Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 2, 2021 (New Series), pp 241 - 253

Measuring interchangeability in school lunch intervention data using concordance correlation coefficient

Gokul, T.[1], Srinivasan, M. R.[1] and Swaminathan, V. D.[2]
[1]*Department of Statistics, University of Madras, Chennai*
[2]*Department of Psychology, University of Madras, Chennai*

Received: 02 March 2020; Revised: 26 October 2020; Accepted: 30 December 2020

Abstract

In this paper, the measure of agreement is evaluated in terms of interchangeability among the set of treatment groups through the concordance correlation coefficient (CCC) for the longitudinal data using generalized linear mixed model (GLMM) and zero inflated model to account for the presence of zero observations in the data. CCC allow us to retain/interchange the treatment groups and provide scope for many experimental researches. Apart from CCC, this study also considers intra class correlation coefficient (CC), precision and accuracy for the evaluation. A simulation study is carried out to evaluate the performance of CCC followed by an application to a psychometric data that reveals there is no interchangeability in the nutritional supplements based on a school lunch intervention study.

Key words: Concordance correlation coefficient; Generalized linear mixed model; Poisson model; Zero inflated model.

1. Introduction

In many psychological studies, the measurement of variables could be ordinal, count or continuous using different methods. Further, in the case of longitudinal data, each subject is measured at different time points to obtain repeated measurements. Generally, measurements of agreement such as kappa measure in statistics are commonly used to identify the degree of concordance between the two or more observers. However, for the longitudinal data, there exists a necessity to assess the agreement between repeated measurements produced by a single observer or among multiple measurement methods.

The concordance correlation coefficient (CCC) proposed by Lin (1989, 1992) is one of the most widely applied procedures to assess agreement between observers on a quantitative scale by measuring the variation of linear relationship between each pair of data from a 45-degree line through the origin and it degenerates into kappa and weighted kappa for binary and ordinal data. The CCC was formerly defined as the Euclidean distance between paired data from two observers and the concordance line, and it was conveniently scaled to [-1, 1] interval, where -1 indicates perfect inverse agreement, 1 indicates perfect agreement, and value 0 is interpreted as complete disagreement. Further, each pair of measurements should fall on the 45-degree line, otherwise some disagreement is present in the data and hence that particular method becomes interchangeable. Traditionally, CCC is said to be a measure of total agreement and it can also be expressed as intra class correlation coefficient (CC) discussed in Carrasco and Jover (2003) and demonstrated its equivalence using mixed effect model by two methods namely variance components and moment method. Carrasco and

Corresponding Author: Srinivasan M.R.
E-mail: mrsrin8@gmail.com

Jover (2005) extended the CCC for measuring agreement with count data by means of intra class CC derived from a GLMM.

Literature is abundant in studying the measure of agreement in psychological studies (Barchard, 2012; Ma *et al.,* 2010) and for more than two observers in discrete data (King and Chinchilli, 2001; Carrasco and Jover, 2005; Carrasco *et al.,* 2009). Lin *et al.,* (2007) studied CCC based on the variance components under linear mixed model for quantitative/ qualitative data. Here, we have extended CCC to generalized version so that this can be applied to various kinds of data based on the generalized linear mixed model (GLMM) for the longitudinal and repeated measure data (Ge *et al.,* 2016). Carrasco (2010) proposed an index to measure the degree of agreement as the extended version of concordance correlation coefficient (CCC) through variance component (VC) approach for count data using GLMM framework. Moreover, CCC has been estimated with weights of a diagonal matrix between various repeated measurements over time for the longitudinal data. In the case of count data, Carrasco (2010) adopted a two-way GLMM with subject and observer specific random effects and random subject observer interaction effects. We extend this to a three-way GLMM and estimate the CCC for longitudinal count data. Thus, the aim of this work is to evaluate the impact of the CCC for count data with the presence of zeroes by means of a generalized expression in the CCC (GCCC) based on the intra class CC through GLMM approach count data. The idea behind the GCCC is first to fit the data using the most appropriate GLMM, and subsequently to develop the expression of the CCC based on the model parameters. Usually, the correlations have been studied for the identification of the relationship as strong or weak, but the methods/treatments are not interchangeable. In the context of CCC, it is expressed to improve the measure of agreement by interchanging the variables. CCC and intra class CC are the statistics that quantify the proportion of variance explained by a random factor in multilevel/hierarchical data.

The focus of this article is to explore the applicability of the measure of agreement CCC and intra class CC for each observer through GLMM for longitudinal data. Further, an extended three-way GLMM for longitudinal count data in the presence of zeros is considered to measure the agreement between the variables together with inter class CC, intra class CC, and total agreement. The paper is organized as follows. Section 2 introduces the dataset considered to examine the application of CCC. The existing methodology of generalized linear mixed model is reviewed together with the measurement of agreements CCC in Section 3. The results based on application of CCC for longitudinal data is discussed in Section 4. Section 5 provides the conclusion.

