Konfound-It!: An R package and interactive web application to carry out sensitivity analysis

|  |
| --- |
| Joshua M. Rosenberg1, Ran Xu2, & Kenneth A. Frank3 |
| 1 University of Tennessee, Knoxville |
| 2 Virginia Tech |
| 3 Michigan State University |
|  |

# Author note

Correspondence concerning this article should be addressed to Joshua M. Rosenberg, 1122 Volunteer Blvd.. E-mail: [jmrosenberg@utk.edu](mailto:jmrosenberg@utk.edu)

Abstract

Abstract.

*Keywords:* keywords

Word count: X

Konfound-It!: An R package and interactive web application to carry out sensitivity analysis

## Introduction

In social science (and educational) research, we often wish to understand how robust inferences about effects are to unobserved (or controlled for) covariates, possible problems with measurement, and other sources of bias. The goal of konfound is to carry out sensitivity analysis to help analysts to quantify how robust inferences are to potential sources of bias. This package provides functions based on developments in sensitivity analysis by Frank and colleagues, which previously have been implemented in Stata and through an Excel spreadsheet, in R through the konfound package.

# Background on sensitivity analysis

Often times, inferences about effects are critiqued because of *biases*, or factors that contribute to an estimated effect being different from its ground truth value. These biases come from many sources, such as not including possibly important control variables or covariates and sample (i.e., selection bias) and measurement-related (i.e., not measuring the complete nature of a psychological construct) biases.

Because of the prevalence of this problem, a number of strategies to address it have emerged. These strategies can broadly be considered as approach for carrying out sensitivity analysis. These past approaches have done XXX but YYY. Frank and colleauges (2000, 2004, 2007, 2013) have extended this past research by doing ZZZ through two frameworks for approaching sensitivity analysis.

The first approach, termed the replacement of cases approach, uses Rubin’s causal model to interpret how much bias there must be to invalidate an inference in terms of replacing observed cases with counterfactual cases. The second quantifies the robustness of causal inferences in terms of correlations associated with unobserved variables in a regression framework.

In the remainder of this paper, we first describe these two approaches in detail. Then, we introduce the **konfound** package in R, which provides a unified interface to carrying out sensitivity analysis using either approach, using either the results from fitted models or values provided by the user. We show examples from linear and non-linear as well as mixed effects (or multi-level) models. We also discuss how the approach can be used to carry out sensitivity analysis for multiple studies. Finally, we describe a web-based application as an easy interface for those newer to sensitivity analysis that can serve as a gateway to more involved uses.

## Replacement of cases approach to sensitivity analysis

This approach to sensitivity analysis focuses upon quantifying how much of an effect would need to be due to bias to invalidate the inference about it. In particular, Frank et al. (2013) use Rubin’s causal model [cite] to characterize how one could invalidate inferences by replacing observed cases with unobserved cases in which there was no effect. This framework enables researchers to identify the *switch point* (Behn & Vaupel, 1982) whereby the bias is large enough to change one’s belief (one’s inferenece) about an effect.

Technically, the approach compares the estimated effect with a given threshold. The threshold defines the point at which evidence from a study would make one indifferent to the choices. The threshold is commonly defined on the basis of statistical significance but can take other values (i.e., it can be based upon the effect size; Frank et al, 2013). (need some more technical details here; and to provide an example)

## Correlation-based approach to sensitivty analysis

In many non-experimental studies, there is a concern that a variable not considered or included in the analysis is related to both the outcome and predictor of interest. In these cases, not considering the omitted, confounding variable may mean that estimates are biased–they are over-confident. In this approach to sensitivity analysis, the strength of the correlation between a hypothetical omitted variable and the outcome and predictor of interest is quantified.

(need to edit this and provide an example) Specifically, Frank (2000) defined the impact of a confounding variable as rx cvr ycv, where rx cv is the correlation between the unobserved confound and the predictor of interest and r ycv, is the correlation between the unobserved confound and the outcome. Frank (2000) shows how to assess how strong the confounding variable (cv) has to correlate with the predictor (X) as well as the outcome (Y) to invalidate an inference of an effect of X on Y.

# Tutorial

In this section, we provide a tutorial of the use of **konfound** to carry out sensitivity analysis through both the replacement of cases and correlation-based approach.

First, you can install konfound with the following:

You can then load konfound with the library() function:

## Example 1: Use of pkonfound() for values from an already-conducted analysis

pkonfound() is used when we have values from an already-conducted analysis (like a regression analysis), such as one in an already-published study or from an analysis carried out using other software.

In the case of a regression analysis, values from the analysis would simply be used as the inputs to the pkonfound() function. For example, in the use below, we simply enter the values for the estimated effect (an unstandardardized beta coefficient) (2), its standard error (.4), the sample size (100), and the number of covariates (3):

## Percent Bias Necessary to Invalidate the Inference:  
## To invalidate an inference, 60.3% of the estimate would have to be due to bias. This is based on a threshold of 0.794 for statistical significance (alpha = 0.05).  
## To invalidate an inference, 60 observations would have to be replaced with cases for which the effect is 0.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.568 with the outcome and at 0.568 with the predictor of interest (conditioning on observed covariates) to invalidate an inference based on a threshold of 0.201 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.568 X 0.568 = 0.323 to invalidate an inference.

