## Prevalence of invalid inferences

Two independent raters assessed the validity of the inferences made within each article, i.e., whether each article contained an inference on the basis of a hypothesis test that is now known to be confounded. Inter-rater reliability was found to be XXX. Quotes substantiating confounded inferences are available in the supplementary materials. 84.7% of articles were found to contain at least one invalid inference on the basis of a hypothesis test that is now known to be confounded. No changes over time were observed, OR = 0.88, 95% CI [0.72, 1.03], *p* = .138, *R*2Tjur = .02.

78.6% of articles were found to contain at least one invalid inference as part of its key claims. No changes over time were observed, OR = 0.94, 95% CI [0.81, 1.08], *p* = .407, *R*2Tjur = .01.