2. School Lunch Intervention Data

The cognitive data is a secondary data consisting of school lunch intervention given to children in rural Kenya (Neumann *et al.,* 2003). The intervention study is designed with three feeding groups of school children and also a control group who received no nutritional supplements. Each treatment group is comprised of 12 centres with children aged 6–14 years. The school lunch intervention was carried out in 9 out of 12 schools and students at the other three schools formed a control group. Data collected in Round 1 served as baseline before the intervention and called as pre-intervention scores. Round 2 was taken during the term after the intervention started and data in rounds 3, 4, and 5 were recorded during the second, fourth, and sixth terms after intervention started as post intervention scores. A total of 554 participants have been recorded including missing entries in the data. Data associated with 374 participants, excluding the missing observations is considered for this study: out of which

188 were boys and 186 were girls. Among the 374, 97 children were given calorie supplement, 127 children were given meat supplement, 78 were given milk and 72 were considered as control group in this study.

For the intervention study, recorded data was on general intelligence factor called G-factors (Raven's coloured progressive matrix) and other three S-factors (Verbal meaning, Arithmetic score and Digit span total), with nutritional supplements in order to study the measurement of agreement of interchangeability. Five repeated rounds of lunch intervention have been recorded from the schools and a summary of the data is presented in Table 1. From Table 1, it is evident that the there is an increasing trend in the overall mean for all the response variable considered. In addition, verbal meaning shows higher mean value since higher order cognitive function involving reasoning abilities on linguistic domain. The recorded data has been subjected to GLMM for count data and inflated GLMM for data with zeroes.

Table 1: Human intelligence - Overall mean for all response group

Time	RCPM	AS	VM	DS
Round 1	17.11	7.02	26.68	4.90
Round 2	17.47	7.16	27.27	5.37
Round 3	18.02	7.52	29.03	6.20
Round 4	18.41	8.04	32.03	6.97
Round 5	19.40	8.75	33.96	7.80

RCPM: Raven's Coloured Progressive Matrices; AS: Arithmetic Score; VM: Verbal Meaning; DS: Digit Span

3. Models and methods

Consider a study where n subjects are measured m times by J observers. Let Y_{ijkl} be the l^{th} reading $(l = 1,...,L)$ made by the j^{th} observer $(j = 1,...,J)$ at the k^{th} time $(k = 1,...,K)$ on the i^{th} subject $(i = 1,...,n)$. Since a sample of subjects is included in the data, the subject and the subject-by-observer interaction effects are assumed to be random effects. If the interest lies only in measuring the agreement among those observers included in the dataset as in many situations, then the same could be considered as observer-specific and are assumed as fixed effect. However, as mentioned by Carrasco and Jover (2003), when defining the agreement index, it is convenient to consider the observer effect as random to account for the systematic differences between observers as a source of disagreement. The index would otherwise, measure consistency rather than agreement. Thus, the present study, has considered variance components model as appropriate to fit the data with, subject-specific, subject-observers, subject–time interaction effects, $u_{ijk} = (\alpha_i, \alpha\beta_{ij}, \alpha\gamma_{ik})$ as sources of variability.

The random-effects vector u_{ijk} are independently distributed from an exponential family with mean $E(Y_{ijkl} | u_{ijk}) = \mu_{ijk}$ and variance $\text{var}(Y_{ijkl} | u_{ijk}) = v_{ijk} = \phi v(\mu_{ijk})$ where $v(\cdot)$ is a user specified variance function and ϕ is a unknown dispersion parameter. Through the link function $g(\mu_{ijk}) = \eta_{ijk}$, the conditional mean associated with a linear predictor is given by

$$\eta_{ijk} = x_{ijk}\beta^* + z_{ijk}u_{ijk}, \tag{1}$$

where x_{ijk} $(p \times 1)$ and z_{ijk} $(q \times 1)$ are independent variables with the fixed effects β^* and the random effects u_{ijk}.

We extend the three-way LMM proposed by Carrasco *et al.*, (2009) and Tsai (2017) to the GLMM incorporating repeated measurements rated by an observer at a certain time for each subject. The extended three-way GLMM can be written as