For this set of values, around 60% would need to be false due to a source of bias for the inference to be invalidated (based on statistical significance and a p-value (or alpha) of .05), possible a very robust effect. An omitted, confounding variable (sometimes referred to as a covariate) would need to have an impact (defined as the product of the confounding variable’s correlation with both the predictor of interest and the outcome) of 0.323, presenting a different interpretation of how robust this (hypothetical) effect is to a variable which is important but not included in the analysis.

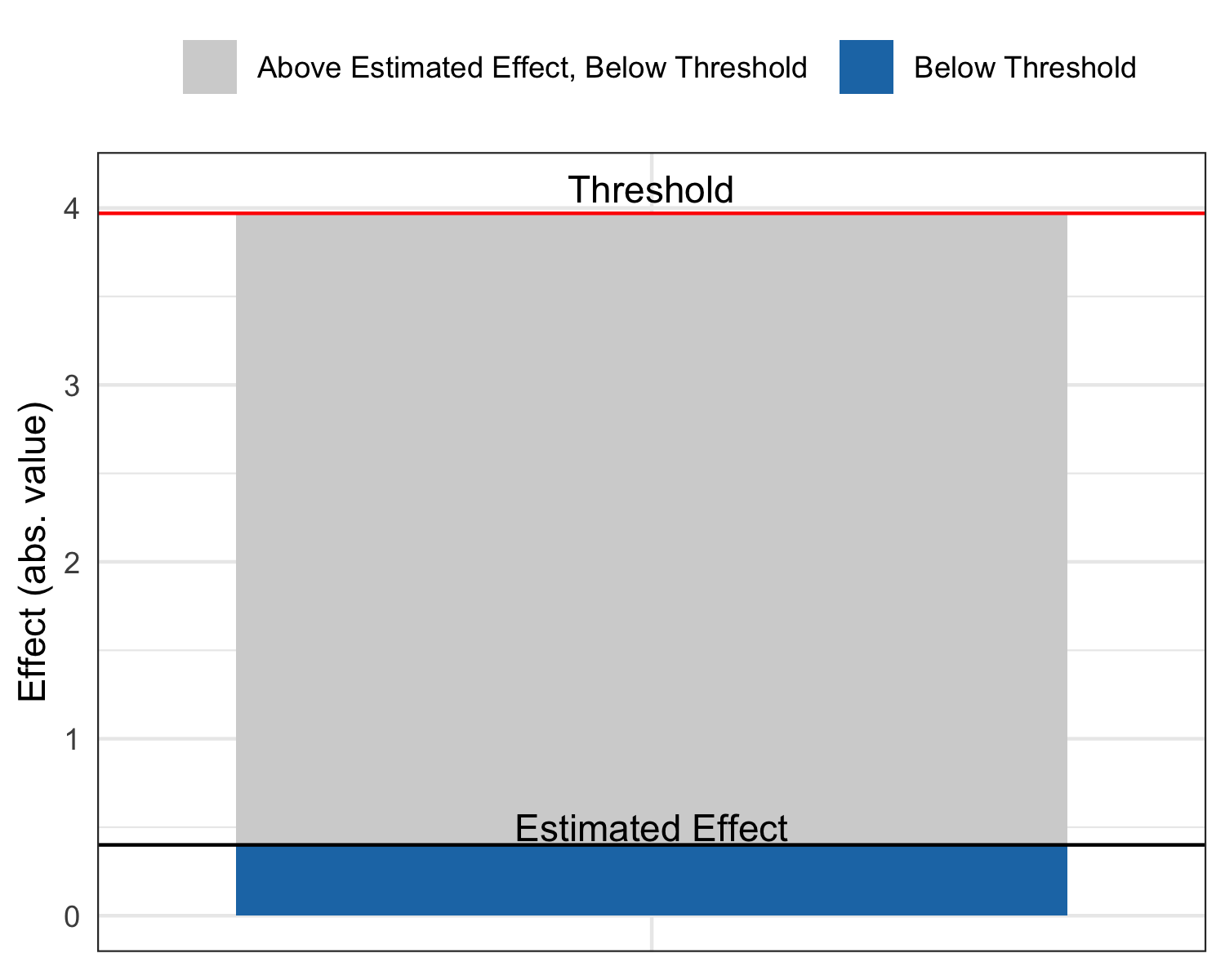
Here is another example, but one in which the unstandardized beta coefficient is smaller than its standard error:

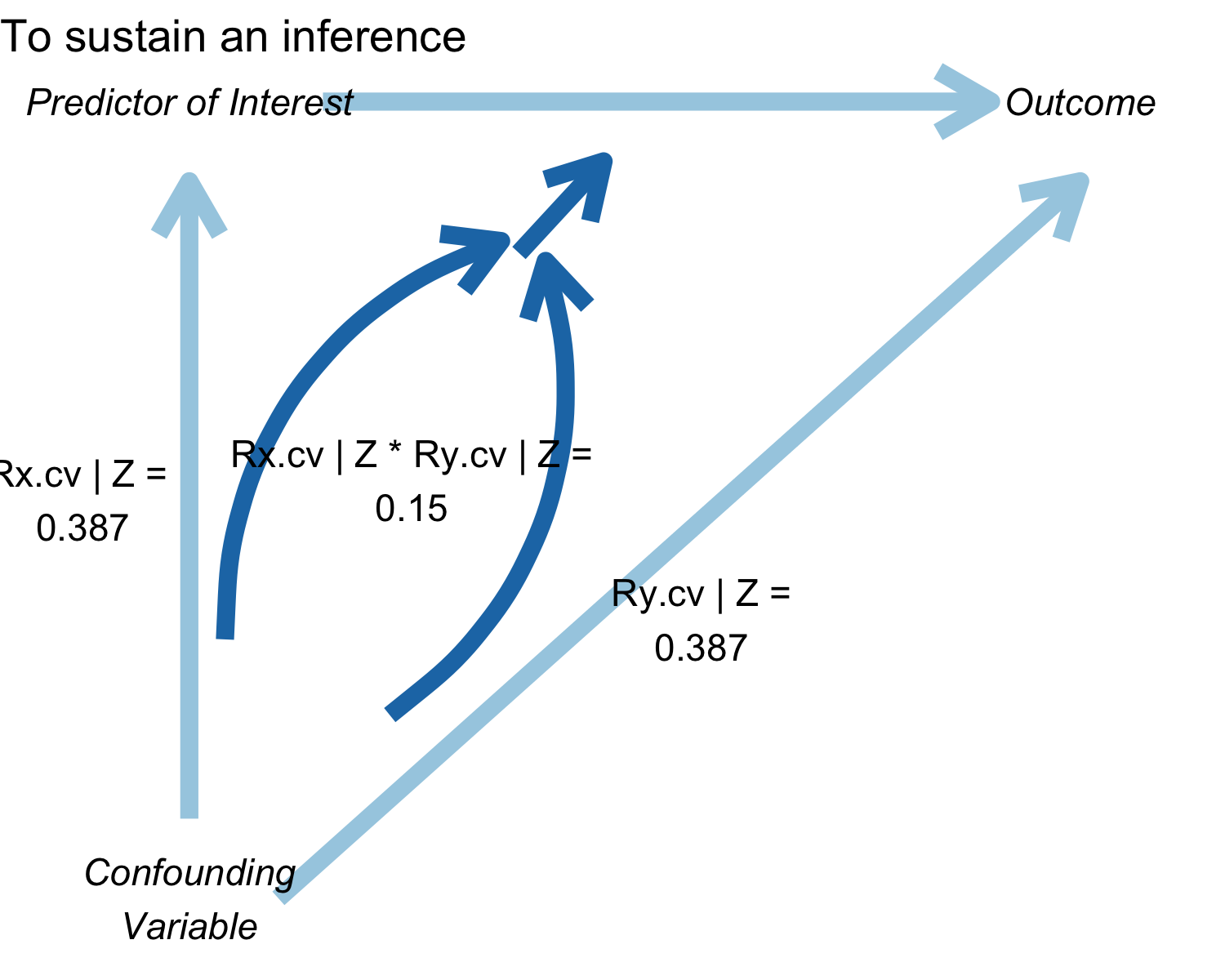
## Percent Bias Necessary to Invalidate the Inference:  
## To sustain an inference, 89.924% of the estimate would have to be due to bias. This is based on a threshold of 3.97 for statistical significance (alpha = 0.05).  
## To sustain an inference, 90 of the cases with 0 effect would have to be replaced with cases at the threshold of inference.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.387 with the outcome and at 0.387 with the predictor of interest (conditioning on observed covariates) to sustain an inference based on a threshold of 3.97 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.387 X 0.387 = 0.15 to sustain an inference.

Note that this use of pkonfound() is equivalent to naming the arguments, i.e. for a different set of values:

## Percent Bias Necessary to Invalidate the Inference:  
## To invalidate an inference, 41.732% of the estimate would have to be due to bias. This is based on a threshold of -1.282 for statistical significance (alpha = 0.05).  
## To invalidate an inference, 83 observations would have to be replaced with cases for which the effect is 0.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.334 with the outcome and at 0.334 with the predictor of interest (conditioning on observed covariates) to invalidate an inference based on a threshold of -0.14 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.334 X 0.334 = 0.112 to invalidate an inference.

We notice that the output includes a message that says we can view other forms of output by changing the to\_return argument. Here are the two plots - for the bias necessary to alter an inference (thresh\_plot) and for the robustness of an inference in terms of the impact of a confounding variable (corr\_plot) that can be returned:





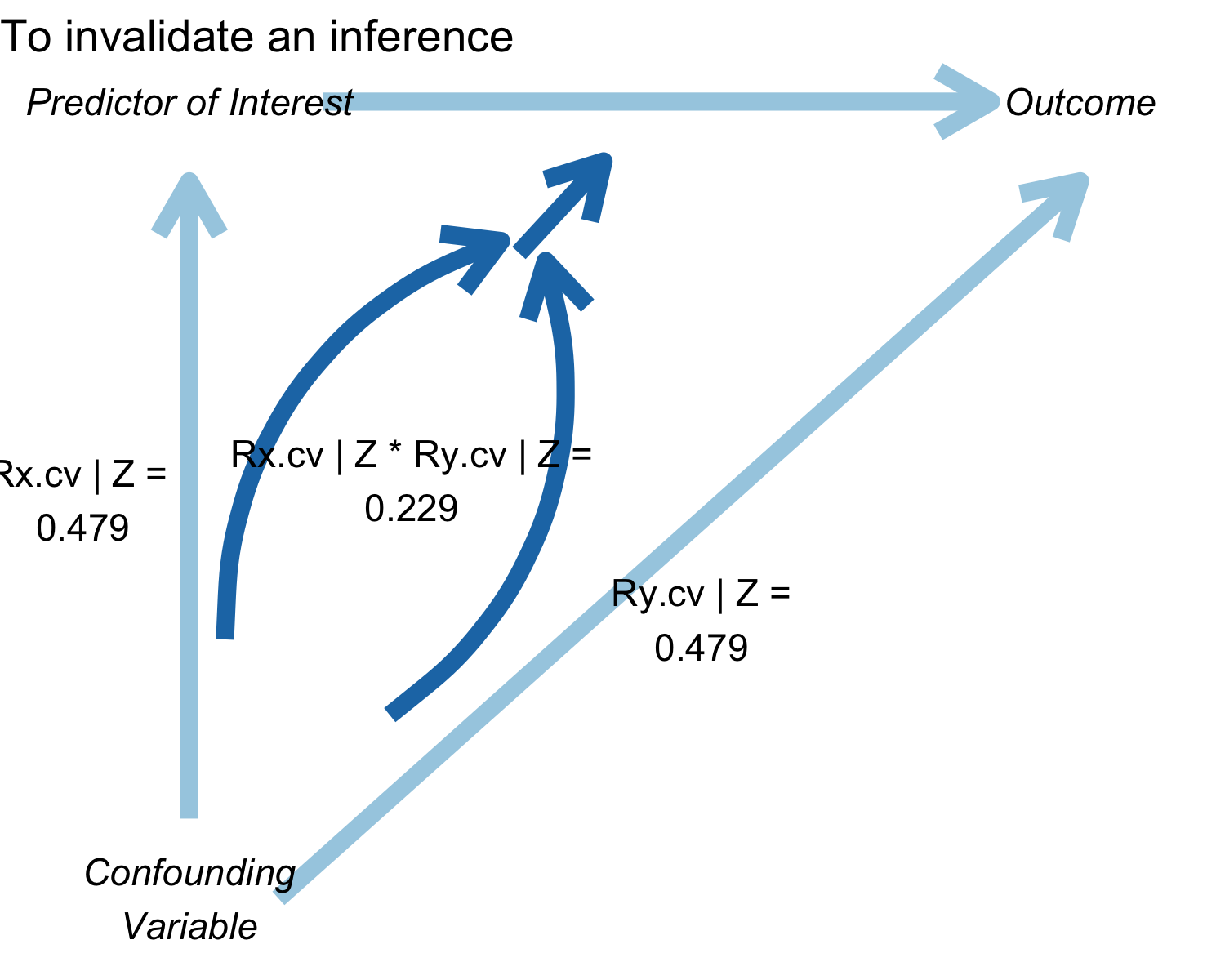
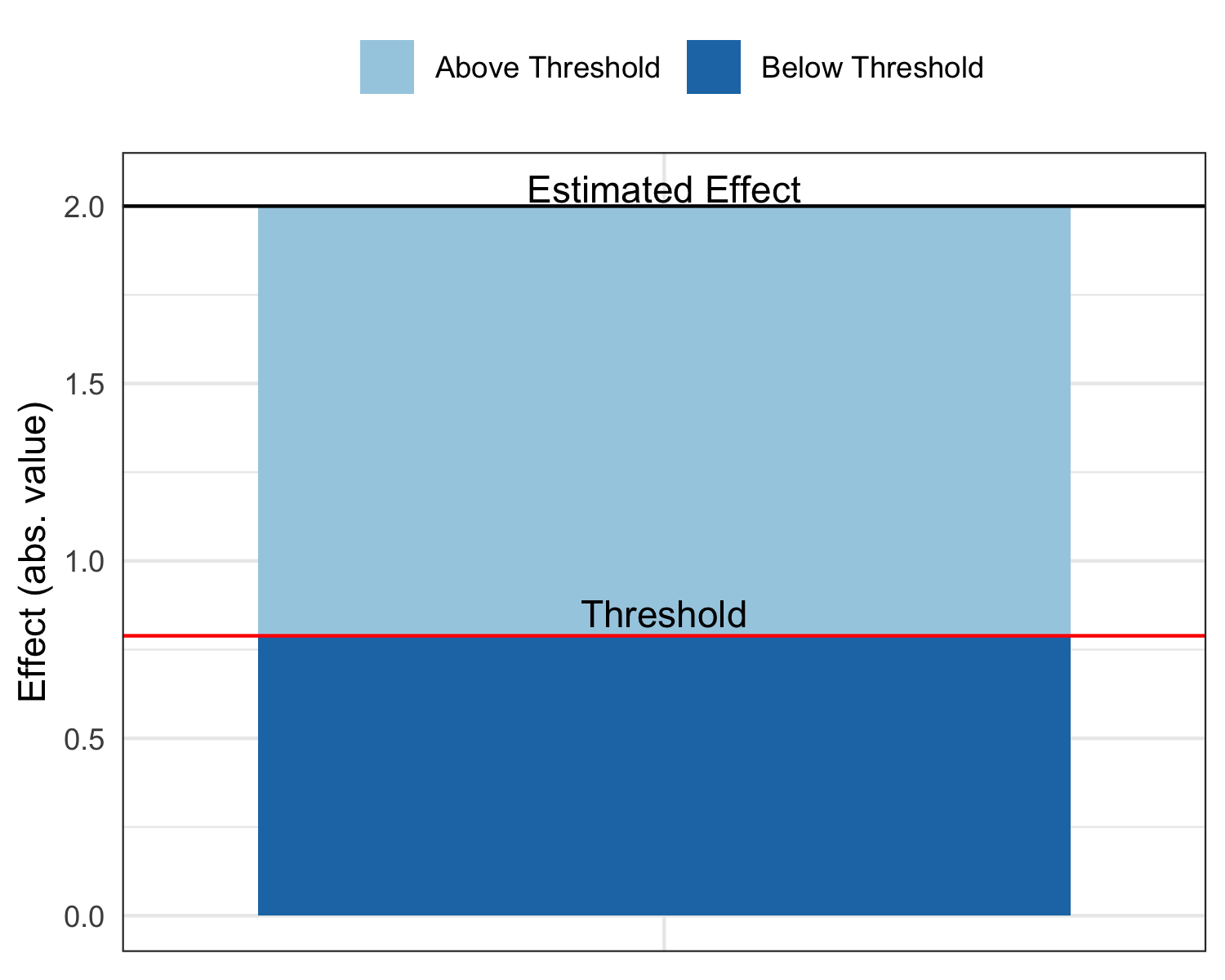
You can also specify multiple forms of output at once.

