Seriously Misleading Results Using Inverse of Freeman-Tukey Double Arcsine Transformation in Meta-Analysis of Single Proportions

Guido Schwarzer[1], Hiam Chemaitelly[2], Laith J. Abu-Raddad[2],

Gerta Rücker[1]

[1] Institute of Medical Biometry and Statistics,

Faculty of Medicine and Medical Center - University of Freiburg,

Germany

[2] Infectious Disease Epidemiology Group,

Weill Cornell Medicine-Qatar, Cornell University,

Qatar Foundation, Education City, Doha,

Qatar

Abstract

Standard generic inverse variance methods for the combination of single proportions are based on transformed proportions using the logit, arcsine, and Freeman-Tukey double arcsine transformations. Generalized linear mixed models are another more elaborate approach. Irrespective of the approach, meta-analysis results are typically back-transformed to the original scale in order to ease interpretation. Whereas the back-transformation of meta-analysis results is straightforward for most transformations, this is not the case for the Freeman-Tukey double arcsine transformation, albeit possible.

In this case study with five studies, we demonstrate how seriously misleading the back-transformation of the Freeman-Tukey double arcsine transformation can be. We conclude that this transformation should only be used with special caution for the meta-analysis of single proportions due to potential problems with the back-transformation. Generalized linear mixed models seem to be a promising alternative.

1 Introduction

A key application of meta-analytical methods is the pooling of proportions, such as prevalence of a specific infection or disease [1, 2, 3, 4]. Classic fixed effect and random effects meta-analysis methods [5] are typically used to combine single proportions. In order to use these methods, proportions are generally transformed using either the log [6], logit [7], arcsine [8], or Freeman-Tukey double arcsine [9] transformations. These transformations are implemented for pure mathematical reasons, e.g. variance stabilization (details on the transformations are given in the appendix and summarized in Table A1). For pooling, the transformed proportions and corresponding standard errors are used in the generic inverse variance method [5]. An alternative yet more elaborate approach based on the logit transformation are generalized linear mixed models [10] which account for the binomial structure of the data and thus avoid the generic inverse variance method. Irrespective of the meta-analysis method and transformation, results are usually presented on the original probability scale after using the corresponding back-transformation.

Whereas the back-transformation of meta-analysis results is straightforward for the log, logit, and arcsine transformations, this is not the case for the Freeman-Tukey double arcsine transformation, albeit possible [11]. In order to calculate the inverse of the Freeman-Tukey double arcsine transformation, a single sample size has to be specified. Accordingly, for a single study a one to one relation exists between transformation and its inverse, however, in a meta-analysis with different sample sizes the value of the back-transformation depends on the specified sample size. Typically, the harmonic mean of sample sizes is used in the back-transformation [11].

1

2 Case study: meta-analysis on prevalence of hepatitic C virus infections

We report results of meta-analyses with five studies estimating the prevalence of hepatitis C virus (HCV) infections in the general population of Nepal which constitute a subset of an unpublished dataset with 28 studies [12]. This unpublished dataset comprises testing for a total of 972 123 individuals among whom 3 696 were HCV antibody positive. The prevalence across studies ranged from 0% to 18.4% with a median of 0.5%. We restrict ourselves to the five-study subset for didactic reasons; the same issues encountered in this subset also exist in the full dataset.

We conducted classic meta-analyses using the arcsine, Freeman-Tukey double arcsine, and logit transformations, respectively. Furthermore, we fit generalized linear mixed models (GLMMs) implicitly using the logit transformation. Details on the statistical methods are provided in the appendix. We used R function `metaprop()` from R package **meta** [13] (see supplementary online material). Results are summarized in Table 1.

Under the fixed effect model, results depicted as transformed proportions (middle column in Table 1) are very similar for the two methods using the arcsine and logit transformations, respectively. Whereas the random effects estimates are also very similar with a slightly smaller confidence interval for the arcsine transformation, the results for the two logit methods are rather different due to a very different estimate for the between-study variance.