$$\eta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk}, \tag{2}$$

where μ is the overall mean, α_i is the subject-specific random effect assumed to be distributed as $\alpha_i \sim N(0,\sigma_\alpha^2)$, β_j is the observer-specific fixed effect, γ_k is the time-specific fixed effect, $\alpha\beta_{ij}$ is the random subject–observer interaction effect assumed to be distributed as $\alpha\beta_{ij} \sim N(0,\sigma_{\alpha\beta}^2)$, $\alpha\gamma_{ik}$ is the random subject–time interaction effect assumed to be distributed as $\alpha\gamma_{ik} \sim N(0,\sigma_{\alpha\gamma}^2)$, and $\beta\gamma_{jk}$ is the fixed observer–time interaction effect and all the three parameters are mutually independent. From the equation (1) and (2), fixed effect is expressed as $\beta^* = (\mu, \beta_1, ..., \beta_J, \gamma_1, ..., \gamma_K, \beta\gamma_{11}, ..., \beta\gamma_{JK})$ and random effect as $u_{ijk} = (\alpha_i, \alpha\beta_{ij}, \alpha\gamma_{ik})$, $u_{ijk} \sim MVN(0,G)$ where G is a diagonal matrix with elements $\sigma_\alpha^2, \sigma_{\alpha\beta}^2$ and $\sigma_{\alpha\gamma}^2$ on the diagonal and zero otherwise.

Following, Barnhart *et al.*, (2005) and Lin *et al.*, (2007), the total, intra class CC and inter class CC based on the GLMM approach can be expressed as follows

$$\rho_{CCC} = \frac{\text{cov}(Y_{ijkl}, Y_{ijk'l'})}{\text{var}(Y_{ijkl})} \tag{3}$$

where $\text{cov}(Y_{ijkl}, Y_{ijk'l'})$ and $\text{var}(Y_{ijkl})$ stand for the marginal covariance of the l^{th} reading $(l = 1,...,L)$ made by the j^{th} observer $(j = 1,...,J)$ at the k^{th} time $(k = 1,...,K)$ on the i^{th} subject $(i = 1,...,n)$. The marginal variance and covariance are developed as

$$\begin{aligned} \text{var}(Y_{ijkl}) &= \text{var}_u\{E(Y_{ijkl} \mid u_{ijk})\} + E_u\{\text{var}(Y_{ijkl} \mid u_{ijk})\} \\ &= \text{var}_u(\mu_{ijk}) + E_u\{\phi h(\mu_{ijk})\} \\ \text{cov}(Y_{ijkl}, Y_{ijk'l'}) &= \text{cov}_u\{E(Y_{ijkl} \mid u_{ijk}), E(Y_{ijk'l'} \mid u_{ijk'})\} + E_u\{\text{cov}(Y_{ijkl}, Y_{ijk'l'} \mid u_{ijk}, u_{ijk'})\}. \end{aligned}$$

Since Y_{ijkl} and $Y_{ijk'l'}$ are conditioned effects considered independent, the marginal covariance reduces to (McCulloch and Searle, 2001)

$$\begin{aligned} \text{cov}(Y_{ijkl}, Y_{ijk'l'}) &= \text{cov}_u\{E(Y_{ijkl} \mid u_{ijk}), E(Y_{ijk'l'} \mid u_{ijk'})\} \\ &= \text{cov}_u(\mu_{ijk}, \mu_{ijk'}). \end{aligned}$$

In addition, the CCC may also be divided into two components namely the precision and accuracy (Lin, 1989; Lin *et al.,* 2007) so that

$$\rho_{CCC} = \rho_p \cdot \chi_a \tag{4}$$

The precision component ρ_p, is the intra class CC considering the observers as fixed in equation (6), *i.e.,* not considering the between-observers variability as a source of disagreement. Additionally, the accuracy index χ_a, measures the distance between the observers' means, *i.e.,* the systematic differences among the observers. Thus, using the expressions of the marginal variance and covariance in the CCC equation (3), the following is obtained:

$$\rho_{GCCC} = \frac{\text{cov}_u(\mu_{ijk}, \mu_{ijk'})}{\text{var}_u(\mu_{ijk}) + E_u\{\phi h(\mu_{ijk})\}} \tag{5}$$

where ϕ is the dispersion parameter and $h(\cdot)$ the corresponding variance function associated random effects, which can be considered as a generalization of CCC (GCCC) to fit GLMM approach. Further, Lin *et al.,* (2007) defined the intra class CC as a measure of proportion of total variance attributable to the subjects and can be expressed as

$$\begin{aligned}\rho_{GCCC}^{\text{intra}} &= \left.\frac{\text{cov}_u(Y_{ijkl}, Y_{ijk'l'})}{\text{var}_u(Y_{ijkl})}\right|_{j(l,l')} \\ &= \frac{\text{var}_u(\mu_{ijk})|_{j(l,l')}}{\text{var}_u(\mu_{ijk})|_{j(l,l')} + E_u\{\phi h(\mu_{ijk})\}}\end{aligned} \tag{6}$$

where $j(l,l')$ be the reading measured m times based on j^{th} observer. Additionally, the conditional variance and covariance of $Y_{ijkl}, Y_{ijk'l'}$ given $j(l,l')$ is defined in equation (6). Furthermore, an intra class CC can also be defined as a measure of intra-observer agreement (Barnhart *et al.,* 2005; Lin *et al.,* 2007) where the observer effect is considered as fixed. This index should be interpreted, for each rater, as a measure of the proportion of the total variance (subjects plus error) attributable to subjects. It is also possible to define an inter class CC if the data have replicated readings $(m > 1)$ by considering the data as the average of those m readings

