Published in final edited form as:

Biom J. 2015 November ; 57(6): 1084–1109. doi:10.1002/bimj.201400131.

Bivariate correlation coefficients in family-type clustered studies

Jingqin Luo^{1,*}, Gina D'Angelo¹, Feng Gao¹, Jimin Ding², and Chengjie Xiong¹

¹Division of Biostatistics, Washington University School of Medicine, 660 S. Euclid Avenue, Box 8067, St. Louis, MO 63110, USA

²Department of Mathematics and Statistics, Washington University, St. Louis, MO 63110, USA

Abstract

We propose a unified approach based on a bivariate linear mixed effects model to estimate three types of bivariate correlation coefficients (BCCs), as well as the associated variances between two quantitative variables in cross-sectional data from a family-type clustered design. These BCCs are defined at different levels of experimental units including clusters (e.g., families) and subjects within clusters and assess different aspects on the relationships between two variables. We study likelihood-based inferences for these BCCs, and provide easy implementation using standard software SAS. Unlike several existing BCC estimators in the literature on clustered data, our approach can seamlessly handle two major analytic challenges arising from a family-type clustered design: (1) many families may consist of only one single subject; (2) one of the paired measurements may be missing for some subjects. Hence, our approach maximizes the use of data from all subjects (even those missing one of the two variables to be correlated) from all families, regardless of family size. We also conduct extensive simulations to show that our estimators are superior to existing estimators in handling missing data or/and imbalanced family sizes and the proposed Wald test maintains good size and power for hypothesis testing. Finally, we analyze a real-world Alzheimer's disease dataset from a family clustered study to investigate the BCCs across different modalities of disease markers including cognitive tests, cerebrospinal fluid biomarkers, and neuroimaging biomarkers.

Keywords

Bivariate correlation; Bivariate linear mixed effects model; Missing data; Power; Random effect; Size; Wald test

1 Introduction

It is often of interest to make statistical inferences on bivariate correlation coefficients (BCCs)—the correlations between two quantitative variables or biomarkers collected through a family-type clustered design in biomedical and epidemiological studies. When

*Corresponding author: luor@wudosis.wustl.edu, Phone: +1-314-362-3718, Fax: +1-314-362-2693.

Additional supporting information including source code to reproduce the results may be found in the online version of this article at the publisher's web-site

Conflict of interest

The authors have declared no conflict of interest.

collected on independent subjects, the simple Pearson product-moment correlation coefficient (r_p) is usually calculated to indicate the bivariate correlation between two biomarkers when they are approximately normally distributed. In a family-type clustered design, however, data are collected on subjects within clusters, for instance, individuals from families, patients from hospitals, students from schools, thus paired observations of two variables measured on subjects within the same cluster are likely correlated at the subject level and the cluster level. Throughout the paper, we call such a study as a family-type clustered study while “cluster” and “family” are used exchangeably. The BCCs between two biomarkers can then be defined in multiple versions with respect to different levels of experimental units, yet they are often confusing to many clinicians whose primary interest is simply in estimating the correlation between different modalities (e.g., across different imaging techniques) of biomarkers. To clarify some of the confusions, we define various BCCs through a bivariate linear mixed effects (BLME) model and provide likelihood-based inferences to the BCCs as well as easy practical implementation in this paper.

The paper is organized as follows. We start by describing the exciting Dominantly Inherited Alzheimer Network (DIAN) study in Section 2 to motivate the need to delineating various BCCs and to point out two common data issues in family-type clustered studies. Section 3 introduces the concepts of our proposed BCCs, briefly reviews existing BCC estimators for clustered studies in the literature and makes contrasts between our BCCs and the alternatives. The BLME model setting in a family-type clustered design as well as definitions and interpretation and hypothesis testing on the BCCs rooted in the BLME are provided in Section 4. Section 5 presents the results from extensive simulation studies to illustrate the superior performance of the proposed BCCs relative to their alternatives across scenarios in presence of the two common data issues as observed in DIAN and to evaluate the size and power of the proposed Wald test. A subset of DIAN cohort was analyzed in Section 6 as a real-world example demonstration. We finally conclude the paper with discussion and summary.

2 The motivating DIAN study

Autosomal dominant Alzheimer’s disease (ADAD) has informed the field of Alzheimer’s disease (AD) research about the molecular and biochemical mechanisms that are believed to underlie the pathological basis of AD. The DIAN (U19 AG032438) was launched in 2008 to establish an international multicenter registry of individuals whose parents carry a known causative mutation of AD in the amyloid precursor protein (APP), presenilin 1 (PSEN1), or presenilin 2 (PSEN2) genes (Mills et al., 2013). Because of the rarity of these mutations, DIAN was designed as a family-type clustered study to enroll multiple family members per family to assure adequate statistical power on relevant scientific hypotheses of the study. The DIAN study collects biological fluids and evaluates participants at entry and longitudinally thereafter with clinical and cognitive batteries, structural, functional, metabolic, and amyloid imaging protocols, with the goal of determining the sequence of changes in presymptomatic gene carriers who are destined to develop AD. A comprehensive set of AD biomarkers, for example, amyloid deposition, cerebrospinal fluid (CSF) A β and tau, magnetic resonance imaging brain atrophy, and positron emission tomography (PET)

imaging with 2-[18F] fluoro-2-deoxy-D-glucose (FDG-PET), has been collected. Interim cross-sectional analyses indicate a cascade of AD biomarker changes that begin at least 20 years before symptomatic onset of disease (Bateman et al., 2012). Analysis of the DIAN database has reinforced the rationale that biomarker changes are ongoing in presymptomatic AD participants.

With the availability of a rich collection of biomarkers, one crucial scientific question in DIAN is to examine the correlations between different modalities of biomarkers (i.e., BCCs) during the preclinical stage when subjects are still cognitively normal at enrollment. In the cross-sectional analysis of DIAN participants at enrollment, a biomarker's measures in DIAN can be conceptualized as a combination of fixed effects and random effects. The fixed effects contain those from important covariates such as mutation status and the estimated age of onset (EAO). The random effects can be further conceptualized as a combination of the latent familial effects (cluster effects) and individuals' random errors (within-cluster subject effect). This conceptualization leads to different versions of BCCs between two biomarkers in DIAN, that is, at the family level or at the within-family subject level.

In a family-type clustered study such as DIAN, two data issues are commonly observed. The first is the "family size" issue, that is, families may contain varying numbers of subjects and many may contain only one single subject. Overall, the whole DIAN database (data freeze 4¹) contained a total of 276 individuals from 109 families: 19 families have two family members, 17 have three members, six have four members, and 11 have more than four members. Critically, 56 (51.37%) families contain only a single individual. The second is the "missing data" issue. For example, only 130 individuals have measurements for all 22 markers across various modalities including cognitive tests, CSF, PET PIB in DIAN. Among all potential pairs of the 22 markers, the missing proportion has a median of 28.6% with a minimum of 3.3% and a maximum of 39.9%. One hundred sixty-four subjects across 70 families have all CSF biomarkers and PET PIB imaging measures, 41 subjects across 29 families have CSF biomarkers measures but not PET PIB, and 43 subjects across 29 families have PET PIB measures but not CSF biomarker measures. The missing data mechanism is assumed to be missing at random since missing occurred mainly because some biomarkers were measured on subjects enrolled later but not earlier when a modality was not employed. The two common data issues pose analytic challenges that need to be dealt with in estimating BCCs.

