Effect of QRPs:

## No censoring:

RE: Increases type I error and power, especially with increased k. Increased ME and RMSE. Decreases coverage, especially at high k.

TF: Increases Type I error under homogeneity, reduces it very slightly under heterogeneity. Very small effect on ME, causing upward bias under homogeneity and having **complex relationship with d = 0.5 and tau and k. Complex relationship with RMSE. Complex relationship with coverage.**

PET-PEESE: Increases Type I error, reduces power. Exacerbates downward bias. (Maybe the Type I errors are in the negative direction?) Increases RMSE. Decreases coverage.

p-curve: Reduces ME in all circumstances, alleviating somewhat positive bias at delta = 0, tau = 0, and tau = 0.2. Undershoots delta = 0.5 tau = 0.2. Reduces RMSE where it alleviates positive bias.

p-uniform: Reduces ME in all circumstances, alleviating somewhat the positive bias at delta = 0, tau = 0 and tau = 0.2. Undershoots delta = 0.5 tau = 0.2. Generally reduces RMSE for delta = 0, but delta = 0.5. Causes overcoverage at delta = 0, tau = 0, undercoverage at delta = 0.5, tau = 0, causes better coverage at delta = 0.5 except for large k delta 0.5 tau 0.2.

3PSM: QRPs reduce power and Type I error. Inflict decrease in ME (meaning greater negative bias) in all conditions. Generally leads to an increase in RMSE because of this bias and a noticeable drop in coverage, particularly at large *k*.

WAAP-WLS: QRPs have a **complex relationship** with Type I error and power. Cause a slight increase in ME under most circumstances but seem to exacerbate downward bias at delta = 0.5, tau = 0.2, k = 60. Complicated! Very small influence on RMSE, usually .01 of a point or less. Generally reduces coverage but can be complex.

## Medium censoring:

RE: QRPs increase Type I error. QRPs increase bias. QRPs increase RMSE. Exception: RMSE reduced slightly for delta = 0.5, tau = 0.2. QRPs generally lead to poorer coverage.

TF: QRPs increase Type I error. QRPs increase ME, chiefly for delta = 0. Exception: ME reduced for delta = 0.5, tau = 0.2. QRPs increase RMSE under the null, very slightly reduce it under H1. QRPs generally lead to poorer coverage when the null was true, better coverage when the null was false.

PET-PEESE: QRPs increase Type I error under homogeneity, have complex relationship with it under heterogeneity (medium p-hacking reduces power and Type I; heavy p-hacking increases power and Type I). QRPs cause negative bias ME. QRPs increase RMSE across the board. Complicated relationship with CI coverage, but generally impairs it.

p-curve: QRPs lead to reduction in mean estimation, which leads to underestimations sometimes and accurate estimations at other times. QRPs reduce RMSE when p-curve would normally be biased (d = 0, tau = 0, or tau = 0.2), increase it otherwise (e.g. d = 0.5, tau = 0).

p-uniform: Same as p-curve. Better CI coverage (and even overcoverage) under H0, poorer coverage under H1, unless the downward bias cancels out the upward bias just so.

3PSM: QRPs reduce Type I error and power. QRPs cause reduction in ME, negative bias. Complex relationship with RMSE – generally reduces RMSE under H0, increases it under H1, but exception at k = 60, delta = 0, tau = 0.2, qrp = high. Improves coverage under H0, but generally hurts it otherwise. Again, a u-shaped relationship with bias makes this complicated.

WAAP-WLS: QRPs increase Type I and Type II error. However, at k = 60, QRPs decrease Type I error rate (which is 85%+). Increases ME under the null, slight reduction in ME under H1. QRPs generally increase RMSE under H0, have minimal influence under H1. Hurts coverage under H0, improves it slightly under H1.

## Heavy censoring

RE: QRPs increase Type I error. QRPs increase mean error when the null is true, but can reduce it slightly at delta = 0.5, tau = 0.2. QRPs increased RMSE, but decreased it for delta = 0.5, tau = 0.2. QRPs caused undercoverage when the null was true and reduced undercoverage when the null was false.

TF: QRPs increase Type I error. Like RE, QRPs increase mean error when the null is true, but decrease it under delta = 0.5, tau = 0.2. QRPs increased RMSE, but decreased it for delta = 0.5, tau = 0.2. QRPs caused undercoverage when the null was true, had minimal effect at delta = 0.5 tau = 0, and improved coverage at delta = 0.5, tau = 0.2.

PET-PEESE: QRPs increase Type I error under homogeneity but decrease it under heterogeneity. Little effect on power under a true effect, although it can reduce power slightly to detect a true effect, especially under heterogeneity. QRPs lead to lower ME, turning into negative bias in most cases. QRPs exacerbated PET-PEESE’s underestimation of true effects. QRP’s relationship with RMSE is very complicated. QRPs reduced coverage except for delta = 0, tau = 0.2.

p-curve: QRPs caused reduction of ME in all circumstances, sometimes leading to negative bias. This reduces RMSE when the null is true or there is heterogeneity and increases it otherwise.

p-uniform: QRPs caused reduction of ME in all circumstances, sometimes leading to negative bias. This reduces RMSE when the null is true or there is heterogeneity and increases it otherwise. QRPs improved coverage when the null was false, but otherwise improved or impaired coverage according to the ME’s being positive, zero, or negative.

3PSM: QRPs generally cause a decrease in Type I error. There is an exception in that high QRPs increase Type I error relative to medium QRPs for k = 60, tau = 0.2. QRPs cause a loss of power under all conditions. QRPs generally caused a decrease in ME, sometimes becoming negative bias. Relationship with RMSE was complicated, interacting with k, delta, and tau. QRPs improved coverage when the null was false, but otherwise improved or impaired coverage according to the ME’s being positive, zero, or negative.

WAAP-WLS: QRPs lead to reduced Type I error rates and lower power. There is an exception in that medium and heavy p-hacking increase type I error at delta = 0, tau = 0.2. Under the null and small k, QRP caused upward bias. Under the alternative and small k, QRP reduced upward bias. At large k, QRP reduced upward bias. Increased RMSE for small k and delta = 0, reduced it otherwise. QRPs generally improved coverage, although there were exceptions.

How to simplify this:  
RE and TF: QRPs generally cause increased bias. However, when there is publication bias and there are true, heterogeneous effects (delta = 0.5, tau = 0.2), p-hacking alleviates bias slightly.

WAAP-WLS: QRPs generally cause increased bias. However, when there is publication bias and there are true, heterogeneous effects, p-hacking alleviates bias slightly. It can also alleviate bias when publication bias is very strong and k is large.

p-uniform, p-curve, 3PSM: QRPs reduce the ME of these estimates. This can cause ME to go negative.

PET-PEESE: QRPs generally caused downward bias.