Arcsine-based transformations for meta-analysis of proportions: Pros, cons, and alternatives

Lifeng Lin | Chang Xu

1Department of Statistics, Florida State University, Tallahassee, Florida
2Department of Population Medicine, College of Medicine, Qatar University, Doha, Qatar

Abstract

Meta-analyses have been increasingly used to synthesize proportions (e.g., disease prevalence) from multiple studies in recent years. Arcsine-based transformations, especially the Freeman–Tukey double-arcsine transformation, are popular tools for stabilizing the variance of each study's proportion in two-step meta-analysis methods. Although they offer some benefits over the conventional logit transformation, they also suffer from several important limitations (e.g., lack of interpretability) and may lead to misleading conclusions. Generalized linear mixed models and Bayesian models are intuitive one-step alternative approaches, and can be readily implemented via many software programs. This article explains various pros and cons of the arcsine-based transformations, and discusses the alternatives that may be generally superior to the currently popular practice.

KEYWORDS

arcsine-based transformation, Bayesian model, generalized linear mixed model, meta-analysis, proportion

1 | BACKGROUND

Many research findings in the health sciences are presented in the form of proportions, such as disease prevalence, case fatality rate, a diagnostic test's sensitivity and specificity, among others.1,2 Meta-analyses have been increasingly used to synthesize proportions that are reported from multiple studies on the same research topic.3-10 Many meta-analyses of proportions are performed using conventional two-step methods. First, a specific transformation is usually applied to each study's proportion estimate for better approximation to the normal distribution, as required by the assumptions of conventional meta-analysis models.11 Second, the meta-analysis is performed on the transformed scale, and the synthesized result is then back-transformed to the original proportion scale that ranges from 0% to 100%.

Of note, one may also directly synthesize proportions without any transformation; however, this approach is not optimal, because the proportion estimates may not be approximately normally distributed, especially for rare events and small sample sizes. The Wald-type confidence intervals (CIs) of proportions may be even outside the range of 0% to 100%.12

Various transformations are available for proportions, including the log, logit, arcsine-square-root, and Freeman–Tukey double-arcsine transformations.13-17 Among them, the Freeman–Tukey double-arcsine transformation is a popular tool in current practice of synthesizing proportions.10 We did a search on Google Scholar on June 17, 2020; for each year between 2000 and 2019, we searched for the exact terms “meta-analysis” and “double-arcsine” to obtain the number of research items using the double-arcsine transformation in meta-analyses. We also searched for the exact term “meta-analysis,” with restriction to article titles, to obtain the rough number of meta-analysis publications in each year, and calculated the corresponding proportion of research items using the double-arcsine transformation. Figure 1 shows that the double-arcsine transformation has been increasingly used over the past two decades.
Despite the raising popularity, several authors have previously expressed concerns about arcsine-based transformations. In addition, many meta-analyses do not even specify the transformation used for synthesizing proportions. Even if a transformation is specified, meta-analysts frequently fail to provide sufficient justification for the selection of the transformation.

This article discusses the purported benefits of the arcsine-based transformations that potentially explain their popularity in current practice. We also introduce how such transformations may be limited, and recommend alternative methods for meta-analysis of proportions that may be superior. We focus on meta-analysis of single proportions, where the arcsine-based transformations are widely used.

2 | METHODS

Suppose a meta-analysis contains \( N \) studies that report single proportions on a common topic. Let \( p_i \) be the proportion estimate from study \( i \) in the meta-analysis (\( i = 1, \ldots, N \)). The proportion is then simply calculated as \( p_i = e_i/n_i \), where \( e_i \) and \( n_i \) denote study \( i \)'s event count and sample size, respectively. The arcsine-square-root transformation is

\[
y_i = \arcsin \sqrt{e_i/(n_i + 1)}
\]

with variance \( \nu_i = 1/(n_i + 0.5) \). Of note, the formula above of the double-arcsine transformation is the version originally presented in the article by Freeman and Tukey. While one may also take the average of the two arcsine values (by dividing by 2), leading to the variance \( \nu_i = 1/(4n_i + 2) \), so that it has the same scale with the arcsine-square-root transformation, such a linear transformation does not affect the back-transformed proportion.