For easier interpretation, results are back-transformed to the original scale. Due to the small prevalences, we express results as HCV infections per 1 000 observations. In Table 1 (right column), the results using the inverse of the Freeman-Tukey double arcsine transformation based on the harmonic mean of 85 are highly irregular with

2

HCV prevalences and confidence limits exactly equal to zero. Under the fixed effect model, all of the other three methods show very similar results. Conversely, under the random effects model, results for the classic meta-analysis method using the logit transformation are very different from the other results.

Looking at Figure 1, we see that the meta-analysis estimators are reasonable summaries of transformed prevalences. On the other hand, back-transformed meta-analysis results are clearly off the mark in Figure 2 with meta-analysis estimators smaller than all individual study results. Note, the back-transformation works as expected for individual study results, e.g., the prevalence is $1/29 = 0.03448$ for study 26 which corresponds to 34.48 HCV infections per 1 000 observations.

The harmonic mean of 85 is obviously the wrong choice in this meta-analysis with sample sizes ranging from 29 to more than 200 000. Figure 3 shows the influence of sample size on meta-analysis results (see also Table A2 in the appendix). For sample sizes between 10 and around 120, results are exactly zero for the back-transformation of the Freeman-Tukey double arcsine transformation. The number of HCV infections per 1 000 observations then steeply increases up to a sample size of 500 when the effect of sample size starts to slowly level out.

As noted earlier, the results of the random effects model are very different for the two logit methods due to different between-study variance estimates. This discrepancy can be explained by looking at the confidence intervals of individual studies in the corresponding forest plots (Figures 4 and 5). Confidence intervals, based on the normal approximation, are much narrower for the two smallest studies in the classic random effects meta-analysis (Figure 4) than the confidence intervals, based on the Clopper-Pearson method taking the binomial distribution into account [14, 15], in the GLMM meta-analysis (Figure 5). Apparently, in these two small studies with only 1 HCV infection and less than 50 observations the assumption of a normally distributed logit transformed proportion is not fulfilled. With increasing numbers of

3

infections and sample sizes, approximate and Clopper-Pearson confidence intervals get closer to each other. Obviously, the very narrow confidence intervals of the two smallest studies result in an inflated between-study variance estimate leading to a larger estimate for the pooled mean HCV prevalence and a much wider confidence interval for the pooled mean HCV prevalence.

3 Discussion

Our case study shows that meta-analysis results based on the back-transformation of the Freeman-Tukey double arcsine transformation [11] can be very misleading and even smaller than all individual study results. We observe similar undesirable results in a meta-analysis using the complete dataset with 28 studies. To our knowledge this is the first publication reporting such an anomaly and erratic results.

In our view, the main reason for this unexpected behaviour is the very extreme pattern of sample sizes which range from 29 to more than 200 000. The harmonic mean of 85 is much smaller than 3 of the 5 sample sizes. For such highly skewed sample sizes, the harmonic mean is by definition rather small which may result in nonsensical back-transformed probabilities.

In order to prevent misleading conclusions for the Freeman-Tukey double arcsine transformation, several sample sizes could be used to evaluate the sensitivity of meta-analysis results, however, this may lead to diverging meta-analysis estimates. In our example, using the arithmetic or geometric mean in the back-transformation (see Table A2 in the appendix) would result in random effects estimates of 1.96 and 1.59 HCV infections per 1 000 observations, respectively. Here, results for the harmonic mean are obviously wrong, however, it is rather unclear whether to rely on the results for the arithmetic or geometric mean. All other transformations (arcsine, logit, log) do not have this intrinsic problem in the presentation of meta-analysis

4

results.

Overall, the arcsine transformation appears to be the best classic method for the meta-analysis of single proportions. However, as application of GLMMs for meta-analysis is nowadays straightforward due to its implementation in common software, there is neither a real reason nor a clear advantage for using an approximate method. Accordingly, we support the viewpoint of [10], [16], [17], and [18] recommending the use of GLMMs for the meta-analysis of single proportions. From our perspective, the only disadvantage of a GLMM is that individual study weights are not available which we consider as a minor drawback; analysts seeing this differently should use the arcsine transformation.

