ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

META-REGRESSION: ADJUSTING COVARIATES DURING META-ANALYSIS

Ravishankar. N, Sreekumaran Nair. N*

Ravishankar. N, PhD Scholar, Department of Statistics, Level 6, Health Sciences Library Building, Manipal University, Manipal.

Sreekumaran Nair. N, Director, Public Health Evidence South Asia (PHESA), Dr. TMA Pai Endowment Chair in Systematic Reviews and Evidence Based Public Health, Professor and Head, Department of Statistics, Level 6, Health Sciences Library Building, Manipal University, Manipal.

Manuscript Info

Manuscript History:

Received: 12 December 2014
Final Accepted: 22 January 2015
Published Online: February 2015

Key words:
Meta-regression, Meta-analysis, Covariates, Fixed effect model, Random effects model, Weighted linear regression, Weighted least squares

****Corresponding Author***

Prof. Sreekumaran Nair. N

Abstract

Meta-regression is a method to access impact of covariates on the effect estimates of studies to be meta-analyzed. It involves a weighted linear regression with the dependent variable being the effect estimate of studies and study level covariates being the independent variables. The weights are assigned inversely proportional to variance of the effect estimates. Pooled estimate adjusted for the covariate can be obtained by centering the covariate at its mean. In such cases, the estimate of intercept provides the pooled estimate adjusted for the covariate. The present paper elucidates the procedure of meta-regression involving a single covariate and guides the readers to perform meta-regression without the aid of any software packages.

1. Introduction

Meta-analysis is a statistical technique to integrate the results of several independent studies with a same effect measure. The end-product of meta-analysis is an overall estimate or a pooled estimate which is obtained as the weighted average of effect estimates of individual studies (Higgins and Green, 2008). Meta-regression is one of the popular methods to access impact of covariates on the effect estimate of studies to be meta-analyzed. It helps to assess the relationship between the dependent variable, which is the effect estimate of studies (log odds ratio, log risk ratio or mean difference) and one or more study level covariates (Borenstein et al., 2009; Dias et al., 2013; Thompson and Higgins, 2002). Further, the meta-regression can also be used to obtain the pooled estimate after adjusting for the effect of covariates (Sutton et al., 2009). This article demonstrates the technique of meta-regression and offers its detailed computational procedure in presence of a single covariate.

2. Methods

Just as in meta-analysis, there are two approaches to meta-regression - fixed effect model and random effects model. In a fixed effect model, it is assumed that all the studies share a common underlying true effect and consequently the observed variation in the effect estimates is due to chance alone. Whereas in a random effects model, the true effect underlying the studies is allowed to differ. Thus the random effects model accounts for the variability both within the studies and also between the studies (Borenstein et al., 2009).

3. Notations and Definitions

3.1. Meta-regression: Fixed effect

$Y_i = \beta o + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + \epsilon_i$, i = 1,2,3,........,n studies

Where, Y_i is effect estimate of i^{th} study, β's are the regression coefficients to be estimated and X_{ip} are the covariates for i^{th} study. Coefficients of the above model are estimated by "Weighted Least Squares" technique, defining weights by the reciprocal of variance of corresponding effect estimate i.e., $W_i = 1/V_i$. However standard error of the estimated regression coefficients needs to be subjected to a minor correction (Sutton et al., 2009; Thompson and Sharp, 1999).

Weighted least squares algorithm

In presence of a single covariate, the regression coefficients are obtained by minimizing $S = \sum_{i=1}^{n} W_i (Y_i - \beta_0 - \beta_1 X_i)^2$ ---- (1).

This gives;

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_i X_i Y_i - \sum_{i=1}^{n} W_i X_i \sum_{i=1}^{n} W_i Y_i}{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_i {X_i}^2 - (\sum_{i=1}^{n} W_i X_i)^2}$$

$$\widehat{\beta_0} = \overline{Y}_w - \widehat{\beta_1} \overline{X}_w$$

Where, $\overline{Y}_w = \frac{\sum_{i=1}^{n} W_i Y_i}{\sum_{i=1}^{n} W_i}$ and $\overline{X}_w = \frac{\sum_{i=1}^{n} W_i X_i}{\sum_{i=1}^{n} W_i}$

$\widehat{\beta_0}$ and $\widehat{\beta_1}$ are the estimators for β_0 and β_1 respectively. The expression for their corresponding variances is given by;