$$\begin{aligned}\rho_{GCCC}^{\text{inter}} &= \frac{\text{cov}(\bar{Y}_{ijk.}, \bar{Y}_{ijk'.})}{\text{var}\ (\bar{Y}_{ijk.})} \\ &= \frac{\text{cov}_u(\mu_{ijk}, \mu_{ijk'})|_{j(l,l')}}{\text{var}_u(\mu_{ijk}) + \dfrac{E_u\{\phi h(\mu_{ijk})\}}{m}}\end{aligned} \tag{7}$$

where $\bar{Y}_{ijk\cdot}$ stands for the average of m readings of the j^{th} observer on the i^{th} subject and along with k^{th} time point. Specifically, two models are considered for the count data based on the presence or absence of zero observations.

3.1. Poisson model

Let $Y \mid X, u$ follows a Poisson distribution and the conditional mean of Y_{ijkl} given u_{ijk} is $\mu_{ijk} = e^{\lambda_i + \alpha_i + \beta_j + \gamma_k}$ and the conditional variance is given by $\text{var}(Y_{ijkl} \mid u_{ijk}) = \mu_{ijk}$. The marginal expectation over the random effects is expressed in terms of its generating function and is given by

$$\begin{aligned} E_u(\mu_{ijk}) &= E_u(e^{\lambda_i + \alpha_i + \beta_j + \gamma_k}) = e^{\lambda} E_u(e^{\lambda_i + \alpha_i + \beta_j + \gamma_k}) \\ &= e^{\lambda} M_u(\alpha_i + \beta_j + \gamma_k), \end{aligned}$$

where $\lambda = \lambda_0 + \lambda_1 X_1 + ... + \lambda_p X_p$. Since, $u \sim MVN(0, G)$ under the assumption of random effects, the expected value can be obtained from $E_u(\mu_{ijk}) = e^{\lambda + \frac{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2}{2}}$. Thus, the specific CCC for the Poisson GLMM as explained by Carrasco (2010) becomes

$$\begin{aligned} \rho_{CCC} &= \frac{E_u(\mu_{ijk})^2 (e^{\sigma_\alpha^2} - 1)}{E_u(\mu_{ijk})^2 (e^{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2} - 1) + E_u(\mu_{ijk})} \\ &= \frac{E_u(\mu_{ijk})(e^{\sigma_\alpha^2} - 1)}{E_u(\mu_{ijk})(e^{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2} - 1) + 1} \end{aligned}$$

Further, CCC would reduce to

$$\rho_{CCC} = \frac{\mu(e^{\sigma_\alpha^2} - 1)}{\mu(e^{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2} - 1) + 1} \tag{8}$$

in the case of no subject-observer interaction. When $\sigma_\alpha^2 \to 0$ or $\sigma_\beta^2, \sigma_\gamma^2 \to \infty$, the CCC will tend to 0 (independence) and conversely, when $\sigma_\alpha^2 \to \infty$ or $\sigma_\beta^2, \sigma_\gamma^2 \to 0$, it will reach a value of 1 (perfect agreement) if $\sigma_\alpha^2 \rangle\rangle \sigma_\beta^2, \sigma_\gamma^2$. It is to be noted that the CCC is defined using the variance components, thus it cannot result in negative values.

3.2. Zero Inflated Poisson (ZIP) model

Let Y_{ijk} denote the longitudinal response for j^{th} observer $(j = 1, ..., J)$ at the k^{th} time $(k = 1, ..., K)$ on the i^{th} subject $(i = 1, ..., n)$. Then, the distribution of Y_{ijk} is expressed as

$$Y_{ijk} = \begin{cases} 0 & \text{with prob. } \phi_{ijk} \\ Poisson\ (\lambda_{ijk}) & \text{with prob. } 1 - \phi_{ijk} \end{cases} \tag{9}$$

where ϕ_{ijk} denotes the probability of the observation arising from the degenerated distribution at zero and λ_{ijk} represents the mean of the Poisson distribution. This formulation incorporates more zeros than permitted under the Poisson assumption (*i.e.*, where $\phi_{ijk}=0$). The probability distribution function of the longitudinal ZIP model can be written as

$$p(Y_{ijk}=0\,|\,Z_{ijk})=\phi_{ijk}+(1-\phi_{ijk})e^{-\lambda_{ijk}}$$
$$p(Y_{ijk}=0\,|\,X_{ijk})=(1-\phi_{ijk})\frac{\lambda_{ijk}^{y_{ijk}}e^{-\lambda_{ijk}}}{y_{ijk}!},\qquad y_{ijk}=1,2,... \tag{10}$$