Our recommendation is purportedly in contrast to advice by [1] promoting the use of the Freeman-Tukey double arcsine transformation over the logit transformation. However, this publication only considered these transformations under the classic meta-analysis model. We agree with [1] that the use of the logit transformation is problematic in inverse variance meta-analyses with small event numbers or sample sizes; this is also visible in our example. These problems with the logit transformation under the classic meta-analysis do not translate to GLMMs. The classic meta-analysis model assumes that treatment estimates of individual studies follow a normal distribution which is obviously critical in studies with small numbers of events and observations. The arcsine and Freeman-Tukey double arcsine transformation are less affected by this normality assumption than the logit transformation. However, GLMMs taking into account the binomial structure of the data are not affected by this problem at all [10, 16].

5

4 Conclusions

Our case study shows that the Freeman-Tukey double arcsine transformation should only be used with special caution for the meta-analysis of single proportions due to potential problems in the back-transformation of meta-analysis results. In our view, a sensitivity analysis using other sample sizes is mandatory for this transformation. Generalized linear mixed models seem to be a promising alternative which is nowadays available in common meta-analysis software.

References

[1] Barendregt Jan J., Doi Suhail A., Lee Yong Yi, Norman Rosana E., Vos Theo. Meta-analysis of prevalence. *Journal of Epidemiology and Community Health.* 2013;67:974-8.

[2] Baxter A. J., Scott K. M., Vos T., Whiteford H. A.. Global prevalence of anxiety disorders: a systematic review and meta-regression. *Psychological Medicine.* 2013;43:897-910.

[3] Chemaitelly Hiam, Chaabna Karima, Abu-Raddad Laith J.. The Epidemiology of Hepatitis C Virus in the Fertile Crescent: Systematic Review and Meta-Analysis. *PLoS One.* 2015;10:e0135281.

[4] Greenaway Christina, Thu Ma Ann, Kloda Lorie A., et al. The Seroprevalence of Hepatitis C Antibodies in Immigrants and Refugees from Intermediate and High Endemic Countries: A Systematic Review and Meta-Analysis. *PLoS One.* 2015;10:e0141715.

[5] Borenstein Michael, Hedges Larry V., Higgins Julian PT, Rothstein Hannah R.. A basic introduction to fixed-effect and random-effects models for meta-analysis *Research Synthesis Methods.* 2010;1:97-111.

[6] Pettigrew Hugh M., Gart John J., Thomas Donald G.. The bias and higher cumulants of the logarithm of a binomial variate *Biometrika.* 1986;73:425–435.

[7] Berkson Joseph. Application of the Logistic Function to Bio-Assay *Journal of the American Statistical Association.* 1944;39:357-365.

[8] Anscombe FJ. The transformation of Poisson, Binomial and Negative-Binomial Data *Biometrika.* 1948;35:246–254.

[9] Freeman Murray F., Tukey John W.. Transformations related to the angular and the square root *Annals of Mathematical Statistics.* 1950;21:607–611.

[10] Stijnen Theo, Hamza Taye H., Ozdemir Pinar. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine.* 2010;29:3046-67.

[11] Miller John J.. The Inverse of the Freeman-Tukey Double Arcsine Transformation *The American Statistician.* 1978;32:138.

[12] Naveira M., Badal K.. The epidemiology of hepatitis C virus in Nepal Under preparation 2018.

[13] Schwarzer Guido. meta: An R package for Meta-Analysis *R News.* 2007;7:40–45.

[14] Agresti Alan, Coull Brent A.. Approximate is Better than Exact for Interval Estimation of Binomial Proportions *The American Statistician.* 1998;52:119–125.

[15] Newcombe Robert G.. Two-sided confidence intervals for the single proportion: comparison of seven methods *Statistics in Medicine.* 1998;17:857–872.