$\mathrm{Var}(\widehat{\beta_1}) = \frac{\widehat{\sigma^2}}{\sum_{i=1}^{n} W_i (X_i - \overline{X}_w)^2}$, where $\widehat{\sigma^2} = \frac{\sum_{i=1}^{n} W_i (Y_i - \hat{Y}_i)^2}{n-2}$ is the Mean Squared Error (MSE)

$$\mathrm{Var}(\widehat{\beta_0}) = \left[\frac{1}{\sum_{i=1}^{n} W_i} + \frac{(\overline{X}_w)^2}{\sum_{i=1}^{n} W_i (X_i - \overline{X}_w)^2}\right] \widehat{\sigma^2}$$

The standard error of $\widehat{\beta_0}$ and $\widehat{\beta_1}$ is obtained as $SE(\widehat{\beta_j}) = \sqrt{\mathrm{Var}(\widehat{\beta_j})}$, j= 0,1 and their corresponding adjustment is $S_j = SE(\widehat{\beta_j}) / \sqrt{MSE}$ (Sutton et al., 2009). The 95% confidence interval (CI) for the fixed effect meta-regression coefficients is given by $\widehat{\beta_j} - Z_{(1-\alpha/2)} S_j \le \beta_j \le \widehat{\beta_j} + Z_{(1-\alpha/2)} S_j$.

3.2. Meta-regression: Random effects

$Y_i = \beta o + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + U_i + \epsilon_i$

This model has two components in its error term U_i and ϵ_i, where U_i represents the between studies variability and ϵ_i represents the within study variability. Therefore the variance of Y_i is $V_i^* = var(U_i + \epsilon_i) = V_i + \tau^2$, where τ^2 is the between studies variability. In case of a single covariate, it is estimated as;

$\widehat{\tau^2} = \frac{Q-(k-1)}{F(W,X)}$ if Q > n – 2, or 0 otherwise

Q is the heterogeneity statistic, given by $Q = \sum_{i=1}^{n} W_i (Y_i - \widehat{\beta_0} - \widehat{\beta_1} X_i)^2$ and $\widehat{\beta_0}$ and $\widehat{\beta_1}$ are estimated from (1).

$$F(W,X) = \sum_{i=1}^{n} W_i - \frac{\sum_{i=1}^{n} {W_i}^2 \sum_{i=1}^{n} W_i X_i - 2 \sum_{i=1}^{n} {W_i}^2 X_i \sum_{i=1}^{n} W_i X_i + \sum_{i=1}^{n} W_i \sum_{i=1}^{n} {W_i}^2 {X_i}^2}{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_i {X_i}^2 - (\sum_{i=1}^{n} W_i X_i)^2}$$

Then a weighted linear regression is carried out with weights $W_i^* = 1/(V_i + \widehat{\tau^2})$ to provide new estimates of β_0 and β_1 (Thompson and Sharp, 1999). The standard errors of the estimated random effects meta-regression coefficients doesn't require any correction. Thus the 95% CI is given by $\widehat{\beta_j} - Z_{(1-\alpha/2)} SE_j \le \beta_j \le \widehat{\beta_j} + Z_{(1-\alpha/2)} SE_j$.

3.3. Testing the significance of estimated regression coefficients

The statistical significance of estimated coefficients of meta-regression is tested by Z statistic (Sutton et al., 2009) of the form;

$Z = \frac{\widehat{\beta_j}}{SE(\widehat{\beta_j})}$, j = 0,1

Under the null hypothesis that the coefficient is zero, Z would follow normal distribution with mean 0 and variance 1.

4. Data

Colditz et al examined the efficacy of BCG vaccine against tuberculosis (TB) in their article "Efficacy of BCG vaccine in the prevention of TB: Meta-analysis of the published literature" (Colditz et al., 1994). It was suggested that there is a relationship between the effect of BCG and the latitude of the area in which the trial was conducted. Data has been presented in table 1.

Table 1: data of the 13 trails conducted to examine the efficacy of BCG vaccine against TB along with corresponding latitudes of the places where the trials were conducted

Slno	Trial name	Authors	Latitude	Vaccination		No vaccination	
				Events	Total	Events	Total
1	Canada	Ferguson & Simes 1933	55	6	306	29	303
2	Northern USA	Aronson 1935	52	4	123	11	139
3	Northern USA	Stein & Aronson 1935	52	180	1541	372	1451
4	Chigago	Rosenthal et al 1937	42	17	1716	65	1665
5	Chigago	Rosenthal et al 1941	42	17	1716	65	1665
6	Georgia (school)	Comstock & Webster 1947	33	5	2498	3	2341
7	Puerto Rico	Comstock et al. 1949	18	186	50634	141	27338
8	UK	Hart & Sutherland 1950	53	62	13598	248	12867
9	Madanapalle	Frimont-Moller et al. 1950	13	33	5069	47	5808
10	Georgia (community)	Comstock et al. 1950	33	27	16913	29	17854
11	Haiti	Vandeviere et al. 1965	18	8	2545	10	629
12	South Africa	Coetzee & Berjak 1965	27	29	7499	45	7277
13	Madras	TB prevention trial 1968	13	505	88391	499	88391

*Effect measure: relative risk (RR)

5. Results

Meta-analysis of the data presented in Table 1 by inverse variance method with RR as the effect measure yielded the estimates as depicted in table 2.