3 | PROS

Because conventional meta-analysis models assume normally distributed data, the various transformations are applied to the proportion data in an effort to yield better approximations to the normal distribution. As shown in the formulas above, the variances of the arcsine-based transformations depend only on the sample sizes, which are typically fixed, known values. However, the variances of the log and logit transformations depend additionally on the event counts, which are random variables.

Involving event counts in the variances implies several limitations of the log and logit transformations. First, it does not meet the assumption of conventional meta-analysis models; that is, the within-study variances are treated as fixed, known values, while the event counts are not. The violation of this assumption may reduce statistical inference accuracy. Second, because the log- or logit-transformed proportion estimates and their variances both depend on the event counts, they are intrinsically correlated. This intrinsic correlation has been well known to cause substantial biases in meta-analytic results,
especially when the sample sizes are small. In addition, in the presence of zero event counts, both log- and logit-transformed proportions cannot be calculated, and a continuity correction must be applied to the zero counts, usually by adding 0.5. This correction may have considerable impact on the synthesized proportion for rare events.

In this sense, the arcsine-based transformations have the important advantage of stabilizing variances, which is likely the main reason that such transformations are widely used in current practice. As their variances depend only on the sample sizes, they can be validly treated as fixed, known values, and have no correlation with the transformed proportion estimates. These transformations also do not need the continuity correction for zero counts. Moreover, compared with the arcsine-square-root transformation, the Freeman–Tukey double-arcsine transformation may stabilize variances better in general.

## CONS

Despite the advantages of the arcsine-based transformations listed above, they also suffer from several critical limitations. First, these transformations lack intuitive interpretations from practical perspectives, especially compared with the traditional log transformation. The arcsine function is mainly used for technical purposes. More specifically, the variances of transformed proportions are approximated by the so-called delta method, which uses the derivative of the transformation function. Taking the benefit of the special structure of the arcsine function’s derivative, the event counts are canceled out in the approximated variances of the arcsine-transformed proportions. Unlike the logit proportion that represents the odds on a logarithmic scale, the arcsine-transformed proportions may not be intuitive for practitioners.

Second, as the proportion estimates are usually heterogeneous, the random-effects meta-analysis method is frequently used, assuming that each study’s underlying true transformed proportion follows the normal distribution across studies. The arcsine-based transformations might violate this assumption, because the arcsine function has a bounded domain, implying truncations for the assumed normal distribution. On the other hand, the logit-transformed proportions can take any real value, and thus, may be more suitable for the between-study normality assumption.

Third, the Freeman–Tukey double-arcsine transformation has a complicated form of back-transformation to the original proportion scale. Compared with other transformations, its back-transformation depends additionally on a sample size that represents the overall synthesized result. This “overall sample size” is not well defined in the meta-analysis setting; it may be selected as the harmonic, geometric, and arithmetic means of study-specific sample sizes, or the inverse of the variance of the synthesized result; it is generally difficult to justify the value used as the “overall sample size.” More importantly, different values may lead to substantially different proportions in some cases, potentially leading to misleading conclusions.

Moreover, numerical problems may occur when using the Freeman–Tukey double-arcsine transformation. Although this transformation refines the usual arcsine-square-root transformation by averaging over the double arcsines for better stabilizing variances, it may have low accuracy at values close to its domain limits, which likely occur in cases of rare events and small sample sizes. Specifically, because the event count $e_i$ is between 0 and $n_i$, as indicated in Equation (1), the transformed proportion must be bounded between $\arcsin\sqrt{1/(n_i + 1)}$ and $\arcsin\sqrt{n_i/(n_i + 1)} + \pi/2$. It is possible that the synthesized result is outside this domain based on a certain “overall sample size.” In this case, the result cannot be back-transformed to the original proportion scale. When such issues occur, one may decide to use the back-transformation of the arcsine-square-root transformation, which is a good approximation of the Freeman–Tukey double-arcsine transformation for sufficiently large sample sizes; however, this might affect the accuracy of the analysis.

