R Package to Estimate Intracluster Correlation Coefficient with Confidence Interval for Binary Data

Hrishikesh Chakraborty, Akhtar Hossain

PII: S0169-2607(17)30577-1
DOI: 10.1016/j.cmpb.2017.10.023
Reference: COMM 4526

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 8 May 2017
Revised date: 21 September 2017
Accepted date: 28 October 2017

Please cite this article as: Hrishikesh Chakraborty, Akhtar Hossain, R Package to Estimate Intracluster Correlation Coefficient with Confidence Interval for Binary Data, *Computer Methods and Programs in Biomedicine* (2017), doi: 10.1016/j.cmpb.2017.10.023

R Package to Estimate Intracluster Correlation Coefficient with Confidence Interval for Binary Data

Hrishikesh Chakraborty[a,*], Akhtar Hossain[b]

[a]*Duke Clinical Research Institute, Duke University, Durham, NC, USA*
[b]*Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA*

Abstract

Background and Objective: The Intracluster Correlation Coefficient (ICC) is a major parameter of interest in cluster randomized trials that measures the degree to which responses within the same cluster are correlated. There are many different types of ICC estimators and its confidence intervals (CI) suggested in the literature for binary data. Studies have compared relative weaknesses and advantages of ICC estimators as well as its CI for binary data and suggested situations where one is advantageous in practical research. The commonly used statistical computing systems currently facilitate estimation of only a very few variants of ICC and its CI. To address the limitations of current statistical packages, we developed an R package, ICCbin, to facilitate estimating ICC and its CI for binary responses using different methods.

Methods: The `ICCbin` package is designed to provide estimates of ICC in 16 different ways including analysis of variance methods, moments based estimation, direct probabilistic methods, correlation based estimation, and resampling method. CI of ICC are estimating using 5 different methods. It also generates cluster binary data using exchangeable correlation structure.

Results: `ICCbin` package provides two functions for users. The function `rcbin()` generates cluster binary data and the function `iccbin()` estimates ICC and it's CI. The users can choose appropriate ICC and its CI estimate from the wide selection of estimates from the outputs.

Conclusions: The `R` package `ICCbin` presents very flexible and easy to use ways to generate cluster binary data and to estimate ICC and it's CI for binary response using different methods. The package `ICCbin` is freely available for use with `R` from the CRAN repository (https://cran.r-project.org/package=ICCbin).We believe that this package can be a very useful tool for researchers to design cluster randomized trials with binary outcome.

Keywords: Randomized Clinical Trials, Intracluster Correlation Coefficient, Confidence Interval of ICC, R Package, ICCbin

*Corresponding author.
E-mail address: rishi.c@duke.edu (Hrishikesh Chakraborty), mhossain@email.sc.edu (Akhtar Hossain)