3 The BCCs

Motivated by the DIAN study, the paper focuses on statistical inferences on BCCs in family-type clustered studies. We propose a BLME model to define and estimate different versions of BCCs between two biomarkers, $Y^{(1)}$ and $Y^{(2)}$, corresponding to different levels of experimental units in a family-type clustered design. The two biomarkers ($Y^{(1)}$ and $Y^{(2)}$) are jointly modeled as a bivariate normal vector determined by fixed effects (i.e., contributions

¹There are two data freezes per year of the DIAN data such that one year after each data freeze, the data are available to researchers to conduct projects and write manuscripts. Please refer to the DIAN document on Data and Tissue Sharing, Notifications, Publications, and Authorship Policies at the link below for details. <http://www.dian-info.org/resourcedb/PDFs/DIAN%20Data%20Sharing%20and%20Publications%20including%20DZNE%2012-3-13.pdf>

from important covariates) and random effects consisting of a bivariate latent familial random effect (cluster effect) and a bivariate within-family individual subject's random error. Naturally, this model structure leads to different versions of BCCs between $Y^{(1)}$ and $Y^{(2)}$, that is, at the family level or at the within-family subject level. The family-level BCC (labeled as r_b) is defined by the latent familial contribution to both biomarkers, mostly due to the shared family characteristics, that is, shared genetic traits and environmental conditions. The family-level BCC r_b can be interpreted as the association across families between the two biomarkers. A negative r_b indicates that an increase of the familial contribution to $Y^{(1)}$ is associated with a decrease of the familial contribution to $Y^{(2)}$ and vice versa. Within each family, $Y^{(1)}$ and $Y^{(2)}$ can be correlated at the individual subject level. Hence, the within-family subject-level BCC (labeled as r_w) can be defined as a conditional correlation between $Y^{(1)}$ and $Y^{(2)}$ when restricted to subjects from the same families. The subject-level BCC r_w thus measures the association of subjects' unique contributions to the biomarkers when the familial contribution is held as constant. A positive r_w indicates that for subjects from the same family, higher values of $Y^{(1)}$ are associated with higher values of $Y^{(2)}$ measured on the same subject and vice versa, regardless of the level of the familial effects on the biomarkers. Finally, when clinicians interpret the correlation between $Y^{(1)}$ and $Y^{(2)}$ collected through a family-type clustered design, they often refer to the overall correlation between the two that encompasses both familial contribution and subjects' unique contribution to the biomarkers' measures. This leads to the unconditional marginal correlation between $Y^{(1)}$ and $Y^{(2)}$ (labeled as r_o), using the overall joint distribution of the variables. Estimation and inference on the BCCs (r_b , r_w , r_o) will be affected, to varying extent, by data structures including number of families, number of subjects per family and level of imbalance between families (Bello et al., 2012). For instance, estimation and inference on the family-level BCC r_b will be limited if data are from a few number of families, the marginal BCC r_o will apparently be dominated by clusters with many subjects and in presence of many single-subject families, r_w will not be adequately estimated.

It is important to differentiate and quantify the three different types of BCCs to render a complete insight into the relationship between two variables, at what level, family or subject, are biomarkers correlated and to what extent. However, different types of clustered designs may emphasize on different BCCs. The family-type clustered design mostly focuses on r_w and the other types of clustered design may usually focus on r_b . A common malpractice summarizes the correlation between two variables by calculating one simple Pearson correlation coefficient (r_p) as if all subjects in a clustered study are independent. This ignores the existence of BCCs at different hierarchies, oversimplifies inferences on BCCs, and may lead to erroneous scientific conclusions.

3.1 Existing BCC estimators

Only a few BCCs between two quantitative variables in clustered designs have been studied in the literature (Bland and Altman, 1995a, 1995b; Thiébaud et al., 2002; Hamlett et al., 2004; Gao et al., 2006; Roy, 2006; Nguyen and Jiang, 2011), but were almost entirely restricted to the clustered design where subjects are repeatedly and completely measured. Clusters usually had the same size, each cluster size was required to have two or more subjects, and there was no missing data or the missing data issue was not considered. Bland

and Altman (1995a) analyzed the measurements of two variables on a set of subjects via the analysis of variance (AOV) method, where a linear model is fitted on one of the two variables to be correlated using the other as a covariate together with the “subject” variable as a fixed effect and the “correlation within subjects” is estimated as the square root of the ratio of the sum of squares (SS) attributable to the covariate variable over the total SS subtracting the SS attributable to the “subject” variable. We label this Bland and Altman estimator as r_w^{BA} in the paper. r_w^{BA} obviously ditches subjects with any missing values. As discussed in details in Bello et al. (2012), treating one of the two variables as a covariate neglects the fact that, like the response variable in the AOV, this variable is also measured with errors and can be influenced by other covariates. Furthermore, choosing to model “subject” as a fixed but not a random effect constrains generalizing inferences/conclusions to the entire population where the subjects are sampled from. Bland and Altman (1995b) also proposed to estimate the “correlation between subjects” as a cluster-size weighted version of the Pearson correlation coefficient using each subject’s averaged data (labeled as r_b^{BA}). Although r_b^{BA} takes into account of varying cluster sizes through weighting, it disregards subjects missing one of the variables. The associated variances of both BCCs were not originally derived in the references, although the p -value associated with the hypothesis testing $H_0: r_w^{BA} = 0$ was given by the F -test p -value accompanying the covariate variable in the AOV analysis and the p -value for r_b^{BA} was calculated in the reference as if testing a simple Pearson correlation coefficient. Nguyen and Jiang (2011) proposed for clustered designs a Pearson correlation alike “hidden correlation” estimator (labeled as r_b^H) without making stringent assumptions, for example, on the covariance structure of measurement errors and normality. However, r_b^H only uses clusters with two or more subjects with no missing values and thus clusters with only one single subject will not be used. Other literatures (Thiébaud et al., 2002; Hamlett et al., 2004; Gao et al., 2006; Roy, 2006) focused on analyzing real-world data taking advantage of the SAS PROC MIXED procedure using different covariance/correlation structure and attention was paid to the influence of various correlation structures.

In comparison with the existing BCC estimators, our approach builds upon the BLME and focuses on the family-type clustered design, not longitudinal or other types of clustered data. We point out the need of efficient BCC estimators in presence of the two common data issues: missing data issue (i.e., subjects missing one of the paired measurements of the two variables) and family size issue (i.e., families with varying size and many having only one single individual), as observed in the DIAN study. We emphasize on estimating all the three types of BCCs (r_b , r_w , r_o) simultaneously in a unified model to delineate the relationship between two variables at different hierarchies, and in doing so, we maximize the use of all data from all subjects (even those missing one of the two variables to be correlated) from all families, regardless of family size. Some existing BCC estimators, for example, r_b^H requires at least two subjects with complete paired measurements within each cluster and the meta-analysis approach (see Section 5) requires at least four subjects due to Fisher’s Z transformation while r_b^{BA} and r_w^{BA} will ditch observations with any missing values. As a result, clusters with only one single subject cannot be used in these exiting estimators in the

literature, although data from those clusters do provide information for estimating relevant variance/covariance parameters and ultimately for the correlations. Instead of being ditched, these data should be fully utilized for efficient estimations. Similarly, the missing data problem may be serious in studies when many subjects have one of the paired measurements missing. Some of the existing methods are constrained by only using subjects with complete pairs of measurements. Though not directly providing information on covariance/correlations, subjects with missing data do provide information on variance estimation and subsequently will indirectly lead to a more precise covariance/correlation estimation, and should be included in the estimation of the BCCs. Thus, approaches that can practically handle these two important data issues would be favored in real application, such as in the motivating DIAN study. The BLME model makes use of every bit of data and handles the two important data issues by borrowing information from other subjects in the same family and from the other families. Taking advantage of the linear mixed effects model implemented in most existing statistical software, it is easy to fit the BLME utilizing all available data (clusters with a single subject and subjects with missing measurements). To our best knowledge, the influence of the two practical issues on BCC estimations has not been previously evaluated in the literature. Through extensive simulations in Section 5, we compare the performance of the proposed BCC estimators to their alternatives in consideration of the two common data issues.

