

## Misuse of Standardization to Meta-analyze Differences in Means

Will G Hopkins, David S Rowlands

Sportscience 27, 28, 2024 ([sportsci.org/2024/MisuseMeta.htm](https://sportsci.org/2024/MisuseMeta.htm))

Internet Society for Sport Science, Auckland, New Zealand; School of Sport Exercise and Nutrition, Massey University, Auckland, New Zealand. [Email](#).

Meta-analysts often use standardized mean differences (SMD) to combine mean effects from studies in which the dependent variable has been measured with different instruments or scales. The SMD is properly calculated as the difference in means divided by a between-subject reference-group, control-group, or pre-intervention standard deviation (SD), usually free of measurement error. When combining mean effects from controlled trials and crossovers, many meta-analysts divide instead by an SD of change scores, resulting in SMDs that have no useful interpretation and that can underestimate or grossly overestimate the magnitude of the intervention. Others standardize using only post-intervention means and pooled SD, which usually results in reduced precision of the SMD and underestimation of the SMD arising from individual responses to the intervention. These misuses of standardization were frequent in recent meta-analyses in medical journals we surveyed; they arise apparently from misleading advice in peer-reviewed publications and from inappropriate use of popular meta-analysis packages. In any case, meta-analysis of any form of SMD increases heterogeneity artifactually via differences in standardizing SD between settings. We therefore favor other approaches to combining mean effects of disparate measures: log transformation of factor effects (response ratios) and of percent effects converted to factors; rescaling of psychometrics to percent of maximum range; and rescaling with minimum clinically important differences. If meta-analysts cannot adduce clinically important thresholds for mean effects, standardization *after* meta-analysis with appropriately transformed or rescaled chosen or pooled pre-intervention SDs is a fallback for assessing magnitudes of a meta-analyzed mean effect in different settings.

Keywords: change-score SD, Cochrane, Comprehensive Meta-analysis (CMA), factor effect, meta-analysis, metafor, RevMan, standardized mean difference (SMD),

[Reprint pdf](#) · [Reprint docx](#) · [Slideshow](#)

We have long known that researchers sometimes misuse standardization to combine mean effects in a meta-analysis. When we encountered a particularly egregious example several years ago, we surveyed several medical journals for the prevalence of misuse. Only ~10% of studies used the correct standard deviations to standardize, so we decided to write an article explaining the wrong and right ways to meta-analyze mean effects.

The emphasis of the article was originally the misuse of standardization, but during the review process, the editor requested a revision into a tutorial in biostatistics covering all the methods for meta-analyzing differences in means. The article

has now (May 7) been accepted for the journal *Statistics in Medicine*, where it will appear with the title "Standardization and Other Approaches to Meta-Analyze Differences in Means."

The slideshow attached to this article was presented by one of us (WGH) in several European universities in November 2023. The slideshow and the above abstract reflect the original emphasis on misuse of standardization, but all the methods for meta-analyzing mean changes are described. Make sure you view the slideshow as a full presentation to get the benefit of the extensive animations.

Published April 2024; updated May 2024

[©2024](#)