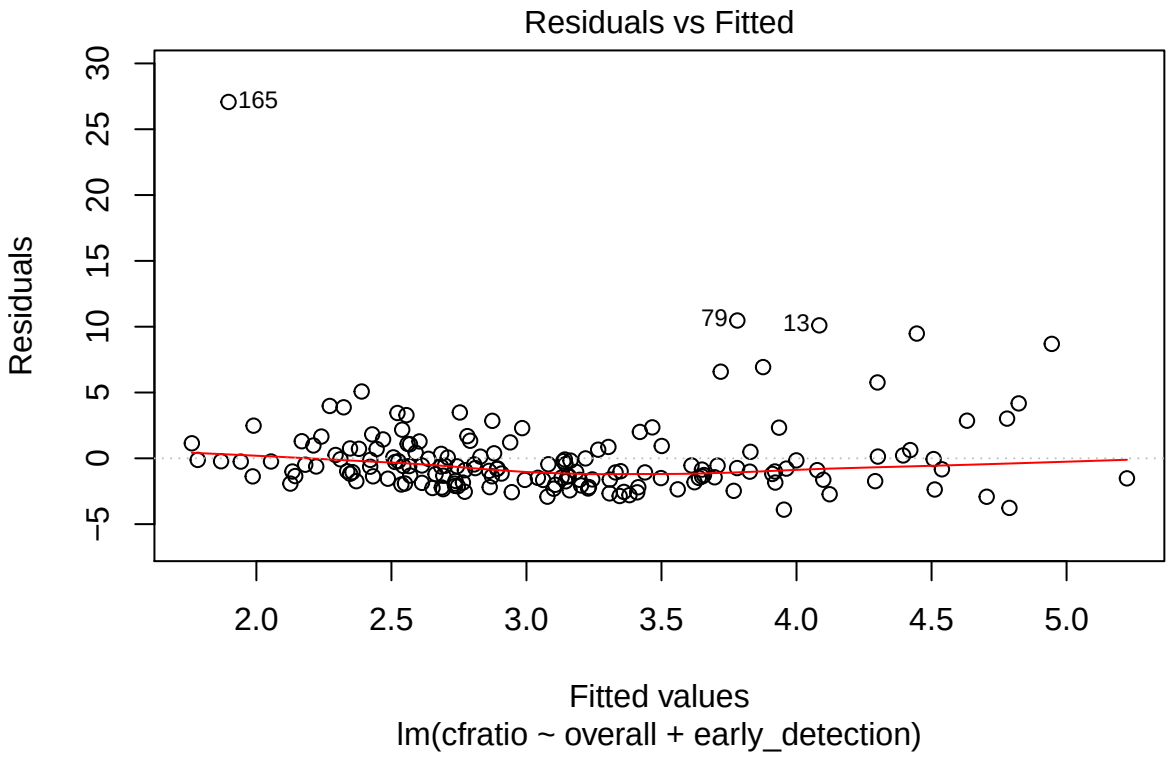

Normal Q−Q
Standardized residuals
0
2
4
6
140
124
13
−2
−1
0
1
2
Theoretical Quantiles
lm(deathpc ~ overall + early_detection)