where $0\le\phi_{ijk}\le 1$ and $0<\lambda_{ijk}<\infty$. Here X_{ijk} and Z_{ijk} can be mutually exclusive, partially, or completely overlapping to achieve modeling flexibility. This model has the same specification as that of the Poisson model but the variance function is expressed by $E_u\{\phi_{ijk}E_u(\mu_{ijk})\}=\phi_{ijk}E_u(\mu_{ijk})$ and the total CCC for ZIP model becomes

$$\rho_{CCC}=\begin{cases}\dfrac{E_u(\mu_{ijk})(e^{\sigma_\alpha^2}-1)}{E_u(\mu_{ijk})(e^{\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2}-1)+\phi_{ijk}} & \text{with prob. } \phi_{ijk}\\[2ex] \dfrac{E_u(\mu_{ijk})(e^{\sigma_\alpha^2}-1)}{E_u(\mu_{ijk})(e^{\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2}-1)+(1-\phi_{ijk})} & \text{with prob. } 1-\phi_{ijk}\end{cases} \tag{11}$$

To estimate the CCC and related quantities for the ZIP case, the data are fitted by penalized quasi-likelihood (PQL) through the SAS GLIMMIX and NLMIXED (Zhu *et al.*, 2015). The goal of the analysis is to determine the degree of agreement between each treatment group in order to decide if a new type of treatment could replace any other treatment. The interchangeability of treatment in this study will be considered only if the CCC value is at least 0.9 (Carrasco 2010). The following section describes the application of CCC for school lunch intervention study.

4. Simulation Study

The aim of the simulation study is to evaluate the impact of zero counts present in the dataset and assess the performance of CCC obtained from ZIP model. The parameter values of each combination were used as initial values to start the estimation process, with regard to the ML procedure. Carrasco (2010) studied to evaluate the impact of the over dispersion when estimating the CCC assuming a Poisson distribution. Similar, CCC estimate from a normal linear mixed model (Normal) whose behavior is related to the classical Lin's sample moment approach (Lin, 1989; Carrasco and Jover, 2003).

Following Carrasco (2010) simulation has been further extended to ZIP model and built three levels of agreement (as discussed in Section 1) through CCC values namely, below 20 as low, between 20 and 200 as medium and above 200 as high. Similarly, simulating the data into multiple zero values and classified as low (below 30), moderate (between 30 and 120), high (between 120 and 300) and extremely high (above 300) and repeated 1000 times. For each case, CCC was estimated using ZIP model using ML method of estimation and the simulation is performed in PROC GLIMMIX and PROC NLMIXED using SAS software. Table 2 explains the CCC values for the above mentioned scenarios.

Table 2: Results of the simulation scenarios

Multiple zeros	Agreement level	Comb.	Mean	CCC
Low	Low	1	11.65	0.2922
	Medium	2	21.06	0.3397
	High	3	29.78	0.4812
Moderate	Low	4	65.18	0.5461
	Medium	5	97.87	0.6424
	High	6	115.75	0.6997
High	Low	7	150.02	0.7495
	Medium	8	208.12	0.7948
	High	9	282.52	0.8542
Extremely	Low	10	565.14	0.8999
	Medium	11	729.45	0.9206
	High	12	1314.19	0.9958

The following are the findings from Table 2:

(i) The CCC values show increasing trend irrespective of the zero counts with the increasing levels.

(ii) In the case of extremely high zero counts, there is an improvement in the measure of agreement with an increasing rate of mean. Similarly, CCC increases as the mean increases for the case of medium and high zero counts in the data.

(iii) It is clear that the CCC value reaches 0.9 when the mean value of the data is extremely high irrespective of the zero counts and this provides an insight in interchangeability of the methods used for the study.

The simulation results paved the way for a better understanding of CCC with different categories of mean and zero counts in the data and motivates us to incorporate the same in the real time data set as explained in the following section.

5. Data Analysis

The school lunch intervention data introduced in Section 2 were considered for data analysis. To capture the measurement of agreement, various assumptions about the distribution of the random effects has been made. There is also baseline covariate information on each subject including age, gender, socio economic status, intake of nutritional supplements such as milk and meat, duration of the follow-up study. Furtther, measurement of G-factor in the analytical ability as assessed by Raven’s coloured progressive matrices and S-factor involving reasoning ability, linguistic ability and immediate memory and are assessed by Verbal meaning, Arithmetic scores and Digit span total respectively. Obviously, it is expected that improvement of children cognitive skills is correlated with the nutrition supplements and this association is studied using GLMM. We considered Poisson regression model and zero inflated Poisson model for analyzing the concept of interchangeability of nutritional supplements such as milk, meat, calories and control groups. From $u_{ijk} = (\alpha_i, \alpha\beta_{ij}, \alpha\gamma_{ik})$, a variance component model is used to fit the data including subjects-specific, subject-observers, subject–time interaction effects as sources of variability. Following Carrasco and Jover (2003), the observer effect is treated to be random since it accounts for the systematic difference between the observers by means of disagreement.