4 The BLME model

Let Y_{fj} be the paired measurements of two variables/biomarkers collected at a time point on

subject j in family (i.e., cluster) f , $Y_{fj} = \begin{pmatrix} Y_{fj}^{(1)} \\ Y_{fj}^{(2)} \end{pmatrix}$ with $f = 1, 2, \dots, F; j = 1, 2, \dots, n_f$ and

$\sum_{f=1}^F n_f = N$, a total of N individuals across F families. Let $\beta = \begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \end{pmatrix}$ be the fixed effects regression coefficients associated with a covariate vector, X_{fj} (known by design), which includes at least one corresponding to the intercept and a dummy variable indicating the two biomarkers (and other covariates if necessary). The BLME model can be written as,

$$Y_{fj} = \begin{pmatrix} Y_{fj}^{(1)} \\ Y_{fj}^{(2)} \end{pmatrix} = \begin{pmatrix} X_{fj}\beta^{(1)} \\ X_{fj}\beta^{(2)} \end{pmatrix} + \begin{pmatrix} \gamma_f^{(1)} \\ \gamma_f^{(2)} \end{pmatrix} + \begin{pmatrix} \varepsilon_{fj}^{(1)} \\ \varepsilon_{fj}^{(2)} \end{pmatrix} = \begin{pmatrix} X_{fj} \\ X_{fj} \end{pmatrix} \beta + \gamma_f + \varepsilon_{fj}. \quad (1)$$

$\gamma_f = \begin{pmatrix} \gamma_f^{(1)} \\ \gamma_f^{(2)} \end{pmatrix}$ denotes the bivariate “family” random effect, that is, the familial contribution to Y_{fj} that captures variations to account for heterogeneity in Y_{fj} across families. Sharing of $\gamma_f^{(i)}$ by $Y_{fj}^{(i)}$, $i = 1, 2$, in all subjects from the same cluster f induces a correlation in the measurements of biomarker i for subjects within a cluster. The “family” random effect γ_f follows a zero mean bivariate normal distribution with a covariance matrix Σ_F ,

$$\begin{pmatrix} \gamma_f^{(1)} \\ \gamma_f^{(2)} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_F = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \right),$$

where $g_{12} = g_{21}$ is the covariance between the familial effects on the two biomarkers' measures. γ_f is independent of $\gamma_{f'}$ for $f \neq f'$. Therefore, the family-level BCC r_b measuring the familial association between the two biomarkers is defined as ("Corr" denoting correlation),

$$r_b = \text{Corr}(\gamma_f^{(1)}, \gamma_f^{(2)}) = \frac{g_{12}}{\sqrt{g_{11} \times g_{22}}}. \quad (2)$$

The error term $\varepsilon_f = \begin{pmatrix} \varepsilon_{fj}^{(1)} \\ \varepsilon_{fj}^{(2)} \end{pmatrix}$ is independent of γ_f and follows a zero mean bivariate normal distribution with a covariance matrix Σ_ε

$$\begin{pmatrix} \varepsilon_{fj}^{(1)} \\ \varepsilon_{fj}^{(2)} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_\varepsilon = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \right).$$

where $s_{12}(=s_{21})$ is the covariance term between the paired measurements on the same subjects. $\varepsilon_{fj}^{(1)}$ and $\varepsilon_{f'k}^{(2)}$, $j \neq k$, are independent of each other given γ_f and $\text{Cov}(\varepsilon_{fj}^{(1)}, \varepsilon_{f'k}^{(2)}) = 0$ across different families ($f \neq f'$). Hence, Σ_ε is the conditional covariance matrix of $\mathbf{Y}_{fj}$ when conditioning on the family random effect γ_f , that is,

$$\text{Var}(\mathbf{Y}_{fj} | \gamma_f) = \Sigma_\varepsilon.$$

Thus, conditional on family effect, the within-cluster BCC r_w can be defined as,

$$r_w = \text{Corr}(Y_{fj}^{(1)}, Y_{fj}^{(2)} | \gamma_f) = \frac{s_{12}}{\sqrt{s_{11} \times s_{22}}}. \quad (3)$$

The marginal covariance matrix of $\mathbf{Y}_{fj}$ encompasses both the within-cluster and the between-cluster variance/covariance,

$$\text{Var}(\mathbf{Y}_{fj}) = \Sigma_F + \Sigma_\varepsilon = \begin{pmatrix} g_{11} + s_{11} & g_{12} + s_{12} \\ g_{21} + s_{21} & g_{22} + s_{22} \end{pmatrix}.$$

Thus, the overall correlation between two biomarkers can be described by the marginal BCC r_o , which is defined as,

$$r_o = \text{Corr}(Y_{fj}^{(1)}, Y_{fj}^{(2)}) = \frac{g_{12} + s_{12}}{\sqrt{g_{11} + s_{11}} \sqrt{g_{22} + s_{22}}}. \quad (4)$$

The regression coefficients β and all parameters in the covariance matrices Σ_F and Σ_ε (and the defined BCCs) are unknown and will be estimated from fitting the BLME.

4.1 Relationships among the BCCs

To see the relationship among different versions of the BCCs, let $\bar{\mathbf{Y}}_f = \begin{pmatrix} \bar{Y}_f^{(1)} \\ \bar{Y}_f^{(2)} \end{pmatrix}$ represent the average of the two variables in family f with a family size of n_f , the variance/covariance matrix of $\bar{\mathbf{Y}}_f$ is (see Supporting Information Appendix SA.1),

$$\text{Var}(\bar{\mathbf{Y}}_f) = E \left[(\bar{\mathbf{Y}}_f - E \bar{\mathbf{Y}}_f)(\bar{\mathbf{Y}}_f - E \bar{\mathbf{Y}}_f)' \right] = \Sigma_F + \frac{1}{n_f} \Sigma_\varepsilon,$$

with E denoting expectation and $'$ denoting vector transpose. Thus, the BCC between the paired averages at the cluster level (labeled as r_b^*) can be readily calculated as,

$$r_b^* = \text{Corr}(\bar{Y}_f^{(1)}, \bar{Y}_f^{(2)}) = \frac{g_{12} + \frac{s_{12}}{n_f}}{\sqrt{\left(g_{11} + \frac{s_{11}}{n_f}\right) \times \left(g_{22} + \frac{s_{22}}{n_f}\right)}}. \quad (5)$$

The Bland and Altman estimator r_b^{BA} (Bland and Altman, 1995b) is an estimator of r_b^* . It is clear that if the cluster size is very large, that is, as n_f goes to infinity for all f , our version of cluster-level BCC r_b (2) can be interpreted as the limit of r_b^* . On the other extreme situation, if there is only one family member sampled from each family cluster, that is, $n_f = 1$ and all subjects are independent, then the family effect and the individual subject effect are completely confounded, and the three versions of BCCs reduce to one single quantity. Finally, if both variances from the familial effects, g_{11} and g_{22} , are small and close to 0, that is, the variation of the familial contribution to the biomarker measures is minimal, then g_{12} is close to 0 too, and hence, our version of the within-family subject-level (i.e., conditional on family effect) BCC r_w (3) can be interpreted as the limit of r_b^* when the familial variances go to 0.