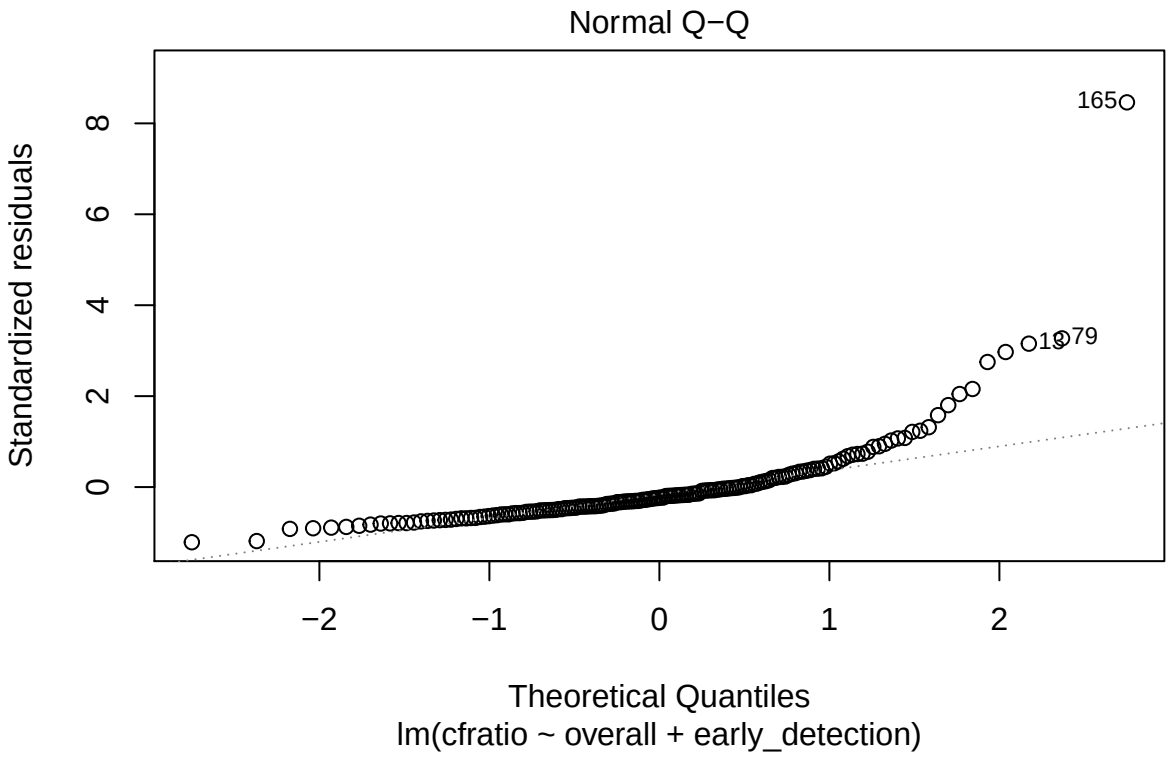

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
140
124
13
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Fitted values
lm(deathpc ~ overall + early_detection)