However, if the measures of agreement in CCC value is at least 0.9, then the particular treatment can be interchanged (Carrasco 2010).

The CCC is estimated using the variance components of a Poisson GLMM and the model is fitted by maximum likelihood (ML) using Gauss-Hermite quadrature through the NLMIXED SAS procedure. For each case, the CCC and its standard error are estimated. Table 3 gives the results of CCC on (8), precision, accuracy, intra class CC based on (3) to (7) for the Poisson model.

Table 3: Human intelligence - Results for Poisson model

	CCC	**Precision**	**Accuracy**	**Intra class CC Observer 1**	**Intra class CC Observer 2**	**Intra class CC Observer 3**
Analytical ability by Raven's coloured progressive matrices test						
Calorie	0.3926 (0.0298)	0.897 (0.0798)	0.4329 (0.0333)	0.897 (0.0798)	0.9029 (0.0805)	0.9085 (0.0810)
Meat	0.4089 (0.0299)	0.8951 (0.0786)	0.4587 (0.0352)	0.8951 (0.0786)	0.9011 (0.0802)	0.9068 (0.0808)
Milk	0.4356 (0.0327)	0.8979 (0.0798)	0.4931 (0.0398)	0.8979 (0.0798)	0.9037 (0.0806)	0.9093 (0.0812)
Control	0.3797 (0.0279)	0.8964 (0.0797)	0.4063 (0.0316)	0.8964 (0.0798)	0.9024 (0.0803)	0.908 (0.0809)
Numerical ability by Arithmetic score						
Calorie	0.3127 (0.0239)	0.6513 (0.0552)	0.3609 (0.0273)	0.6513 (0.0551)	0.6634 (0.0565)	0.6755 (0.0577)
Meat	0.3263 (0.0242)	0.6465 (0.0545)	0.3934 (0.0298)	0.6465 (0.0545)	0.6589 (0.0559)	0.6712 (0.0570)
Milk	0.3589 (0.0256)	0.6526 (0.0553)	0.4082 (0.0302)	0.6526 (0.0554)	0.6647 (0.0566)	0.6768 (0.0581)
Control	0.3004 (0.0213)	0.6477 (0.0546)	0.3589 (0.0255)	0.6477 (0.0546)	0.6602 (0.0561)	0.6725 (0.0572)
Linguistic ability by Verbal meaning						
Calorie	0.4871 (0.0383)	0.8298 (0.0729)	0.5329 (0.0432)	0.8298 (0.07287)	0.8484 (0.0746)	0.8653 (0.0768)
Meat	0.4936 (0.0399)	0.8234 (0.0723)	0.5412 (0.0447)	0.8234 (0.0723)	0.8426 (0.0736)	0.8601 (0.0764)
Milk	0.5031 (0.0403)	0.8329 (0.0731)	0.5532 (0.0452)	0.8329 (0.0731)	0.8511 (0.0752)	0.8678 (0.0775)
Control	0.4724 (0.0372)	0.829 (0.0728)	0.5216 (0.043)	0.8290 (0.0728)	0.8477 (0.0743)	0.8647 (0.0767)
Immediate memory by Digit span total						
Calorie	0.3264 (0.0226)	0.8422 (0.0743)	0.3721 (0.0277)	0.8422 (0.0743)	0.8567 (0.076)	0.8702 (0.0778)
Meat	0.3315 (0.0233)	0.8383 (0.0730)	0.3824 (0.0288)	0.8383 (0.073)	0.8532 (0.0753)	0.867 (0.077)
Milk	0.3561 (0.0256)	0.8452 (0.0744)	0.3987 (0.0293)	0.8452 (0.0242)	0.8594 (0.0762)	0.8727 (0.078)
Control	0.3129 (0.0213)	0.8387 (0.07308)	0.3621 (0.0262)	0.8387 (0.0248)	0.8537 (0.0757)	0.8674 (0.0773)

The following are the observations from Table 3:

(i) Concerning Raven's coloured progressive matrices with milk supplement produces a high degree of agreement (with CCC 0.4356) than other treatments but insufficient to declare that the treatment is interchangeable.

(ii) Based on the Koo and Li (2016) guidelines for interpreting the intra class CC be classified as poor (below 0.5), moderate (0.5 to 0.75) and excellent (above 0.9). In this study, the intra class CC was closer to one in all cases, so one could conclude that the methods adopted are reliable. However, we can see that there are few cases where the intra class CC falls between 0.6 and 0.8, but still are not closer to zero. Thus, we can say that these methods are reliable to fit under GLMM.