To see the importance to differentiate the various BCCs from the simple Pearson correlation coefficient r_p in family-type clustered studies, we compute a simple naive Pearson correlation ignoring the cluster structure in the data, and examine how it is related to the BCCs we have defined. Consider the ideal situation when all families have the same cluster size, that is, $n_f = n$ and there is no missing data from subjects within each family. The equation below gives the naive sample Pearson correlation coefficient in a family-type clustered design ignoring the cluster structure:

$$\hat{r}_p = \frac{\sum_{f=1}^F \sum_{j=1}^n (Y_{fj}^{(1)} - \bar{Y}^{(1)})(Y_{fj}^{(2)} - \bar{Y}^{(2)})}{\sqrt{\left[\sum_{f=1}^F \sum_{j=1}^n (Y_{fj}^{(1)} - \bar{Y}^{(1)})^2 \right] \left[\sum_{f=1}^F \sum_{j=1}^n (Y_{fj}^{(2)} - \bar{Y}^{(2)})^2 \right]}}, \quad (6)$$

where the sample mean of the variable i in family f is, $\bar{Y}^{(i)} = \frac{\sum_{f=1}^F \sum_{j=1}^n Y_{fj}^{(i)}}{n \times F}$, $i = 1, 2$. According to the BLME assuming no covariates, it can be derived (see Supporting Information Appendix SA.2) that,

$$E \left[\sum_{f=1}^F \sum_{j=1}^n (Y_{fj}^{(1)} - \bar{Y}^{(1)})(Y_{fj}^{(2)} - \bar{Y}^{(2)}) \right] = n \times (F-1) \times g_{12} + (n \times F - 1) \times s_{12}$$

and

$$E \left[\sum_{f=1}^F \sum_{j=1}^n (Y_{fj}^{(i)} - \bar{Y}^{(i)})^2 \right] = n \times (F-1) \times g_{ii} + (n \times F - 1) \times s_{ii}, \text{ for } i=1, 2.$$

Therefore, the sample estimator $\hat{r}_p(6)$ may be used as an estimator, though not necessarily unbiased, of the population quantity below in the BLME context,

$$r_p = \frac{n \times (F-1) \times g_{12} + (n \times F - 1) \times s_{12}}{\sqrt{n \times (F-1) \times g_{11} + (n \times F - 1) \times s_{11}} \sqrt{n \times (F-1) \times g_{22} + (n \times F - 1) \times s_{22}}}. \quad (7)$$

In special situations when $n = 1$ and $F > 1$ (i.e., all subjects are independent) and $F = 1$ and $n > 1$ (i.e., all subjects are from one family), $r_p(7)$ becomes equivalent to $r_o(4)$ and $r_w(3)$, respectively. In other cases, r_p is a function of both the total number of clusters (F) and the cluster size within each family (n). If F is fixed and n goes to infinity (i.e., with very large

families), r_p approaches to $\frac{(F-1) \times g_{12} + F \times s_{12}}{[\sqrt{(F-1) \times g_{11} + F \times s_{11}}] [\sqrt{(F-1) \times g_{22} + F \times s_{22}}]}$, a function of the number of families and hence could not be reasonably conceptualized as a correlation that measures the association of two biomarkers across the population of families and/or the population of subjects within families. On the other hand, if n is fixed and F goes to infinity, r_p approaches to the marginal BCC $r_o(4)$. Thus, the various BCCs we have defined in the paper reflect natural extensions of the classical Pearson correlation in the more complex family-type clustered study designs.

4.2 Estimators and associated variances

All parameters in the BLME (1) can be efficiently estimated by the restricted maximum likelihood (REML) method. Organize all the variance/covariance parameters in the BLME into a vector, $\boldsymbol{\varnothing} = (g_{11}, g_{12}, g_{22}, s_{11}, s_{12}, s_{22})$. Denote the REML estimator by $\hat{\boldsymbol{\varnothing}} = (\hat{g}_{11}, \hat{g}_{12},$

$\hat{g}_{22}, \hat{s}_{11}, \hat{s}_{12}, \hat{s}_{22}$) and $\hat{\varnothing}$ can be obtained by fitting the proposed BLME in standard software packages such as SAS and R. The estimators to each proposed BCCs (similarly to r_b^*) can be subsequently estimated by replacing each parameter by its REML estimator,

$$\hat{r}_b = \frac{\hat{g}_{12}}{\sqrt{\hat{g}_{11} \times \hat{g}_{22}}}, \hat{r}_w = \frac{\hat{s}_{12}}{\sqrt{\hat{s}_{11} \times \hat{s}_{22}}}, \hat{r}_o = \frac{\hat{g}_{12} + \hat{s}_{12}}{\sqrt{\hat{g}_{11} + \hat{s}_{11}} \sqrt{\hat{g}_{22} + \hat{s}_{22}}}.$$

We can derive the variance on the estimators of the three BCCs via the multivariate delta method,

$$\text{Var}(\hat{r}_*) = \left(\frac{\partial r_*}{\partial \varnothing} \right)' \text{CoVar}(\hat{\varnothing}) \frac{\partial r_*}{\partial \varnothing},$$

with $\frac{\partial r_*}{\partial \varnothing}$ representing the vector of partial derivatives of a BCC r_* (denoting any of the BCCs including r_b, r_w, r_o , and r_b^*) with respect to each element in $\varnothing$ and $\text{CoVar}(\hat{\varnothing})$ denoting the asymptotic estimation on the covariance matrix of $\hat{\varnothing}$. More specifically, the partial derivatives of a BCC with each relevant parameter can be derived as follows (see Supporting Information Appendix SA.3 for derivations on r_b^*):

$$\frac{\partial r_w}{\partial \varnothing} = \left(0, 0, 0, \frac{\partial r_w}{\partial s_{11}}, \frac{\partial r_w}{\partial s_{12}}, \frac{\partial r_w}{\partial s_{22}} \right),$$

$$\text{with } \frac{\partial r_w}{\partial s_{11}} = -0.5 \frac{s_{12}}{\sqrt{s_{11}^3 s_{22}}}, \frac{\partial r_w}{\partial s_{12}} = \frac{1}{\sqrt{s_{11} s_{22}}}, \frac{\partial r_w}{\partial s_{22}} = -0.5 \frac{s_{12}}{\sqrt{s_{11} s_{22}^3}}.$$

For $\frac{\partial r_b}{\partial \varnothing} = \left(\frac{\partial r_b}{\partial g_{11}}, \frac{\partial r_b}{\partial g_{12}}, \frac{\partial r_b}{\partial g_{22}}, 0, 0, 0 \right)$, the three associated partial derivatives $\frac{\partial r_b}{\partial g_{ij}} (i, j=1, 2)$ have exactly the same expressions as the above three partial derivatives of r_w with respect to $s_{ij} (i, j=1, 2)$, thus we only need to substitute $r_w, s_{11}, s_{12}, s_{22}$ simultaneously by $r_b, g_{11}, g_{12}, g_{22}$ correspondingly to obtain $\frac{\partial r_b}{\partial \varnothing}$. For $\frac{\partial r_o}{\partial \varnothing} = \left(\frac{\partial r_o}{\partial g_{11}}, \frac{\partial r_o}{\partial g_{12}}, \frac{\partial r_o}{\partial g_{22}}, \frac{\partial r_o}{\partial s_{11}}, \frac{\partial r_o}{\partial s_{12}}, \frac{\partial r_o}{\partial s_{22}} \right)$, we have

$$\frac{\partial r_o}{\partial g_{11}} = \frac{\partial r_o}{\partial s_{11}} = -0.5 \frac{g_{12} + s_{12}}{\sqrt{g_{22} + s_{22}} \sqrt{(g_{11} + s_{11})^3}},$$

$$\frac{\partial r_o}{\partial g_{12}} = \frac{\partial r_o}{\partial s_{12}} = \frac{1}{\sqrt{g_{11} + s_{11}} \sqrt{g_{22} + s_{22}}},$$

$$\frac{\partial r_o}{\partial g_{22}} = \frac{\partial r_o}{\partial s_{22}} = -0.5 \frac{g_{12} + s_{12}}{\sqrt{g_{11} + s_{11}} \sqrt{(g_{22} + s_{22})^3}}.$$

The estimation on $\text{Var}(r_*)$ will be subsequently obtained by plugging in the REML estimates of the parameters into the above relevant expressions.