## ALTERNATIVES

From a statistical perspective, event counts are typically assumed to follow binomial distributions, and all transformations discussed above are applied to the binomial data for approximations to normal distributions within studies. With advances in statistical computing techniques, these approximations in the two-step methods may be unnecessary, because they can be feasibly replaced with one-step meta-analysis methods, including generalized linear mixed models (GLMMs) or Bayesian hierarchical models.

GLMMs directly model event counts with binomial likelihoods and fully account for within-study uncertainties. They use a specific link function to transform study-specific latent true proportion to a linear scale, on which random effects are specified in a manner similar to the conventional two-step methods. The logit link is the canonical link function for proportions (ie, binomial data), while many other links, such as the log and probit links, may be also used. GLMMs with the log and logit links correspond to the log and logit transformations used in the two-step methods, but they do not have any of the aforementioned limitations. Specifically, GLMMs do not involve estimating (transformed) proportions and their variances at the within-study level. Therefore, they do not suffer from the problems caused by the intrinsic correlation between the log- or logit-transformed proportions and their sample variances approximated by the delta method. GLMMs can also effectively model zero event counts without continuity corrections. More importantly, compared with the arcsine-based transformations, the GLMM with the logit link produces more interpretable results.

Similarly, the multilevel structure of meta-analyses can be naturally modeled under the Bayesian framework. Bayesian methods assign priors to the unknown parameters, including the overall proportion and the heterogeneity variance on the transformed scale; the conclusions are drawn from the posterior distributions of these parameters. As one of the Bayesian methods’ benefits, researchers
GLMMs and Bayesian models have been seldom used in meta-analysis applications so far, despite the fact that the current literature offers many software programs to implement these alternative approaches for synthesizing proportions (as well as other measures), including SAS, R, and Stata. Bayesian models can be fitted via BUGS, JAGS, Stan, and others that are designed for general purposes of Bayesian analyses. When the number of studies or the number of events in a meta-analysis is small (say, <10), GLMMs and Bayesian models may have issues about their algorithm convergence (more specifically, for maximizing likelihood and deriving posterior samples from the Markov chain Monte Carlo, respectively). In such situations, although the conventional two-step methods might successfully produce results, the synthesized proportion may be subject to large biases and thus should be interpreted with great caution.

As stated above, this article has focused on meta-analysis of single proportions, where the arcsine-based transformations are widely used. GLMMs and Bayesian methods are also available for jointly modeling multiple proportions, such as sensitivity and specificity of diagnostic tests.

6 | DISCUSSION

Compared with the traditional logit transformation, the arcsine-based transformations for proportions mainly benefit from their stabilized variances that depend only on sample sizes. However, they do not have intuitive interpretations, and the limitations of the logit transformation can be easily overcome by using GLMMs or Bayesian models. These alternatives are straightforward, one-step methods and are generally superior to the conventional two-step methods that require transformations of proportions within studies. Importantly, in some cases, the one-step methods may lead to substantially different results from the two-step methods. In future studies, it is worthwhile to explore the performance of the different methods with various transformations or link functions based on a large collection of empirical meta-analysis datasets, and quantitatively investigate the differences between the synthesized proportions produced by these methods.

The limitations of the arcsine-based transformations in meta-analysis, however, do not nullify their use in individual studies. We have focused on the synthesis of proportions, and GLMMs or Bayesian models are advantageous for producing such synthesized proportions. When meta-analysts want to present estimates from individual studies, transformations of proportions are still useful. For example, meta-analysts frequently use the forest plot to visualize the distributions of study-specific estimates, and the funnel plot to assess potential publication bias or small-study effects. Both plots depend on each study’s point estimate of proportion, CI, and SE, which cannot be obtained by GLMMs or Bayesian models. The arcsine-based transformations can be preferably used at the within-study level, while they are not recommended at the between-study level. In fact, many early articles on the arcsine-based transformations were discussed in the setting of an individual study; these articles did not directly suggest extending the arcsine-based transformations to the meta-analysis setting.

In summary, we highly recommend the use of GLMMs or Bayesian models for synthesizing proportions; nowadays, many software programs are readily available for implementing them. Most meta-analyses of proportions published in recent years continue to use the Freeman-Tukey double-arcsine transformation, and the rate is increasing (Figure 1); it is a time for change.

REFERENCES


---