[16] Warton David I., Hui Francis K. C.. The arcsine is asinine: the analysis of proportions in ecology. *Ecology.* 2011;92:3-10.

[17] Hamza Taye H., Houwelingen Hans C., Stijnen Theo. The binomial distribution of meta-analysis was preferred to model within-study variability. *Journal of Clinical Epidemiology.* 2008;61:41-51.

[18] Jaeger T. Florian. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language.* 2008;59:434-446.

[19] Veroniki Areti Angeliki, Jackson Dan, Viechtbauer Wolfgang, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis *Research Synthesis Methods.* 2016;7:55-79.

8

TABLES

Table 1: Estimates and 95% confidence intervals of HCV prevalence meta-analyses using arcsine, Freeman-Tukey double arcsine, and logit transformations, respectively. Legend: GLMM (fixed) = logistic regression; GLMM (random) = random intercept logistic regression; between-study variance estimate $\hat{\tau}^2$.

Transformation (meta-analysis model)	Transformed proportion	HCV infections per 1 000 observations
Arcsine (fixed)	0.044 [0.042; 0.046]	1.94 [1.77; 2.13]
Double arcsine (fixed)	0.044 [0.042; 0.046]	0.00 [0.00; 0.00]
Logit (fixed)	-6.231 [-6.323; -6.139]	1.96 [1.79; 2.15]
GLMM (fixed)	-6.238 [-6.330; -6.147]	1.95 [1.78; 2.14]
Arcsine (random, $\hat{\tau} = 0.0003$)	0.044 [0.042; 0.046]	1.94 [1.76; 2.13]
Double arcsine (random, $\hat{\tau} = 0.0020$)	0.044 [0.041; 0.048]	0.00 [0.00; 0.00]
Logit (random, $\hat{\tau} = 1.1758$)	-5.451 [-6.649; -4.254]	4.27 [1.29; 14.01]
GLMM (random, $\hat{\tau} = 0.0000$)	-6.238 [-6.330; -6.147]	1.95 [1.78; 2.14]

FIGURE LEGENDS

Figure 1: Forest plot of HCV meta-analysis with Freeman-Tukey double arcsine transformation and without back-transformation of results. Abbreviation: PFT = Freeman-Tukey double arcsine transformed proportion.

Figure 2: Forest plot of HCV meta-analysis with Freeman-Tukey double arcsine transformation and back-transformation according to [11].

Figure 3: Influence of sample size on results of HCV meta-analysis using inverse of Freeman-Tukey double arcsine transformation according to [11].

Figure 4: Forest plot of HCV meta-analysis using classic method and logit transformation. Confidence intervals for individual studies are based on normal approximation for logit transformed proportions.

Figure 5: Forest plot of HCV meta-analysis using generalized linear mixed model. Confidence intervals for individual studies are based on Clopper-Pearson method [14, 15].

10

APPENDIX

A Statistical methods

We consider a meta-analysis of K studies where each study reports the number of events, a_k, and the number of observations n_k, $k = 1, \ldots, K$. We assume that the number of events follows a binomial distribution. Specifically, cell count $a_k \sim$ Binomial(n_k, p_k), where p_k denotes the probability of the event in study k. These probabilities are estimated from the observed number of events and sample sizes by $\hat{p}_k = a_k/n_k$.

A.1 Transformations

In this subsection we briefly introduce the arcsine, Freeman-Tukey double arcsine, and logit transformations in the context of a single study. In the next subsection, the use of these transformations in meta-analyses will be described.