## Percent Bias Necessary to Invalidate the Inference:  
## To invalidate an inference, 60.557% of the estimate would have to be due to bias. This is based on a threshold of 0.789 for statistical significance (alpha = 0.05).  
## To invalidate an inference, 121 observations would have to be replaced with cases for which the effect is 0.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.479 with the outcome and at 0.479 with the predictor of interest (conditioning on observed covariates) to invalidate an inference based on a threshold of 0.14 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.479 X 0.479 = 0.229 to invalidate an inference.

## Created 3 forms of output. To access type:   
##   
## model\_output$raw\_output  
## model\_output$thresh\_plot  
## model\_output$corr\_plot

When we type the name of the object, we see that we created three types of output that we can access as follows:

## # A tibble: 1 x 8  
## action inference percent\_bias\_to… replace\_null\_ca… unstd\_beta  
## <chr> <chr> <dbl> <dbl> <dbl>  
## 1 to\_in… reject\_n… 60.6 121 2  
## # … with 3 more variables: beta\_threshhold <dbl>,  
## # omitted\_variable\_corr <dbl>, itcv <dbl>



Finally, you can return the raw output, for use in other analyses.

## # A tibble: 1 x 8  
## action inference percent\_bias\_to… replace\_null\_ca… unstd\_beta  
## <chr> <chr> <dbl> <dbl> <dbl>  
## 1 to\_su… fail\_to\_… 89.9 90 0.4  
## # … with 3 more variables: beta\_threshhold <dbl>,  
## # omitted\_variable\_corr <dbl>, itcv <dbl>

## Use of konfound() for models fit in R

Where pkonfound() can be used with values from already-conducted analyses, konfound() can be used with models (lm(), glm(), and lme4::lmer()) fit in R.

### Example 2A: For linear models fit with lm()

##   
## Call:  
## lm(formula = mpg ~ wt + hp + qsec, data = mtcars)  
##   
## Coefficients:  
## (Intercept) wt hp qsec   
## 27.61053 -4.35880 -0.01782 0.51083

## Percent Bias Necessary to Invalidate the Inference:  
## To sustain an inference, 41.327% of the estimate would have to be due to bias. This is based on a threshold of -0.031 for statistical significance (alpha = 0.05).  
## To sustain an inference, 13 of the cases with 0 effect would have to be replaced with cases at the threshold of inference.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.322 with the outcome and at 0.322 with the predictor of interest (conditioning on observed covariates) to sustain an inference based on a threshold of -0.031 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.322 X 0.322 = 0.104 to sustain an inference.