4.3 Hypothesis testing

With a BCC and its associated variance estimated, the Wald test statistic $z = \frac{\hat{r}_* - c}{\sqrt{\text{Var}(\hat{r}_*)}}$ (r_* denoting any of the defined BCCs) can be constructed to test the null hypothesis, $H_0: r_* = c$, where c is a given constant. The Wald test statistic asymptotically follows a standard normal distribution under the null hypothesis. The associated $100(1 - \alpha)\%$ Wald confidence interval (CI) on r_* can be calculated as $(\hat{r}_* - z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{r}_*)}, \hat{r}_* + z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{r}_*)})$, where $z_{1-\frac{\alpha}{2}}$ is

the corresponding $1 - \frac{\alpha}{2}$ critical value of cumulative standard normal distribution. Fisher's Z transformation can be implemented on an estimator: $\hat{r}_*^z = 0.5 \times \log \frac{1+\hat{r}_*}{1-\hat{r}_*}$ and $\text{Var}(\hat{r}_*^z)$ can be easily derived via delta method to conduct the Wald test in Fisher's Z scale on the null

hypothesis $H_0: r_*^z = c^z$ is Fisher's Z transformation of the constant c , $c^z = 0.5 \times \log \frac{1+c}{1-c}$. Asymptotically, the Wald test is equivalent to its two alternatives, the likelihood ratio test and the score test. We will investigate the performance of the proposed Wald test in the original scale and Fisher's Z scale in the simulation studies of Section 5.2, particularly in terms of small sample sizes.

4.4 Practical implementation

We have pointed out the two major analytic challenges (the imbalanced family size issue, especially when many families have only one single individual and the missing data issue where one of the paired variable is missing) in correlating two biomarkers in real-world family-type clustered studies such as DIAN. Existing BCC estimators usually handle data with multiple subjects per family and no missing values. In presence of many single-subject families, existing approaches usually kick out those families. Deletion of single-subject families and subjects with missing values when estimating the BCCs not only will significantly reduce the accuracy of the estimators but also raise the question whether the estimators are even valid and interpretable. Linear mixed models are especially appealing due to its ability to handle imbalanced and incomplete data. We thus proposed the unified BLME to estimate all three versions of the BCCs between two biomarkers by utilizing every piece of data, regardless of the degree of uneven family sizes and missing data. *SAS PROC MIXED* is commonly used to efficiently fit univariate linear mixed effects models (alternatives include the *lme4* (Bates et al., 2014) and *nlme* (Pinheiro et al., 2013) package implemented in R). However, by properly organizing the data, we can solve the BLME (1) via the univariate linear mixed effects model utilizing *SAS PROC MIXED* to obtain $\hat{\phi}$ and $\text{CoVar}(\hat{\phi})$. The detailed data organization and SAS script for practical implementation are provided in Supporting Information Appendix SA.4.

5 Simulation studies

The missing data and imbalanced family size issues (where many families have only one single member) are common in real-world family-type clustered studies. The main purpose of the simulation studies is to examine the performance of the proposed BCC estimators under different scenarios in presence/absence of the missing data or/and imbalanced family size issue in comparison to the existing alternatives in Section 5.1 and to evaluate the size and power of the proposed Wald test on a BCC in Section 5.2. We computed several BCCs in the literature as mentioned in Section 3.1, the Bland and Altman estimator r_w^{BA} and r_b^{BA} as an alternative to r_w (3) and r_b^* (5), respectively, and the hidden correlation r_b^H as an alternative estimator to r_b (2). Additionally, for the within-family BCC r_w , when there are families with a family size $n_f = 4$, it is both intuitive and tempting to obtain an estimator through a two-stage process using the meta-analysis approach, mimicking the meta analysis on correlation coefficients across independent studies (Hedges and Olkin, 1985; Hunter and Schmidt, 2004). At the first stage, the Pearson correlation coefficients (r_p) between two variables within each family were estimated and Fisher's Z (Fisher, 1921) transformed r_p were derived with their approximate variances determined by family size as $\frac{1}{n_f - 3}$. At the second stage, a random effect meta-analysis model was fitted through the R package "metafor" (Viechtbauer, 2010) to estimate the within-family BCC and Fisher's Z transformation was reversely applied to obtain the final estimator on the within-family subject-level BCC (labeled as r_w^{Meta}) as an alternative estimator to r_w (3). Notice that r_w^{Meta} ignored data from families with $n_f < 4$ due to the approximate variance. Throughout the simulations, we fixed the variance parameters: $g_{11} = 2$, $g_{22} = 3$, $s_{11} = 1$, and $s_{22} = 2$, but varied the covariance parameters to result in a range of BCCs: $r_b = \{-0.6, 0, 0.3, 0.6\}$ and $r_w = \{-0.8, 0, 0.5, 0.8\}$, ultimately deciding r_o . A total of 1000 datasets were independently simulated under each combination of (r_b , r_w) in all simulation studies.

5.1 Simulations on BCC estimation

For estimation of the BCCs, we assumed a total number of 200 families $F = 200$ and a fixed five individuals per family, that is, $n_f = 5$, $f = 1, 2, \dots, 200$. We first randomly generated a complete dataset according to the BLME (1) under each combination of (r_b , r_w). To introduce imbalance or/and missing, a simulated dataset went subsequently through the following two sequential operations: (i) randomly select a percentage (0% or 30%) of families (labeled as "%single" in Tables 1, 2, and 3) to keep data from only one of the subjects in these selected families; (ii) randomly select a percentage (0% or 30%) of subjects (labeled as "%miss" in Tables 1, 2, and 3) across families so that the selected subjects have either one (randomly determined) of the paired observations missing. We thus employed for the simulations study a factorial design incorporating (four cases of) r_b , (four cases of) r_w , (two cases of) percentage of single-subject families, and (two cases of) missing percentage, a total of 64 scenarios. Obviously, 0% of single-subject families and 0% missing corresponded to a complete untouched dataset initially simulated. In the worst scenario, 30% of 200 families contain a single subject while 30% of all subjects have one of the paired measurements missing. All the proposed BCCs and their existing alternatives were then

calculated under each same dataset. We evaluated performance of the proposed BCC estimators based on bias, root mean squared error (RMSE) and coverage probability of the resulting 95% Wald CIs (i.e., the percentage of times the 95% CIs contain a true parameter), both in the original scale and Fisher's Z scale. Among the alternative BCC estimators in comparison, bias, RMSE and coverage will be reported for r_b^H and r_w^{Meta} while only bias will be reported for the others as their variances are not originally possible/available.

The estimation results on the family-level BCC r_b and its alternative r_b^H are presented in Table 1. The results on the family-level BCC using the family-level averages (r_b^*) and its alternative r_b^{BA} are displayed in Supporting Information Table S1. Table 2 displays the results on the subject-level BCC r_w , together with its alternatives of r_w^{BA} and r_w^{Meta} . In Table 3, the overall BCC r_o was compared to the naive Pearson correlation coefficient r_p . In estimating all the different versions of BCCs, our proposed estimators performed superior to their corresponding alternatives with comparable or even smaller biases and RMSEs relative to their alternatives. The resulting coverage was approximately at the nominal 95% level in all scenarios while Fisher's Z scale mostly led to higher coverage than the original scale. The issues of missing data and family size exerted their influence on all the BCC estimators under evaluation and led to larger biases and RMSEs and reduced coverage than under the ideal balanced and no missing data situations. However, our proposed BCC estimators were least influenced in comparison to their alternatives. Among the alternative estimators, some rendered comparable results to ours when the two issues were absent. In estimating the subject-level BCC, $\hat{r}_w^{BA}$ performed almost the same as r_w and sometimes even better in terms of bias and its biases only slightly increased in presence of the two data issues. The good performance could probably be attributed to the use of a balanced family size in our simulation settings. Since variance estimation is not possible with $\hat{r}_w^{BA}$ in the original reference, evaluation of this estimator in terms of RMSE and coverage was not reported. The meta-analysis approach ($\hat{r}_w^{Meta}$) performed poorly with large biases and RMSEs and the coverage was terribly low except when the true subject-level BCC is zero. In estimating the cluster-level BCC, $\hat{r}_b^H$ performed consistently well across simulations with resulting RMSEs almost the same as $\hat{r}_b$ in the absence of the two data issues but in their presence, the resulting biases and RMSEs clearly increased while coverage obviously decreased in most cases. The hidden correlation $\hat{r}_b^H$ was influenced more by the proportion of missing data than the family size issue, while this could also be observed on $\hat{r}_b^{BA}$ in comparison to $\hat{r}_b^*$ and r_p to r_o . Intolerance of missing values is likely due to the use of the sum of squares method in deriving relevant covariance parameters in these BCC estimators. The naive estimator of Pearson correlation coefficient (i.e., ignoring the familial structure) is commonly calculated in practice to summarize the correlation between two variables even when data has a clustered structure. In our setting with a relatively large number of families ($F = 200$) and a small fixed and balanced family size ($n_f = 5$), the simulation results in Table 3 showed that the naive Pearson correlation (r_p) seemed to be a reasonable estimator of r_o based on bias (only) in the simulated setting.