A.1.1 Arcsine transformation

The arcsine-transformed event probability θ_k^{AS} [8] is defined as

$$\theta_k^{AS} = \arcsin \sqrt{p_k} \,.$$

An estimate of θ_k^{AS} is given by replacing p_k with $\hat{p}_k$. The main advantage of this transformation is the property of variance stabilization. The approximate variance of $\hat{\theta}_k^{AS}$ is calculated using

$$\widehat{\mathrm{Var}}\,(\hat{\theta}_k^{AS}) = \frac{1}{4\,n_k}$$

where the approximation improves as n_k increases. Notice that the approximate variance of $\hat{\theta}_k^{AS}$ only depends on the sample size. A confidence interval for θ_k^{AS} can

11

be constructed as

$$\hat{\theta}_k^{AS} \pm z_{1-\frac{\alpha}{2}} \text{ S.E.} (\hat{\theta}_k^{AS})$$

with standard error S.E. $(\hat{\theta}_k^{AS}) = \sqrt{\widehat{\text{Var}} (\hat{\theta}_k^{AS})}$ and $z_{1-\frac{\alpha}{2}}$ denoting the $1 - \frac{\alpha}{2}$ quantile of the standard normal distribution.

A.1.2 Freeman-Tukey double arcsine transformation

The Freeman-Tukey double arcsine-transformed event probability θ_k^{FT} [9] is an average of two arcsine-transformed probabilities. Its estimate is given by

$$\hat{\theta}_k^{FT} = 0.5 \left(\arcsin \sqrt{\frac{a_k}{n_k + 1}} + \arcsin \sqrt{\frac{a_k + 1}{n_k + 1}} \right)$$

The Freeman-Tukey double arcsine transformation was introduced in order to improve on the variance stabilizing property of the arcsine transformation. The approximate variance of $\hat{\theta}_k^{FT}$ is

$$\widehat{\text{Var}} (\hat{\theta}_k^{FT}) = \frac{1}{4 \, n_k + 2}$$

where the approximation – again – improves as n_k increases. A confidence interval for θ_k^{FT} can be constructed following the same methodology for that of the arcsine transformed probability described above.

A.1.3 Logit transformation

The logit transformation is another classic transformation [7] defined as

$$\theta_k^{LO} = \log \left(\frac{p_k}{1 - p_k} \right) .$$

Again, an estimate of θ_k^{LO} is given by replacing p_k with $\hat{p}_k$. The approximate variance of $\hat{\theta}_k^{LO}$ is

$$\widehat{\text{Var}} (\hat{\theta}_k^{LO}) = \frac{1}{a_k} + \frac{1}{n_k - a_k} .$$

12

It is clear from this variance formula that the approximate variance of a logit transformed proportion can become infinite if the number of events is zero or equal to the sample size. Typically, in this situation, a small increment is added to each denominator in order to yield a finite variance estimate.

A confidence interval for θ_k^{LO} can be constructed following the same methodology for that of the arcsine transformed probability described earlier.

A.2 Meta-analysis of single proportions

We briefly describe both the classic meta-analysis method assuming approximate normally distributed study effects (i.e., prevalence measures) as well as the generalized linear mixed model taking the binary structure of the data into account.

All methods are available in R function `metaprop()` from R package **meta** [13].

A.2.1 Classic random effects model

Classic fixed effect and random effects meta-analysis methods using the inverse variance method [5] can be implemented to combine single proportions. As the random effects model is a generalization of the fixed effect model, we only introduce the random effects model which is defined as

$$
\begin{aligned}
\hat{\theta}_k &= \theta_k + \epsilon_k, & \epsilon_k \overset{\text{i.i.d.}}{\sim} N(0, \sigma_k), \\
\theta_k &= \theta + u_k, & u_k \overset{\text{i.i.d.}}{\sim} N(0, \tau^2).
\end{aligned}
$$

where the ϵ's and u's are independent. This model contains two sources of variation: the within-study variances σ_k^2, $k = 1, \ldots, K$, and the between-study variance τ^2. The classic meta-analysis methods assume that the σ's are known, i.e., that the variances σ_k^2 are estimated without error by $\hat{\sigma}_k^2$. The estimated effects $\hat{\theta}_k$ and corresponding standard errors σ_k (which are assumed known) are used to estimate

13

τ^2 with the restricted maximum likelihood method [19]. Results are very similar using the classic DerSimonian and Laird estimator which is still the default in most statistical software for meta-analysis. The fixed effect model is a special case when $\tau^2 = 0$. Accordingly, results of fixed effect and random effects meta-analysis are identical if the estimate $\hat{\tau}^2$ equals zero.