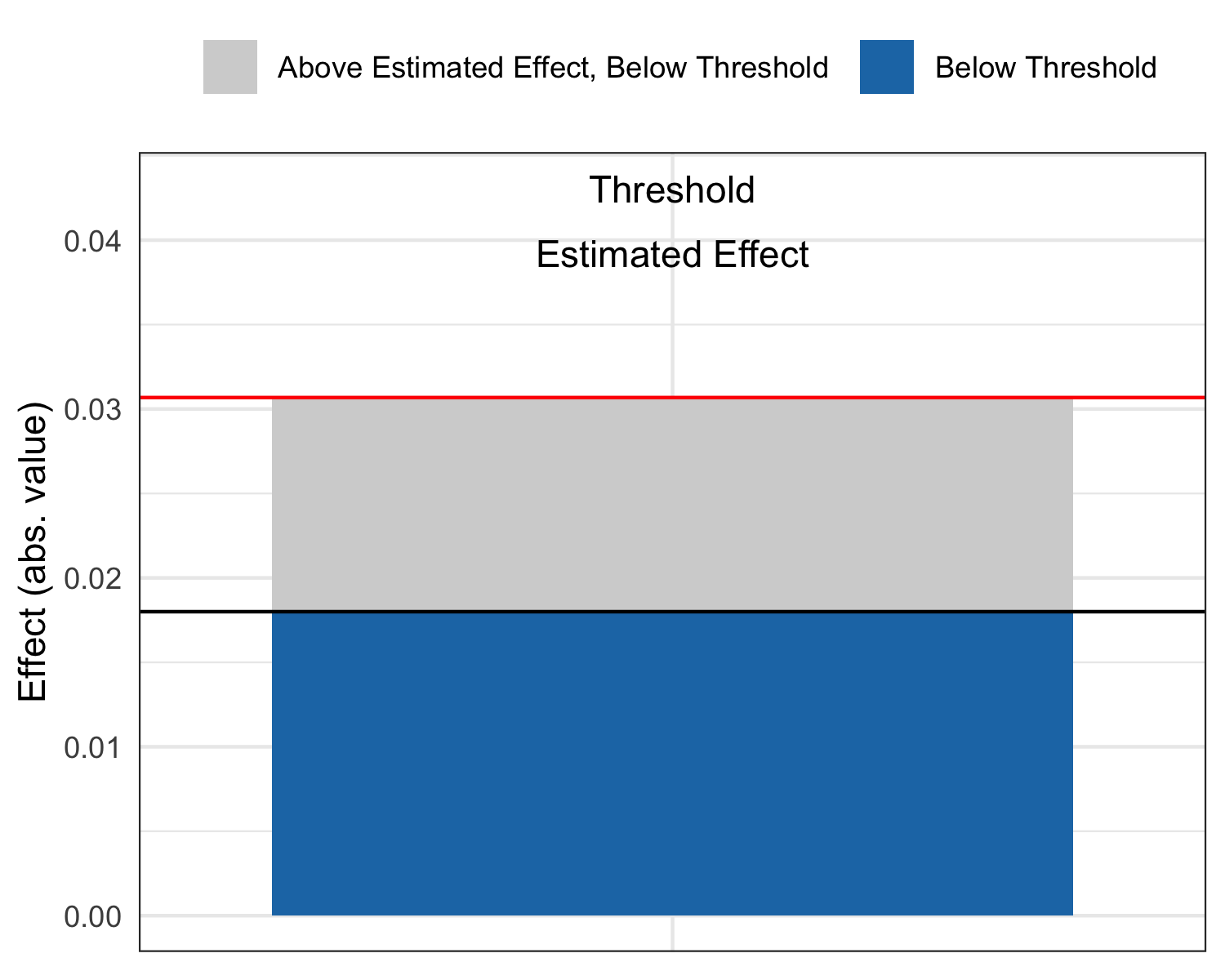
Like with pkonfound(), we can also output multiple forms of output at once with konfound():

## Percent Bias Necessary to Invalidate the Inference:  
## To sustain an inference, 41.327% of the estimate would have to be due to bias. This is based on a threshold of -0.031 for statistical significance (alpha = 0.05).  
## To sustain an inference, 13 of the cases with 0 effect would have to be replaced with cases at the threshold of inference.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.322 with the outcome and at 0.322 with the predictor of interest (conditioning on observed covariates) to sustain an inference based on a threshold of -0.031 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.322 X 0.322 = 0.104 to sustain an inference.

## Created 3 forms of output. To access type:   
##   
## konfound\_output$raw\_output  
## konfound\_output$thresh\_plot  
## konfound\_output$corr\_plot

Again, we can type each of those, i.e.:

## # A tibble: 1 x 8  
## action inference percent\_bias\_to… replace\_null\_ca… unstd\_beta  
## <chr> <chr> <dbl> <dbl> <dbl>  
## 1 to\_su… fail\_to\_… 41.3 13 -0.018  
## # … with 3 more variables: beta\_threshhold <dbl>,  
## # omitted\_variable\_corr <dbl>, itcv <dbl>



We can also test all of the variables as predictors of interest:

## # A tibble: 3 x 8  
## var\_name t df action inference pct\_bias\_to\_chang… itcv r\_con  
## <chr> <dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl>  
## 1 wt -5.79 29 to\_inva… reject\_null 51.5 0.585 0.765  
## 2 hp -1.2 29 to\_sust… fail\_to\_re… 38.7 -0.102 0.319  
## 3 qsec 1.16 29 to\_sust… fail\_to\_re… 40.5 -0.106 0.326

Whereas this cannot be carried out with pkonfound(), with konfound() you can also return a table with some key output from the correlation-based approach.

## Dependent variable is mpg

## Warning: Unknown or uninitialised column: 'itcv'.

## Warning: Unknown or uninitialised column: 'impact'.

## # A tibble: 4 x 7  
## term estimate std.error statistic p.value itcv impact  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 27.6 8.42 3.28 0.003 NA NA   
## 2 wt -4.36 0.753 -5.79 0 0.243 NA   
## 3 hp -0.018 0.015 -1.19 0.244 NA 0.511  
## 4 qsec 0.511 0.439 1.16 0.255 NA 0.073

If the impact threshhold is greater than the impacts of the Zs (the other covariates) then an omitted variable would have to have a greater impact than any of the observed covariates to change the inference. Note that in fields in which there is a lot known about covariates given the outcome of interest, then the omitted ones are likely less important than those that are known an included (i.e., we have a good sense of the factors that matter in terms of educational achievement).