To further investigate when naive Pearson correlation performs best in estimating r_o , we conducted more simulation studies similar to the aforementioned simulation settings (without considering the two data issues) but varied family size $n_f = (1, 3, 5, 8, 10, 20)$ and total number of families $F = (50, 100, 200, 400)$, that is, a factorial design involving combinations of (r_w, r_b, n_f, F) . The averaged Pearson correlation coefficients of 1000 randomly simulated datasets per scenario were calculated and the biases (i.e., deviation of mean estimation from the true r_o) are displayed against family size in Supporting Information Fig. S1. Overall, estimation when nonzero r_w and r_b had the same sign (leading to a relatively large absolute r_o) yielded less variable biases across family size n_f and across total number of families F than otherwise. When the two had opposite signs, bias under $n_f = 1$ tended to be the smallest than larger n_f . Under a fixed combination scenario of r_w and r_b , biases shrank with increasing total numbers of families. The bias curves for $F = 50$ showed the greatest variability along family size while the two curves of $F = 200$ and 400 , close to each other, showed the least variability. Interestingly, the paired panel of curves for r_o of the same magnitude (e.g., 0.6716, 0.3795, 0.2921, 0.0874) but opposite signs appeared almost like reflections across the zero horizontal line (which was more obviously for small n_f), suggesting that the naive Pearson correlation may underestimate (in absolute magnitude) r_o . This may pertain to the use of a balanced family size in the simulation.

5.2 Simulations on size and power of hypothesis testing

We have proposed in Section 4.3 the Wald test on testing a BCC. In this subsection, we aimed to evaluate at the 5% level the size and power of the two-sided hypothesis testing, $H_0 : r_* = c$ against $H_a : r_* \neq c$, where r_* denotes any of the proposed BCCs but here with a focus on r_b and r_w , in the original scale and Fisher's Z scale. We employed the same simulation settings as in Section 5.1 but further varied $F = (10, 20, 30, 50, 100, 200)$ and thus across combinations of (r_b, r_w) , percentage of single-subject families, missing percentage, Fisher's Z scale. For a true nonzero BCC, letting $c = 0$ gives the power of the test. Letting c equal to each of the true values of a BCC used in simulation yields the size of the test. For a true zero BCC, we can only evaluate size of the test.

The size and power (in the unit of percentage) of the Wald test on r_b along total number of families (F) are separately presented in Fig. 1 and Supporting Information Fig. S2 while those on r_w in Fig. 2 and Supporting Information Fig. S3. Generally, the Wald test on r_b achieved good size even at small $F \geq 20$ and the size approached gradually to the nominal 5% level with an increasing F . At $F = 10$, the type-I error rates were inflated, ranging from about 15% in absence of the two data issues to 20% in their presence across combinations, but Fisher's Z transformation led to improved rates though slightly deflated. With 100 families or more, the size was all acceptable around 5% across all scenarios. The two data issues impacted the size greatly at small F_s but only subtly at large F_s and the single-subject issue seemed to exhibit greater influence than the missing issue. The size curves for r_b s of the same magnitude but opposite signs (specifically, $r_b = -0.6$ and 0.6) almost overlapped at $r_w = 0$ but otherwise, the size curves deviated from each other, more obviously at small F_s , in Fisher's Z scale and in presence of the data issues. The deviation pattern seemed to depend on the sign and magnitude of r_w . Under a negative $r_w (= -0.8)$, the size of r_b of the same negative sign was closer to the 5% level ($r_b = -0.6$ closer to the 5% horizontal line

than $r_b = 0.6$) while the opposite could be seen under a positive $r_w (= 0.8)$. The difference seemed to increase with larger absolute magnitude of r_w . As expected, the power for testing $r_b = 0$ increased with F for a fixed r_b to (nearly) 100% at $F = 200$ while at a fixed F , the power for a r_b of larger absolute magnitude was higher and the two data issues led to diminished power. Even for $r_b = 0.3$, the power could attain 80% with 150 families. Due to the trade-off between size and power, Fisher's Z scale led to slightly lower power compared to the original scale but the difference became negligible with increasing F . Influence of r_w on power of testing r_b was similar to the previous observations on size: at a fixed r_w (e.g., $= -0.8$), power was greater for testing r_b of the same sign ($r_b = -0.6$) than the opposite sign ($r_b = 0.6$) at small F s but the difference reduced with an increasing F and when r_w moved toward zero. The Wald test on r_w also yielded good results in terms of size and power. In absence of the two data issues, the size curves for testing r_w started with inflated size and larger variability at small F s, moved close to 5% around $F = 50$ but stabilized with less variation around 6% at $F = 200$. Fisher's Z improved upon the original scale. The presence of the two data issues corresponded to worse size and enlarged variability along total number of families. No obvious influence from r_b was visible. With five subjects per family, the power of testing $r_w = 0$ was high above 80% even for small F as r_w in our simulation setting has absolute magnitude $= 0.5$, except when 30% of all families had a single subject and 30% of observations were missing one of the paired variable.

6 Application to biomarker correlations in DIAN

AD is an age-related brain-damaging disorder that results in progressive cognitive impairment and death. Accumulating research evidence suggests that neurodegenerative processes associated with AD begin years prior to the symptomatic onset of AD when the disease is clinically at the early prodromal stage or even the latent stage (Katzman, 1976). Many recent clinico-pathologic studies have also suggested a time window prior to the symptomatic onset of AD during which no clinical diagnosis could be rendered, but neuropathological changes of AD, notably senile plaques and neurofibrillary tangles, have accumulated (Morris and Price, 2001; Bennett et al., 2006; Price et al., 2009). These observations have led to a major paradigm shift in the study of AD, that is, the focus of identifying individuals at high risk of AD when they are still cognitively normal through biomarkers prior to the substantial development of clinical symptoms as these may be the groups of individuals in which targeted therapies may have the greatest chance of preserving normal brain function. Two major types of biomarkers have been studied. Both the 42 amino acid form of β -amyloid ($A\beta_{42}$) and tau, a microtubule-associated phosphoprotein, are dispersed to the CSF, and CSF tau protein (CSF_tau) and β -amyloid protein (CSF_ $A\beta_{42}$) have been shown as potential diagnostic biomarkers that differentiate early AD from normal aging individuals (Sjögren et al., 2001; Andreasen et al., 2011). Neuroimaging biomarkers have also been reported to show early changes in AD progression. Pittsburgh compound B (PIB), a fluorescent analog of thioflavin T has been used in PET scans to image β -amyloid plaques in neuronal tissue of brain regions (Mintun et al., 2006). The validity of these biomarkers in identifying subjects at early stages of AD, however, must be established through appropriate correlations with cognitive outcomes.

The DIAN (Section 2) is an international network that enrolled a unique cohort of individuals who are at high risk of AD by virtue of carrying a type of causative mutation for AD. A comprehensive set of AD biomarkers (e.g., CSF, PET PIB) has been collected in DIAN. In addition, a large cognitive battery has also been administered to assess subjects' cognitive changes. A crucial scientific question of DIAN is to examine the correlations across different modalities of biomarkers as well as cognitive tests during the preclinical stage when subjects are still cognitively normal.