Table 2: Estimates of meta-analysis

Type	Pooled estimate	95% CI	
		Lower	Upper
Fixed effect	0.64	0.59	0.69
Random effects	0.48	0.34	0.68

Meta-regression was performed considering ln(RR) as the dependent variable and latitude as the covariate. Tables 3 and 4 contain the results of fixed effect meta-regression and random effects meta-regression respectively.

Table 3: Fixed effect meta-regression

Ln(RR)	Regression β	Std.Error	Z	P value	95% CI Lower	Upper
Intercept	0.25	0.07	3.13	0.001	0.08	0.42
Latitude	-0.02	0.002	-10.00	<0.001	-0.03	-0.01

Table 4: Random effects meta-regression

Ln(RR)	Regression β	Std.Error	Z	P value	95% CI Lower	Upper
Intercept	0.07	0.37	0.19	0.85	-0.66	0.79
Latitude	-0.02	0.01	-2.00	0.04	-0.04	-0.004

Thus for a given latitude, the estimate of log risk ratio is obtained by the regression equations; ln(RR) = 0.25 – 0.02*(latitude) for fixed effects and ln(RR) = 0.07 – 0.02*(latitude) for random effects.
It is important to note that in both the cases; the regression coefficient of the latitude is statistically significant, indicating that latitude has a significant impact on the effect estimate of studies.

5.1. Pooled estimate adjusted for the covariate

The pooled estimate adjusted for covariate can be obtained by centering the covariate at its mean (subtracting the mean value of the covariate from the value of covariate of each study). However rest of the meta-regression procedure remains the same. In such a situation, the pooled estimate adjusted for covariate is obtained as the estimate of intercept (Sutton et al., 2009).
The results of this analysis have been presented in table 5 and 6.

Table 5: Pooled estimate adjusted for latitude – fixed effect

Coefficient	β	95% CI Lower	Upper
Intercept	-0.58	-0.67	-0.49
Latitude	-0.02	-0.03	-0.01

Therefore, ln(RR) = -0.58 – 0.02*(X_i – 34.69), i = 1,2,....,13. X_i is the value of latitude for the i^{th} study and 34.69 is the mean of the latitude values of all 13 studies. Exponential of the estimate of intercept gives the pooled estimate adjusted for the effect of latitude, which is 0.56 with 95% CI (0.51, 0.62).

Table 6: Pooled estimate adjusted for latitude – random effects

Coefficient	β	95% CI Lower	Upper
Intercept	-0.74	-1.05	-0.44
Latitude	-0.02	-0.04	-0.004

Thus, the adjusted pooled estimate for random effects model is 0.48 (0.35, 0.65).

5.2. A close comparison of the unadjusted pooled estimate and pooled estimate adjusted for the effect of covariate

Figures 1 and 2 depict the forest plot with the raw pooled estimate and adjusted pooled estimate.