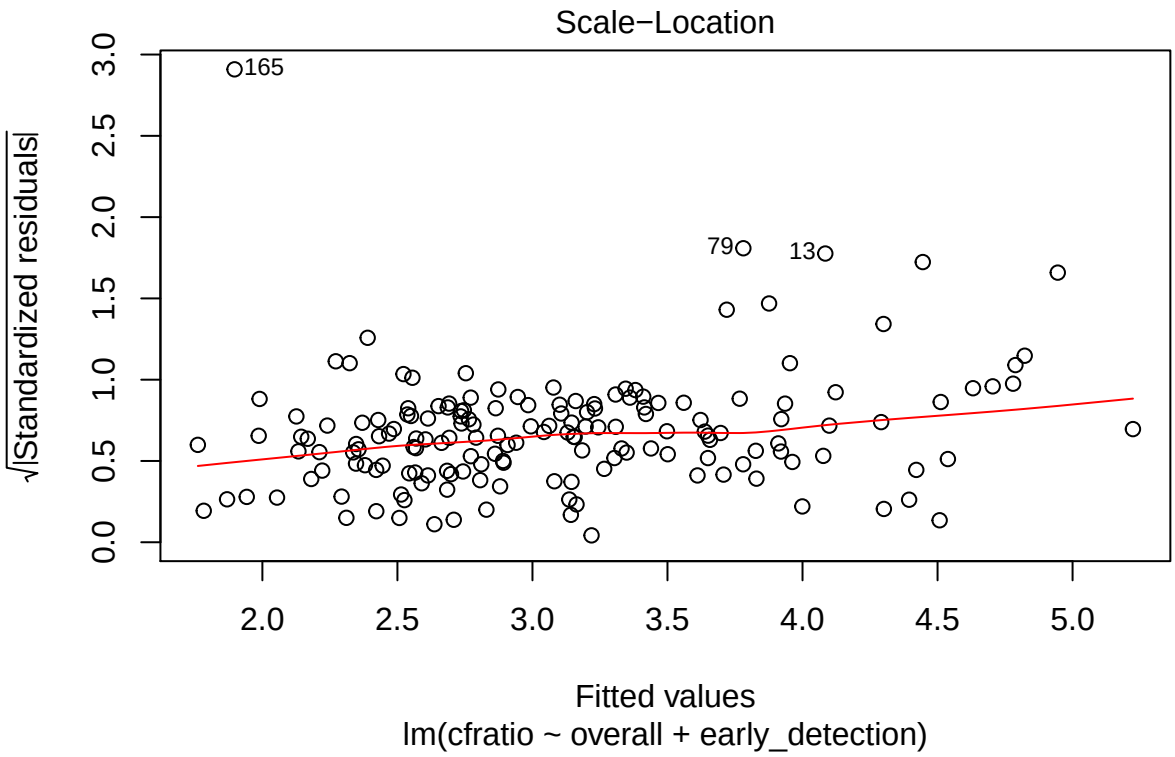


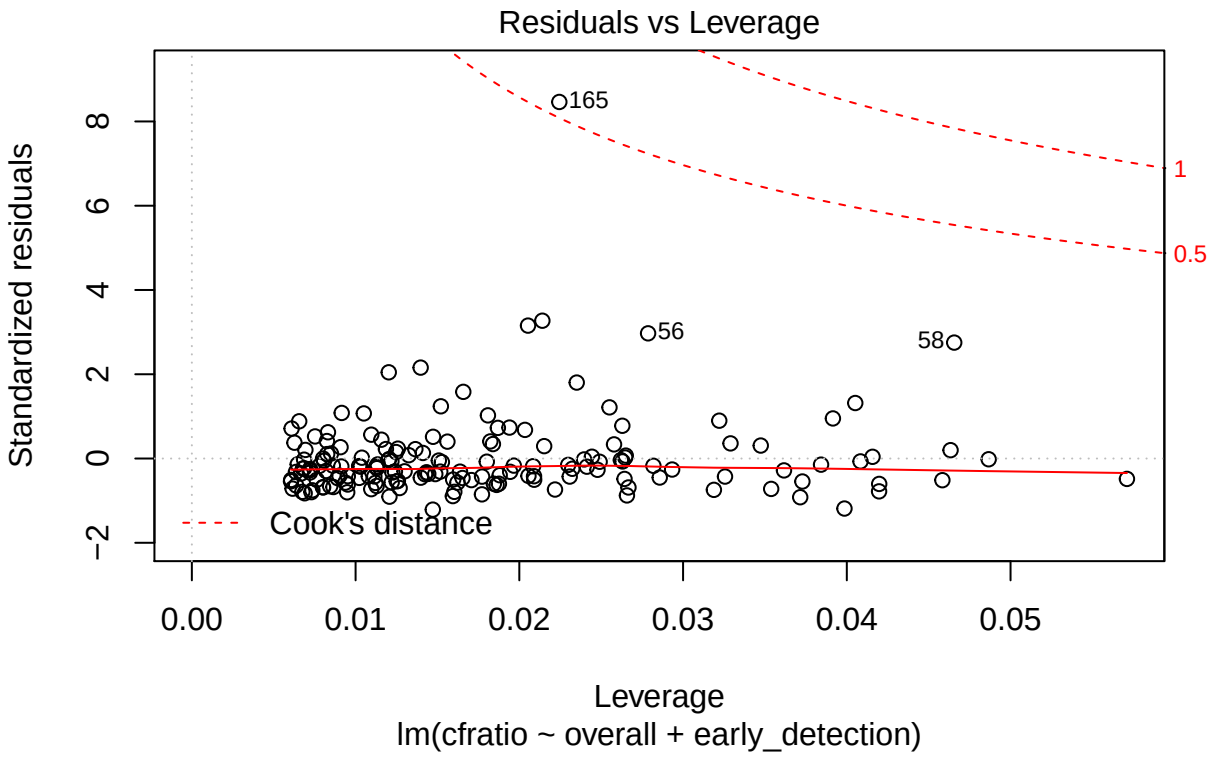

### Case-fatality ratio

```
##
## Call:
## lm(formula = cfratio ~ overall + early_detection, data = regression_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8972 -1.6346 -0.7568  0.6336 27.0723
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.908148   0.844456   1.075    0.284
## overall          0.054690   0.037008   1.478    0.141
## early_detection -0.002558   0.023149  -0.111    0.912
##
## Residual standard error: 3.236 on 165 degrees of freedom
## Multiple R-squared:  0.0495, Adjusted R-squared:  0.03798
## F-statistic: 4.296 on 2 and 165 DF,  p-value: 0.01518
```


Residuals vs Fitted
Residuals
Fitted values
lm(cfratio ~ overall + early_detection)
165
79
13
30
25
20
15
10
5
0
−5
2.0
2.5
3.0
3.5
4.0
4.5
5.0


Normal Q−Q
Standardized residuals
0
2
4
6
8
165
79
−2
−1
0
1
2
Theoretical Quantiles
lm(cfratio ~ overall + early_detection)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
3.0
165
79
13
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Fitted values
lm(cfratio ~ overall + early_detection)


Residuals vs Leverage
Standardized residuals
Leverage
lm(cfratio ~ overall + early_detection)
Cook's distance
1
0.5
165
56
58
−2
0
2
4
6
8
0.00
0.01
0.02
0.03
0.04
0.05