### Example 2B: For generalized linear models fit with glm()

Effects for these models are interpreted on the basis of average partial (or marginal) effects (calculated using the margins package).

## Percent Bias Necessary to Invalidate the Inference:  
## To sustain an inference, 80.978% of the estimate would have to be due to bias. This is based on a threshold of 0.013 for statistical significance (alpha = 0.05).  
## To sustain an inference, 17334 of the cases with 0 effect would have to be replaced with cases at the threshold of inference.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 5.535 with the outcome and at 5.535 with the predictor of interest (conditioning on observed covariates) to invalidate an inference based on a threshold of 1.003 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 5.535 X 5.535 = 30.636 to invalidate an inference.

## NULL

As with models fit with lm() (and use of pkonfound()), multiple forms of output can be specified with the to\_return argument to konfound(), i.e. konfound(m2, age, to\_return = c("raw\_output", "corr\_plot", "thresh\_plot")).

### Example 2C: For mixed effects (or multi-level) models fit with the lmer() function from the lme4 package

konfound also works with models fit with the lmer() function from the package lme4, for mixed-effects or multi-level models. One challenge with carrying out sensitivity analysis for fixed effects in mixed effects models is calculating the correct denominator degrees of freedom for the t-test associated with the coefficients. This is not unique to sensitivity analysis, as, for example, lmer() does not report degrees of freedom (or p-values) for fixed effects predictors (see this information in the lme4 FAQ [here](http://bbolker.github.io/mixedmodels-misc/glmmFAQ.html#why-doesnt-lme4-display-denominator-degrees-of-freedomp-values-what-other-options-do-i-have)). While it may be possible to determine the correct degrees of freedom for some models (i.e., models with relatively simple random effects structures), it is difficult to generalize this approach, and so in this package the Kenward-Roger approximation for the denominator degrees of freedom as implemented in the pbkrtest package (described in [Halekoh and Højsgaard, 2014](https://www.jstatsoft.org/htaccess.php?volume=59&type=i&issue=09&paper=true)).

Here is an example of the use of konfound() with a model fit with lmer():

## Warning in bind\_rows\_(x, .id): binding factor and character vector,  
## coercing into character vector

## Warning in bind\_rows\_(x, .id): binding character and factor vector,  
## coercing into character vector

## Percent Bias Necessary to Invalidate the Inference:  
## To invalidate an inference, 84.83% of the estimate would have to be due to bias. This is based on a threshold of 1.588 for statistical significance (alpha = 0.05).  
## To invalidate an inference, 137 observations would have to be replaced with cases for which the effect is 0.  
##   
## Impact Threshold for a Confounding Variable:  
## An omitted variable would have to be correlated at 0.817 with the outcome and at 0.817 with the predictor of interest (conditioning on observed covariates) to invalidate an inference based on a threshold of 0.155 for statistical significance (alpha = 0.05).  
## Correspondingly the impact of an omitted variable (as defined in Frank 2000) must be 0.817 X 0.817 = 0.667 to invalidate an inference.

## NULL

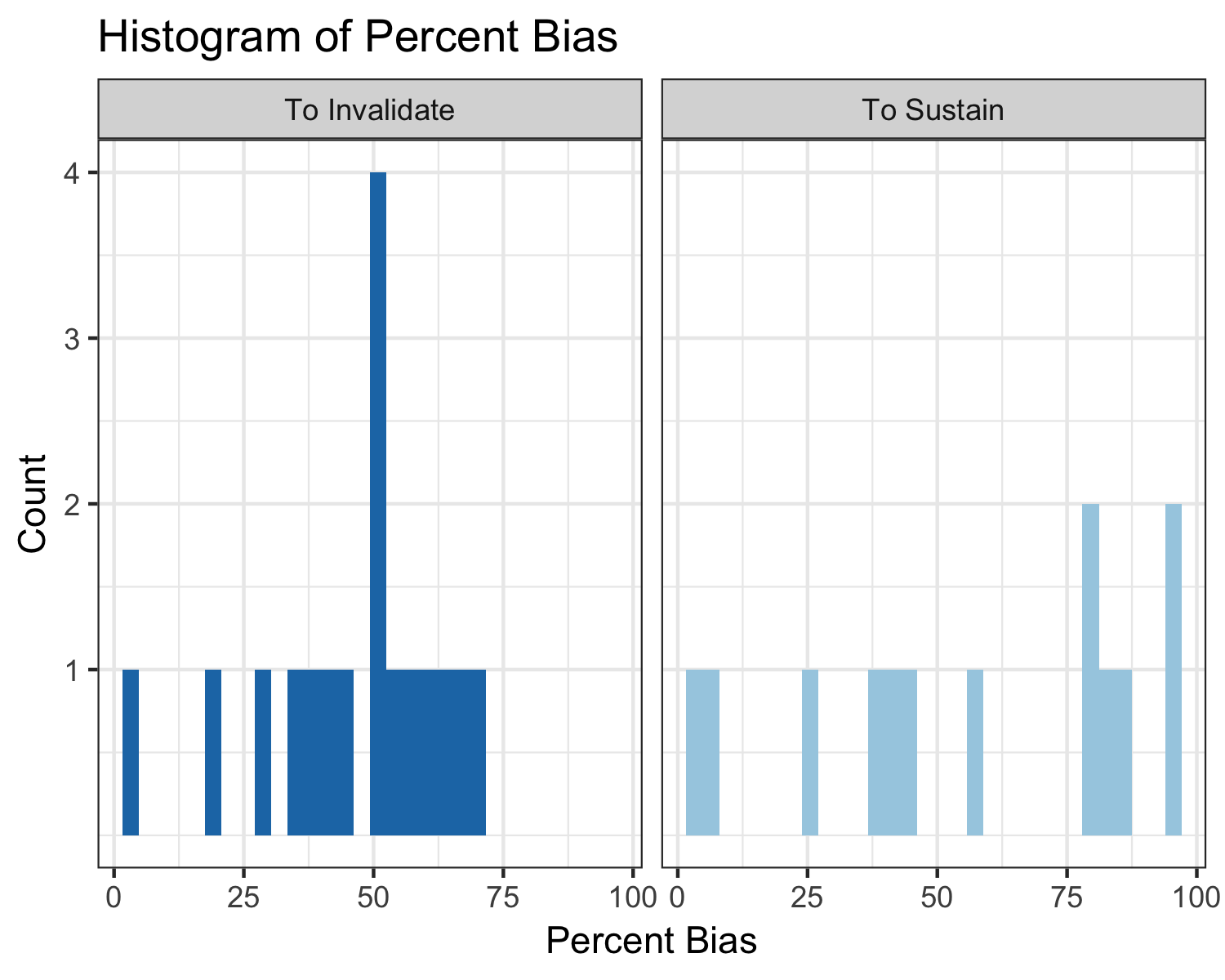
### Example 2D: Use of mkonfound() for meta-analyses that include sensitivity analysis