(iii) For the S factor through responses namely Arithmetic, Verbal meaning and Digit span, the treatment group milk produces a higher degree of agreement than other treatments.

Further, to handle the zero counts in the data, we fit the Zero inflated Poisson model on (9), the CCC behaves the same as in Poisson model yielding that there can be no interchangeability in treatments.

Table 4: Human intelligence - Results for Zero Inflated Poisson model

	CCC	Precision	Accuracy	Intra class CC Observer 1	Intra class CC Observer 2	Intra class CC Observer 3
Analytical ability by Raven's coloured progressive matrices test						
Calorie	0.3327 (0.0237)	0.9008 (0.0801)	0.3818 (0.0279)	0.9008 (0.0801)	0.8777 (0.0783)	0.8846 (0.0786)
Meat	0.3428 (0.0241)	0.8953 (0.0789)	0.3836 (0.028)	0.8953 (0.0791)	0.9062 (0.0807)	0.9117 (0.0812)
Milk	0.3538 (0.0256)	0.8914 (0.079)	0.3863 (0.0281)	0.8914 (0.0789)	0.9068 (0.0808)	0.9123 (0.0814)
Control	0.3288 (0.0232)	0.9 (0.08)	0.3806 (0.0278)	0.9 (0.0800)	0.8907 (0.07987)	0.8969 (0.0792)
Numerical ability by Arithmetic score						
Calorie	0.2684 (0.0152)	0.6554 (0.0556)	0.3017 (0.0205)	0.6554 (0.055)	0.6381 (0.0531)	0.6507 (0.0552)
Meat	0.2736 (0.0171)	0.6448 (0.0541)	0.3069 (0.021)	0.6448 (0.0541)	0.6711 (0.0568)	0.6831 (0.0582)
Milk	0.2883 (0.0189)	0.6452 (0.0546)	0.3082 (0.0214)	0.6452 (0.0545)	0.6615 (0.0562)	0.6737 (0.0573)
Control	0.2491 (0.0144)	0.6552 (0.0554)	0.3002 (0.0201)	0.6552 (0.0554)	0.6273 (0.0527)	0.64 (0.0536)
Linguistic ability by Verbal meaning						
Calorie	0.4186 (0.0313)	0.8346 (0.0726)	0.4949 (0.0397)	0.8346 (0.0725)	0.8091 (0.0705)	0.8294 (0.0715)
Meat	0.4318 (0.0325)	0.831 (0.0720)	0.4919 (0.0395)	0.831 (0.0719)	0.8368 (0.0729)	0.8548 (0.0758)
Milk	0.4536 (0.0352)	0.8218 (0.0712)	0.4962 (0.0399)	0.8218 (0.0712)	0.8441 (0.0739)	0.8613 (0.0765)
Control	0.4003 (0.0301)	0.876 (0.0781)	0.4938 (0.0396)	0.676 (0.0579)	0.8532 (0.07539)	0.8678 (0.0774)

	Immediate memory by Digit span total					
Calorie	0.2201 (0.0118)	0.8484 (0.0746)	0.3092 (0.0218)	0.8484 (0.0746)	0.804 (0.0702)	0.8214 (0.0709)
Meat	0.2239 (0.0126)	0.8382 (0.0735)	0.3131 (0.0223)	0.8382 (0.0732)	0.8534 (0.0755)	0.8671 (0.0771)
Milk	0.2282 (0.014)	0.8335 (0.0726)	0.3134 (0.0227)	0.8335 (0.0722)	0.8687 (0.0776)	0.8812 (0.0784)
Control	0.2164 (0.011)	0.7287 (0.0631)	0.2014 (0.0101)	0.7853 (0.0753)	0.8760 (0.0781)	0.8914 (0.0789)

The result in Table 4 based on (3) to (7) revealed that CCC on (11) is higher for the treatment milk in the entire response group namely Raven's coloured progressive matrices, Arithmetic score, Verbal meaning and Digit span total. This is significant because identifying the correct nutritional supplements in the development of cognitive function improves the intelligence of school children.

6. Concluding Remarks

The field of psychometric studies focus on developing a proper measure that can accurately summarize or give an idea of an individual's intellectual abilities and mental state. Longitudinal studies are common in many psychometric studies particularly on cognitive ability of school children and psychometric factors involving mental illness. Further, the policy makers are interested in identifying suitable interventions by providing nutritional supplements like milk, meat, pulses *etc.*, to enhance the intellectual abilities of the students. There is also a need to examine the interchangeability of various nutritional supplements provided to children.

Gokul *et al.*, (2021) proposed a joint model, based on GLMM approach, for Kenya school lunch intervention study and suggested that the nutritional supplements show gradual improvement in cognitive behavior among the students. However, the choice of nutritional supplements, also play an important and unique role in promoting children's growth and development. There have been arguments and counter arguments through various studies that nutritional supplements like meat, milk and calories provide suitable interventions in the intellectual abilities of children. It is in this direction the present study considers CCC approach as an appropriate measure to study the agreement or otherwise of various nutritional supplements in enhancing the mental abilities (Carrasco 2010).