To address the scientific aim of DIAN and to demonstrate the estimation of BCCs between markers, we analyzed in DIAN (data freeze 4) 83 individuals from 50 independent families who were cognitively normal at enrollment and were mutation positive. Age at enrollment of these individuals varied between 19 and 61. We selected four biomarkers for analysis: two CSF biomarkers, CSF_A β 42 and CSF_tau, and two PET PIB imaging markers at two brain regions, temporal (PIB_TMP) and precuneus (PIB_PRECUNEUS). We also selected two neuropsychological cognitive markers for correlation analyses: “mixed”—*the number of mixed/rearranged correctly identified* (Naveh-Benjamin, 2000) and “paper”—*mentally fold the display into an object that represents the folded paper* (Salthouse, 1989). Noticeably, 76% of the families contained only one single individual. Among all paired measurements of the six markers, the proportion of individuals with a missing observation on any pair of disease markers ranged from 6.02% to 26.51% with a median of 20.5%.

With the CSF and PET PIB markers log-transformed and the cognitive markers median centered and interquartile range scaled, a BLME model was independently fitted between each pair of the markers with the covariate EAO (estimated age of onset) (Katzman, 1976). The variance/covariance parameters were estimated and the resulting BCC estimates were calculated with their associated 95% CIs as presented in Table 4 (a–c) where a star indicated significance at the 5% level with a 95% Wald CI not crossing zero (also confirmed by conducting Wald test). The studentized residuals from BLME were plotted for visual examination of model goodness of fit. Between the two PIB markers (PIB_TMP and PIB_PRECUNEUS), the within-family subject-level BCC r_w was estimated to be 0.9 (95% CI: 0.83–0.98), indicating an almost perfect correlation of the two imaging biomarkers between subjects from the same families. Their family-level BCC r_b and overall BCC r_o estimations were also high (0.85, 95% CI: 0.6–1.0 and 0.89, 95% CI: 0.83–0.95, respectively). Taken together, the β -amyloid plaques measured in the neuronal tissue of these two brain regions were highly correlated at the family level and at the subject level, whether conditioning on the family effect or not. On the other hand, the two CSF markers (CSF_A β 42 and CSF_tau) were negatively correlated both at the family level and at the subject level, reflecting their different biological and neuropathological roles at the early stage of AD, similar to what has been reported in the later stage of AD (Shaw et al., 2009). Albeit showing a high correlation at the family level ($r_b = 0.95$, 95% CI: 0.5–1.0) and a moderate overall BCC ($r_o = 0.39$, 95% CI: 0.17–0.6), the two cognitive markers (“mixed” and “paper”) showed almost no bivariate correlation at the subject level when conditioning on the effect of families ($r_w = -0.05$, 95% CI: -0.39–0.28). Hence, the correlation of the two cognitive tests, whereas reflecting different cognitive domains, is largely due to the familial effects. These observations reflect the importance of differentiating different

correlations at different levels of experimental units, the very objective of the current work. The most scientifically important correlations in DIAN are those of disease markers across different modalities. The two PIB imaging markers in the temporal and precuneus region were both significantly negatively correlated with CSF_A β 42 within families and meanwhile highly positively correlated with CSF_tau across families and overall, confirming their validity as noninvasive surrogates for measuring β -amyloid plaques and tau. Both PET PIB β -amyloid plaques were correlated with at least one of the two cognitive outcomes moderately at the subject level and highly at the family level, justifying the potential of PET PIB to track early changes of AD. Across all BCCs, CSF_A β 42 seemed to be more correlated with the cognitive marker “mixed” than with “paper” while CSF_tau showed slightly higher correlation with “paper,” suggesting that CSF biomarkers may indicate changes in different cognitive domains at the early disease progression. The two PIB imaging markers were both highly correlated with the cognitive marker “mixed” in the negative direction at the family level though the overall BCC seemed to be higher with “paper,” again highlighting the importance of differentiating different levels of correlations. For comparison purpose, we also computed, after adjustment for EAO, the partial Pearson correlation pretending all individuals were independent to estimate r_o and the subject-level BCC r_w was estimated via the meta-analysis approach (where a random effect meta-analysis model was fitted to the partial Pearson correlation coefficients across subjects per family) but results on both were omitted due to space limitation. The overall partial Pearson correlation results were mostly similar to $\hat{r}_o$ while as the meta-analyses could be done only on three families, the results were very different from $\hat{r}_w$ due to limited data.

7 Discussion

We have proposed a unified approach based on the BLME to estimate different versions of BCCs to summarize relationship between two quantitative variables at the within-family subject level, the family (i.e., cluster) level and marginally in family-type clustered data. We have provided statistical inferences on these BCCs based on REML of the BLME, and presented the implementations using the standard software SAS. By simulation studies, we have evaluated the performance of our BLME-based BCCs relative to their alternatives and have observed their superiority in various scenarios corresponding to smaller biases, smaller RMSEs and good coverage, especially when the two common data issues (missing data and family size) are present. The influences of many families having a single individual or/and many subjects in families missing one of the paired measurements on our proposed BCCs is almost negligible but on the contrary have led to larger bias or/and RMSEs in the alternative estimators. Our simulation results also showed good performance of the Wald test in terms of type-I error rate and power in the original scale and Fisher’s Z scale, even under relatively small sample sizes. We have analyzed a real-world AD dataset from DIAN to demonstrate the usefulness of the proposed BCCs in shedding lights on bivariate relationships between AD biomarkers across different modalities while accounting for covariates. The fact that BCCs between biomarkers in DIAN are not the same at different levels of the experimental units speaks to the importance of a clear definition on different versions of BCCs in a family-type clustered design.

Normality is the fundamental distributional assumption for the BLME, as well as for most of the alternative estimators proposed in the literature. We simulated all datasets from the bivariate normal distribution in the simulation studies. While variables can be nonnormally distributed in real datasets, proper transformation (e.g., log-transformation or Box-Cox) could be performed for normal approximation before fitting the BLME though BLME has exhibited robust inference in nonserious deviation from normality, although transformations may change interpretation of the BCCs at various hierarchies. Nonetheless, residuals from the BLME should be examined for model goodness of fit and diagnosis. When goodness of fit is questionable, model assumptions can be modified to accommodate heterogeneous residual variances or the model can be extended to generalized linear mixed effects models for nonnormal data such as binary data and count data. In practice of model implementation, it may happen that the variance of a random effect is estimated to be 0, usually suggesting that there is not enough variation in the response variable attributable to the random effect, after accounting for covariates and residuals in the model. Readers are referred to Kiernan et al. (2012) and Piepho et al. (2014) for good recommendations on practical computations, for example, rescaling the covariates and the outcome variables to be on the same scale and trying different starting values for parameters. In analysis of BCC among multiple variables as in the real data analysis section, we have fitted a separate BLME model to each pair of variables. For future research, a multivariate mixed effects model can be employed to model the BCCs among all variables simultaneously.

Based on our simulation results, the hidden correlation coefficient (Nguyen and Jiang, 2011) serves as a good estimator comparable to our proposed r_b when the two data issues (imbalanced family size and missing) are absent while the Bland and Altman estimation of the within-family subject-level BCC (Bland and Altman, 1995a) performs quite well in terms of bias in the balanced cluster size simulation setting though evaluations in terms of RMSE and coverage are not possible as the variance was not originally available. We have tried a meta-analysis approach to estimate the subject-level BCC in the family-type clustered design (see Section 5) but it showed poor performance probably because of the small family size (five subjects per family) relative to the usual much larger sample sizes in its application to independent studies. Thus, we do not recommend its use in practice for clustered studies of small cluster sizes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants NIH/NCI P30CA091842, NCI U10CA180860-01, and Komen promise grant PG12220321 for Dr. Luo, and NIH/NIA R01 AG029672 and NIH/NIA R01 AG034119 and in part by grants P50 AG005681, P01 AG003991, P01 AG026276, and U01 AG032438 from the National Institute on Aging for Dr. Xiong. The authors thank the Clinical Core, biomarker core, and imaging core of DIAN for subject assessments and data collection.