Given estimates $(\hat{\theta}_k, \hat{\sigma}_k)$, the random effects estimate of θ, denoted by $\hat{\theta}_R$, is

$$\hat{\theta}_R = \frac{\sum\limits_{k=1}^{K} \hat{\theta}_k/(\hat{\sigma}_k^2 + \hat{\tau}^2)}{\sum\limits_{k=1}^{K} 1/(\hat{\sigma}_k^2 + \hat{\tau}^2)} \,,$$

which is a weighted average of the individual effect estimates $\hat{\theta}_k$ with weights $w_k = 1/(\hat{\sigma}_k^2 + \hat{\tau}^2)$.

The variance of $\hat{\theta}_R$ is estimated by

$$\widehat{\mathrm{Var}}\,(\hat{\theta}_R) = \frac{1}{\sum\limits_{k=1}^{K} w_k}$$

and a (1-α) confidence interval for $\hat{\theta}_R$ can be calculated using

$$\hat{\theta}_R \,\pm\, z_{1-\frac{\alpha}{2}}\, \mathrm{S.E.}\,(\hat{\theta}_R)$$

with standard error $\mathrm{S.E.}\,(\hat{\theta}_R) = \sqrt{\widehat{\mathrm{Var}}\,(\hat{\theta}_R)}$.

A fixed effect meta-analysis can be conducted by assuming a between-study variance $\tau^2 = 0$ resulting in a fixed effect estimate $\hat{\theta}_F$.

Instead of $\hat{\theta}_k$ and $\hat{\sigma}_k$, we use $\hat{\theta}_k^{AS}$ and $\mathrm{S.E.}\,(\hat{\theta}_k^{AS})$ for the arcsine method, $\hat{\theta}_k^{FT}$ and $\mathrm{S.E.}\,(\hat{\theta}_k^{FT})$ for the Freeman-Tukey double arcsine method, and $\hat{\theta}_k^{LO}$ and $\mathrm{S.E.}\,(\hat{\theta}_k^{LO})$ for the logit method. We denote the corresponding fixed effect and random effects estimates as $\hat{\theta}_F^{AS}$, $\hat{\theta}_R^{AS}$, $\hat{\theta}_F^{FT}$, $\hat{\theta}_R^{FT}$, $\hat{\theta}_F^{LO}$, and $\hat{\theta}_R^{LO}$, respectively.

14

A.2.2 Generalized linear mixed model

An excellent tutorial [10] describes how generalized linear mixed models can be utilized in the meta-analysis of event outcomes. One special case considered in the paper is the meta-analysis of single proportions which – like the classic meta-analysis model – assumes a normal distribution for the effect size (i.e., transformed proportion) across studies. However, a binomial distribution is assumed for the number of events within a study, i.e., $a_k \sim \text{Binomial}(n_k, p_k)$. Using the above defined logit transformed proportion θ_k^{LO}, this relation can be re-expressed in the following way to define the random effects model

$$
\begin{aligned}
a_k &\overset{\text{i.i.d.}}{\sim} \text{Binomial}\left(n_k, \frac{\exp(\theta_k^{LO})}{1 + \exp(\theta_k^{LO})}\right), \\
\theta_k^{LO} &\overset{\text{i.i.d.}}{\sim} \theta + u_k \quad u_k \overset{\text{i.i.d.}}{\sim} N(0, \tau^2).
\end{aligned}
$$

This model uses the binomial likelihood $\exp(\theta_k^{LO})^{a_k} / (1 + \exp(\theta_k^{LO})^{n_k}$ instead of the likelihood from the normal distribution [10] and is also known as a random intercept logistic regression model which implicitly uses the logit transformation. Accordingly, the GLMM estimates $\hat{\theta}_F^{GL}$ and $\hat{\theta}_R^{GL}$ correspond to the logit transformed probabilities in the fixed effect and random effects model, respectively.