In this study, the concept of interchangeability in the treatment groups through CCC, intra class CC, proposed by Tsai and Lin (2018) is adopted for analyzing the longitudinal school lunch intervention data. Further, in psychological studies, data are often of count or ordinal in nature involving more number of zeroes. Thus, we considered Poisson GLMM for count data, and inflated models in the presence of zero observations to capture the measure of agreement through the concordance correlation co-++efficient. The performance and applicability of the CCC has been first demonstrated with a simulation study followed by the Kenyan real time dataset from a psychological study. The results of CCC based on the real data suggests lesser degree of agreement for the interchangeability among the four considered treatment groups. The study has established in a limited way that the nutritional supplement of milk as an appropriate intervention for the growth of intellectual abilities among children. On the whole, the methodology provides an insight to researchers working on longitudinal data with zeros to derive the benefit of using CCC method based on GLMM as a suitable measure of agreement.

Acknowledgements

The authors are grateful to an anonymous referee for the constructive and insightful comments made on the manuscript which led to a highly improved version.

References

Barchard, K. A. (2012). Examining the reliability of interval level data using root mean square differences and concordance correlation coefficients. *Psychological Methods*, **17**(**2**), 294-308.

Barnhart, H. X., Song, J., and Haber, M. (2005). Assessing intra, inter, and total agreement with replicated measurements. *Statistics in Medicine*, **24**, 1371-1384.

Carrasco, J. L. (2010). A generalized concordance correlation coefficient based on the variance components generalized linear mixed models for over dispersed count data. *Biometrics*, **66**, 897-904.

Carrasco, J. L., and Jover, L. (2003). Estimating the generalized concordance correlation coefficient through variance components. *Biometrics*, **59**, 849-858.

Carrasco, J. L., and Jover, L. (2005). Concordance correlation coefficient applied to discrete data. *Statistics in Medicine*, **24**, 4021-4034.

Carrasco, J. L., King, T. S., and Chinchilli, V. M. (2009). The concordance correlation coefficient for repeated measures estimated by variance components. *Journal of Biopharmaceutical Statistics*, **19**, 90-105.

Ge, L., Tu, J. X., Zhang, H., Wang, H., He, H., and Gunzler, D. (2016). Modern methods for longitudinal data analysis, capabilities, caveats and cautions. *Shanghai Archives of Psychiatry*, **28(5)**, 293-300.

Gokul, T., Srinivasan, M. R., and Gallo, M. (2021). Modelling human intelligence using mixed model approach. *In: Mariani P., Zenga M. (eds) Data Science and Social Research II. Studies in Classification, Data Analysis, and Knowledge Organization*, 199-216.

King, T. S., and Chinchilli, V. M. (2001). A generalized concordance correlation coefficient for continuous and categorical data. *Statistics in Medicine*, **20**, 2131-2147.

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic*, **15**, 155-163.

Lin, L. I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, **45**, 255-268.

Lin, L. I.-K. (1992). Assay validation using the concordance correlation coefficient. *Biometrics*, **48**, 599-604.

Lin, L. I.-K., Hedayat, A. S., and Wu, W. A. (2007). Unified approach for assessing agreement for continuous and categorical data. *Journal of Biopharmaceutical Statistics*. **17**, 629-652.

Ma, Y., Tang, W., and Yu, Q. (2010). Modeling concordance correlation coefficient for longitudinal study data. *Psychometrika*, **75**(**1**), 99-119.

McCulloch, C. E., and Searle S. R. (2001). *Generalized, Linear and Mixed Models*. New York: Wiley.

Neumann, C. G., Bwibo, N. O., Murphy, S. P., Sigman, M., Whaley, S., Allen, L. H., Guthrie, D., Weisszz, R. E., and Demment, M. W. (2003). Animal source foods improve dietary quality, micronutrient status, growth and cognitive function in Kenyan school children: background, study design and baseline findings. *The Journal of Nutrition*, **133**, 3941S-3949S.

Tsai, M. Y., and Lin, C. C. (2018). Concordance correlation coefficients estimated by variance components for longitudinal normal and Poisson data. *Computational Statistics and Data Analysis*, **121**, 57-90.

Tsai, M. Y. (2017). Concordance correlation coefficients estimated by generalized estimating equations and variance components for longitudinal repeated measurements. *Statistics in Medicine*, **36**, 1319-1333.

Zhu, H., Luo, S., and DeSantis, S. M. (2015). Zero-inflated count models for longitudinal measurements with heterogeneous random effects. *Statistical Methods in Medical Research*, **26**(**4**), 1-16.