We can also use konfound to carry out sensitivity analysis as part of meta-analyses. For example, here, d represents output from a number (30 in this case) of past studies, read in a CSV file from a website:

## t df  
## 1 7.076763 178  
## 2 4.127893 193  
## 3 1.893137 47  
## 4 -4.166395 138  
## 5 -1.187599 97  
## 6 3.585478 87

## # A tibble: 30 x 7  
## t df action inference pct\_bias\_to\_change\_i… itcv r\_con  
## <dbl> <int> <chr> <chr> <dbl> <dbl> <dbl>  
## 1 7.08 178 to\_invali… reject\_null 68.8 0.378 0.614  
## 2 4.13 193 to\_invali… reject\_null 50.6 0.168 0.41   
## 3 1.89 47 to\_sustain fail\_to\_reje… 5.47 -0.012 0.11   
## 4 -4.17 138 to\_invali… reject\_null 50.3 0.202 0.449  
## 5 -1.19 97 to\_sustain fail\_to\_reje… 39.4 -0.065 0.255  
## 6 3.59 87 to\_invali… reject\_null 41.9 0.19 0.436  
## 7 0.282 117 to\_sustain fail\_to\_reje… 85.5 -0.131 0.361  
## 8 2.55 75 to\_invali… reject\_null 20.6 0.075 0.274  
## 9 -4.44 137 to\_invali… reject\_null 53.0 0.225 0.475  
## 10 -2.05 195 to\_invali… reject\_null 3.51 0.006 0.077  
## # … with 20 more rows

We can also return a plot summarizing the percent bias needed to sustan or invalidate an inference across all of the past studies:



# Examples of publishable write-ups

(from Beymer) (for correlation-based approach) (ref appendix)

# References

Behn, R.D. and Vaupel, J.W., 1982. Quick analysis for busy decision makers (p. 3). New York: Basic Books. Cohen, J., Cohen, P., West, S.G. and Aiken, L.S., 1983. Applied multiple regression for the behavioral sciences. Laurence Erlbaum, Hillsdale, NJ. Frank, K. A. and Min, K. 2007. Indices of Robustness for Sample Representation. Sociological Methodology. Vol 37, 349-392. Frank, K.A. 2000. Impact of a Confounding Variable on the Inference of a Regression Coefficient. Sociological Methods and Research, 29(2), 147-194 Frank, K.A., Gary Sykes, Dorothea Anagnostopoulos, Marisa Cannata, Linda Chard, Ann Krause, Raven McCrory. 2008. Extended Influence: National Board Certified Teachers as Help Providers. Education, Evaluation, and Policy Analysis. Vol 30(1): 3-3 Frank, K.A., Maroulis, S., Duong, M., and Kelcey, B. 2013. What would it take to Change an Inference?: Using Rubin’s Causal Model to Interpret the Robustness of Causal Inferences. Education, Evaluation and Policy Analysis. Vol 35: 437-460. Pan, W., and Frank, K.A. 2004. An Approximation to the Distribution of the Product of Two Dependent Correlation Coefficients. Journal of Statistical Computation and Simulation, 74, 419-443 Pan, W., and Frank, K.A., 2004. A probability index of the robustness of a causal inference. Journal of Educational and Behavioral Statistics, 28, 315-337. Rubin, D. B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” Journal of Educational Psychology 66:688–701.’