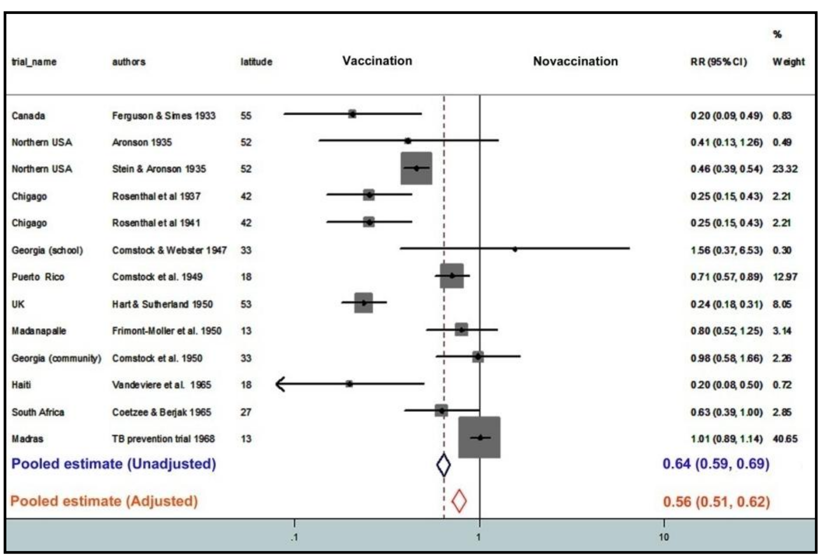

Figure 1: Forest plot – fixed effect

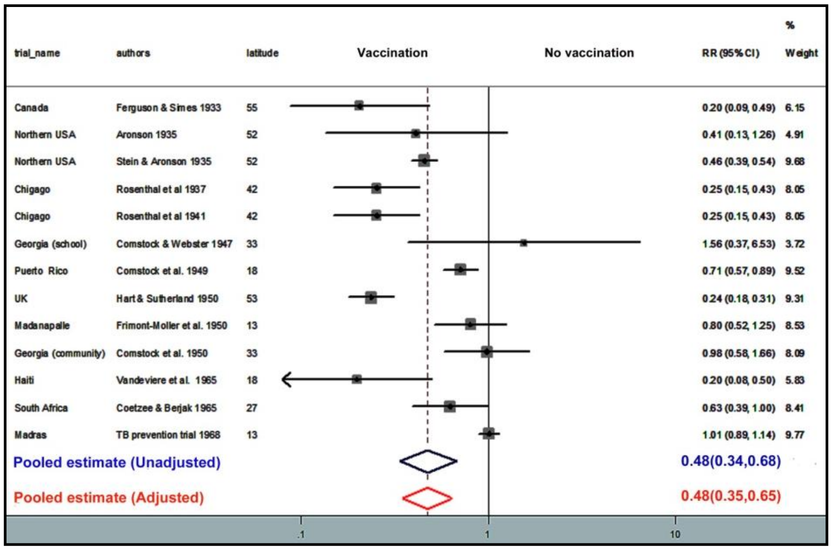

Figure 2: Forest plot – random effects

6. Discussion and Conclusion

Meta-regression is a technique used to explore the relationship between studylevel covariates and effect estimates of studies and further obtain the pooled estimate after adjusting for the covariates. It involves a weighted linear regression with weights assigned inversely proportional to the variance of effect estimates. The present article aims to provide a glimpse of the methodological details of meta-regression involving a single covariate. Both fixed effect and random effects meta-regression was performed using the data on efficacy of BCG vaccine against TB (Colditz et al., 1994) considering latitude as the covariate. The pooled estimate after adjusting for latitude was found to differ from the unadjusted estimate in case of fixed effect meta-regression, while in case of random effects meta-regression, both adjusted and unadjusted estimates were found to be the same; however there was a small shift in the confidence interval. This can be attributed to the fact that in random effects meta-analysis weights are designed to account for both within study and between studies heterogeneity (Dias et al., 2013). Meta-regression with multiple covariates is not recommended if the number of studies is small (Borenstein et al., 2009). As a thumb rule it is suggested to have atleast 10 studies for each covariate. The statistical packages estimate the meta-regression coefficients by the method of Restricted Maximum Likelihood (REML), which provides estimates with higher precision.

7. Acknowledgements

We acknowledge Public Health Evidence South Asia (PHESA) and Dr. TMA Pai Endowment Chair in Systematic Reviews and Evidence Based Public Health, Manipal University for supporting this work. We cordially thank Graham A Colditz and colleagues for their published data which has been utilized for the exercise.

8. Conflict of Interest

None

9. References

Borenstein, M., Hedges, L.V. and Higgins, J.P.T. (2009): Introduction to Meta-Analysis, John Wiley & Sons, Ltd, New York.

Colditz, G.A., Brewer, T.F., Berkey, C.S., Wilson, M.E., Burdick, E., Fineberg, H.V. and Mosteller, F. (1994): Efficacy of BCG Vaccine in the Prevention of Tuberculosis: Meta-analysis of the Published Literature. JAMA., 271:698-702.

Dias, S., Sutton, A.J., Welton, N.J. and Ades, A.E. (2013): Evidence synthesis for decision making 3: heterogeneity, subgroups, meta-regression, bias, and bias-adjustment. Med Des Making., 33:618-640.

Higgins, J.P.T. and Green, S. (2008): Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1, Wiley Blackwell Publications, England.

Sutton, A.J., Abrams., K.R., Jones, D.R., Sheldon, T.A. and Song, F. (2004): Methods for Meta-analysis in Medical Research, John Wiley & Sons, Ltd, England.
Thompson, S.G. and Higgins, J.P.T. (2002): How should meta-regression analyses be undertaken and interpreted? Stat Med., 21: 1559–1573.
Thompson, S.G. and Sharp, S.J. (1999): Explaining Heterogeneity in Meta-analysis: A Comparison of Methods, Stat Med., 18: 2693-2708.