Estimation of GLMMs for meta-analysis of single proportions is straightforward with R function `metaprop()` by specifying argument `method = "GLMM"`.

In principle, individual study weights could be derived from the likelihood contribution of each individual study, however, this information is at the moment not available in the utilized R software. Alternatively, the width of the Clopper-Pearson confidence intervals which also takes the binomial data structure into account [14, 15] could be used to get approximate study weights.

A.3   Back-transformations

For a single study, several statistical methods exist to calculate a confidence interval for a single proportion [14, 15]. These methods do not use the arcsine or the Freeman-Tukey double arcsine transformations and therefore, the back-transformation is not strictly relevant for individual study results. However, in a meta-analysis context, the back-transformation of the (double) arcsine as well as the logit transformation is essential to report results on the original scale, i.e., as proportions.

A.3.1   Arcsine back-transformation

The back-transformation / inverse of the arcsine transformation is defined as

$$p_k^{AS} = \sin(\theta_k^{AS})^2 \,.$$

This back-transformation can be used for a single study as well as the result of a meta-analysis, e.g., for the random effects estimate $\hat{\theta}_R^{AS}$ and its lower and upper confidence limits.

A.3.2   Inverse of Freeman-Tukey double arcsine transformation

Miller [11] introduced the back-transformation of the Freeman-Tukey double arcsine transformation which was published almost thirty years after the initial publication [9]. For study k, the back-transformation is defined as

$$p_k^{FT} = 0.5 \left(1 - \operatorname{sgn}(\cos(\theta_k^{FT}))\sqrt{1 - \left(\sin(2\,\theta_k^{FT}) + \left[\sin(2\,\theta_k^{FT}) - 1/\sin(2\,\theta_k^{FT})\right]/n_k\right)^2} \right) .$$

This rather complex back-transformation arises from using an average of two arcsine transformed proportions. The sample size n_k is included in the back-transformation which is no problem for a single study. However, in a meta-analysis with dif-

16

ferent sample sizes, a single sample size has to be specified to apply the back-transformation. Miller [11] suggested to use the harmonic mean of the sample sizes, i.e., $\tilde{n} = K \left/ \sum_{k=1}^{K} \frac{1}{n_k} \right.$. Accordingly, this harmonic mean $\tilde{n}$ and the meta-analysis estimate $\hat{\theta}_F^{FT}$ or $\hat{\theta}_R^{FT}$ are used in the back-transformation.

A.3.3 Inverse of logit transformation

The inverse of the logit transformation is defined as

$$p_k^{LO} = \frac{\exp(\theta_k^{LO})}{1 + \exp(\theta_k^{LO})} \,.$$

This well-known back-transformation can be used both for a single study and in a meta-analysis setting (classic method or GLMM).

Table A1: Definition and properties of prevalence transformations with number of events a and total sample size n.

Transformation	Estimate	Approximate variance	Comments
log [6]	$\log(a/n)$	$\dfrac{1}{a} - \dfrac{1}{n}$	Infinite estimate and variance for zero events
logit [7]	$\log\left(\dfrac{a/n}{1 - a/n}\right)$	$\dfrac{1}{a} + \dfrac{1}{n-a}$	Infinite estimate and variance for zero or all events
arcsine [8]	$\arcsin\sqrt{a/n}$	$\dfrac{1}{4\,n}$	Variance stabilizing; defined for zero events
Double arcsine [9]	$0.5\left(\arcsin\sqrt{a/(n+1)} + \arcsin\sqrt{(a+1)/(n+1)}\right)$	$\dfrac{1}{4\,n+2}$	Outperforms arcsine for small prevalences; sample size needed in back-transformation [11]

18

Table A2: Estimated number of HCV infections per 1 000 observations for additional sample sizes in fixed effect and random effects meta-analyses using the back-transformation of the Freeman-Tukey double arcsine method.

Sample size	HCV infections per 1 000 observations		Mean
	fixed effect	random effects	
85	0.000	0.000	harmonic
500	1.083	1.097	
1 000	1.486	1.500	
1 254	1.575	1.590	geometric
10 000	1.902	1.917	
46 892	1.941	1.956	arithmetic
100 000	1.947	1.962	
1 000 000	1.951	1.966	

19

SUPPLEMENTARY ONLINE MATERIAL

File: The code using R package **meta** is provided in the file `hcv.R`.

File: The HCV dataset is available in the file `hcv.csv`.

20

Study	HCV Infections	Total	Freeman–Tukey Double Arcsine Transformed Proportion	PFT	95%–CI	Weight (fixed)	Weight (random)
study 1	422	217154		0.044	[0.042; 0.046]	92.6%	78.3%
...dy 10	32	16557		0.044	[0.037; 0.052]	7.1%	20.5%
study 13	1	676		0.046	[0.009; 0.084]	0.3%	1.0%
...dy 18	1	44		0.181	[0.034; 0.328]	0.0%	0.1%
study 26	1	29		0.222	[0.042; 0.403]	0.0%	0.0%
Fixed effect model		234460		0.044	[0.042; 0.046]	100.0%	––
Random effects model				0.044	[0.041; 0.048]	––	100.0%

Heterogeneity: $I^2 = 44\%$, $\tau^2 < 0.0001$

0.025   0.05   0.075

forest1.eps

Study	HCV Infections	Total	HCV infections per 1000 observations	Events	95%–CI	Weight (fixed)	Weight (random)
study 1	422	217154		1.94	[1.76; 2.13]	92.6%	78.3%
study 10	32	16557		1.93	[1.32; 2.66]	7.1%	20.5%
study 13	1	676		1.48	[0.00; 6.34]	0.3%	1.0%
study 18	1	44		22.73	[0.00; 95.00]	0.0%	0.1%
study 26	1	29		34.48	[0.00; 142.11]	0.0%	0.0%
Fixed effect model		234460		0.00	[0.00; 0.00]	100.0%	––
Random effects model				0.00	[0.00; 0.00]	––	100.0%

Heterogeneity: $I^2 = 44\%$, $\tau^2 < 0.0001$

0   5   10   15

forest2.eps

Study	HCV Infections	Total	HCV infections per 1000 observations	Events	95%–CI	Weight (fixed)	Weight (random)
study 1	422	217154		1.94	[1.77; 2.14]	92.4%	27.0%
...dy 10	32	16557		1.93	[1.37; 2.73]	7.0%	26.4%
study 13	1	676		1.48	[0.21; 10.42]	0.2%	15.7%
...dy 18	1	44		22.73	[3.19; 144.48]	0.2%	15.5%
study 26	1	29		34.48	[4.84; 207.91]	0.2%	15.4%
Fixed effect model		**234460**		**1.96**	**[1.79; 2.15]**	**100.0%**	--
Random effects model				**4.27**	**[1.29; 14.01]**	--	**100.0%**

Heterogeneity: $I^2 = 72\%$, $\tau^2 = 1.3825$

0   5   10   15

forest3.eps

Study	HCV Infections	Total	HCV infections per 1000 observations	Events	95%-CI
study 1	422	217154		1.94	[1.76; 2.14]
study 10	32	16557		1.93	[1.32; 2.73]
study 13	1	676		1.48	[0.04; 8.21]
study 18	1	44		22.73	[0.58; 120.24]
study 26	1	29		34.48	[0.87; 177.64]
Fixed effect model		**234460**		**1.95**	**[1.78; 2.14]**
Random effects model				**1.95**	**[1.78; 2.14]**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$

0   5   10   15

forest4.eps

Harmonic
mean

Events per 1000 observations

Sample size used in back−transformation

| Double arcsine − fixed effect model |
| Double arcsine − random effects model |
| Arcsine − fixed effect model |
| Arcsine − random effects model |

samplesize.eps