References

- Andreasen LM, Davidsson P, Vanmechelen E, Vanderstichele H, Winblad B, Blennow K. Evaluation of CSF-tau and CSF-AB42 as diagnostic markers for Alzheimer disease in clinical practice. *Archives of Neurology*. 2011; 58:373–379. [PubMed: 11255440]
- Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, Marcus DS, Cairns NJ, Xie X, Blazey TM, Holtzman DM, Santacruz A, Buckles V, Oliver A, Moulder K, Aisen PS, Ghetti B, Klunk WE, McDade E, Martins RN, Masters CL, Mayeux R, Ringman JM, Rossor MN, Schofield PR, Sperling RA, Salloway S, Morris JC. for the Dominantly Inherited Alzheimer Network. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *The New England Journal of Medicine*. 2012; 367:795–804. [PubMed: 22784036]
- Bates D, Maechler M, Bolker B, Walker S. lme4: linear mixed-effects models using Eigen and S4. R package version 1.1–5. 2014 Retrieved from <http://CRAN.R-project.org/package=lme4>.
- Bello NM, Stevenson JS, Tempelman RJ. Invited review: milk production and reproductive performance: modern interdisciplinary insights into an enduring axiom. *Journal of Dairy Science*. 2012; 95:5461–5475. [PubMed: 22863097]
- Bennett D, Schneider J, Arvanitakis Z, Kelly J, Aggarwal N, Shah R, Wilson R. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*. 2006; 66:837–1844.
- Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1—correlation within subjects. *BMJ*. 1995a February;310:446. [PubMed: 7873953]
- Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2—correlation between subjects. *BMJ*. 1995b Mar.310:633. [PubMed: 7703752]
- Fisher RA. On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*. 1921; 1:3–32.
- Gao, F.; Thompson, P.; Xiong, C.; Miller, JP. Analyzing multivariate longitudinal data using SAS®; Proceedings of the Thirty-first Annual SAS Users Group International Conference; 2006. p. 187-131.
- Hamlett A, Ryan L, Wolfinger R. On the use of PROC MIXED to estimate correlation in the presence of repeated measures. *SAS Users Group International, Proceedings of the Statistics and Data Analysis Section Paper*. 2004; 198:1–7.
- Hedges, LV.; Olkin, I. *Statistical Methods for Meta-Analysis*. London, UK: Academic Press; 1985.
- Hunter, J.; Schmidt, FL. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. 2nd. Thousand Oaks, CA: Sage; 2004.
- Katzman R. The prevalence and malignancy of Alzheimer disease: a major killer. *Archives of Neurology*. 1976; 33:217–218.
- Kiernan K, Tao J, Gibbs P. Tips and strategies for mixed modeling with SAS/STAT procedures. *SAS Global Forum*. 2012; 2012:332.
- Lorenz D, Datta S, Harkema S. Marginal association measures for clustered data. *Statistics in Medicine*. 2011; 30:3181–3191. [PubMed: 21953204]
- Mills SM, Mallmann J, Santacruz AM, Fuqua A, Carril M, Aisen PS, Althage MC, Belyew S, Benzinger TL, Brooks WS, Buckles VD, Cairns NJ, Clifford D, Danek A, Fagan AM, Farlow M, Fox N, Ghetti B, Goate AM, Heinrichs D, Hornbeck R, Jack C, Jucker M, Klunk WE, Marcus DS, Martins RN, Masters CM, Mayeux R, McDade E, Morris JC, Oliver A, Ringman JM, Rossor MN, Salloway S, Schofield PR, Snider J, Snyder P, Sperling RA, Stewart C, Thomas RG, Xiong C, Bateman RJ. Preclinical trials in autosomal dominant Alzheimer's disease: implementation of the DIAN-TU. *Revue Neurologique*. 2013; 169:737–743. [PubMed: 24016464]
- Mintun MA, Larossa GN, Sheline YI, Dence CS, Lee SY, Mach RH, Klunk WE, Mathis CA, DeKosky ST, Morris JC. [11C] PIB in a nondemented population: potential antecedent marker of Alzheimer disease. *Neurology*. 2006; 67:446–452.
- Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology*. 1993; 43:2412–2414. [PubMed: 8232972]

- Morris JC, Price JL. Pathologic correlates of nondemented aging, mild cognitive impairment, and early stage Alzheimer's disease. *Journal of Molecular Neuroscience*. 2001; 17:101–118. [PubMed: 11816784]
- Naveh-Benjamin M. Adult age differences in memory performance: tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2000; 26:1170–1187.
- Nguyen T, Jiang J. Simple estimation of hidden correlation in repeated measures. *Statistics in Medicine*. 2011; 30:3403–3415. [PubMed: 21997471]
- Piepho HP, Müller BU, Jansen C. Analysis of a complex trait with missing data on the component traits. *Communications in Biometry and Crop Science*. 2014; 9:26–40.
- Pinheiro J, Bates D, DebRoy S, Sarkar D, Team TR. nlme: linear and nonlinear mixed effects models. R package version 3.1–121. 2013 Retrieve from <http://CRAN.R-project.org/package=nlme>.
- Price JL, McKeel DW, Buckles VD, R package version 3.1–109. Roe CM, Xiong C, Grundman M, Hansen LA, Petersen RC, Parisi JE, Dickson DW, Smith CD, Davis DG, Schmitt FA, Markesbery WR, Kaye J, Kurlan R, Hulette C, Kurland Higdon R, Jukull W, Morris JC. Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease. *Neurobiol Aging*. 2009; 30:1026–1036. [PubMed: 19376612]
- Roy A. Estimating correlation coefficient between two variables with repeated observations using mixed effects model. *Biometrical Journal*. 2006; 48:286–301. [PubMed: 16708779]
- Salthouse TM. Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1989; 15:507–516.
- Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Dean R, Siemers E, Potter W, Lee VM, Rojanowski JQ. Alzheimer's Disease Neuroimaging Initiative. Cerebrospinal fluid biomarker signature in Alzheimer's Disease Neuroimaging Initiative subjects. *Annals of Neurology*. 2009; 65:403–413. [PubMed: 19296504]
- Sjögren M, Vanderstichele H, Agren H, Zachrisson O, Edsbacke M, Wikkelsø C, Skoog I, Wahlund LO, Marcusson J, Nägga K, Andreasen N, Davidsson P, Vanmeceken E, Blennow K. Tau and A β 42 in cerebrospinal fluid from healthy adults 21–93 years of age: establishment of reference values. *Clinical Chemistry*. 2001; 47:1776–1781. [PubMed: 11568086]
- Thiébaud R, Jacqmin-Gadda H, Chêne G, Leport C, Commenges D. Bivariate linear mixed models using SAS PROC MIXED. *Computer Methods and Programs in Biomedicine*. 2002; 69:249–256. [PubMed: 12204452]
- Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010; 36:1–48.

Figure 1. Size (in the unit of %) of the Wald test on r_b (indicated by colored lines with point shapes) along total number of families (x -axis) by r_w (at row panels) and the scenarios of single subject and missing percentage (at column panels), in the original scale (solid lines) and Fisher's Z scale (dashed lines).

Figure 2.

Size (in the unit of %) of the Wald test on r_w (indicated by colored lines with point shapes) along total number of families (x -axis) by r_b (at row panels) and the scenarios of single subject and missing percentage (at column panels), in the original scale (solid lines) and Fisher's Z scale (dashed lines).

Table 1

Simulation results on estimating the cluster-level BCC (r_b) using our proposed estimator ($\hat{r}_b^{(H)}$) and the hidden correlation coefficient ($\hat{\rho}_b^H$) across scenarios.

True parameters					$\hat{r}_b$				$\hat{\rho}_b^H$			
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*
-0.6	-0.8	-0.6716	0	0	-0.0009	0.0705	0.95	0.954	-0.0008	0.0704	0.942	0.949
-0.6	-0.8	-0.6716	0	0.3	-0.0006	0.0717	0.947	0.954	0.0837	0.1262	0.76	0.699
-0.6	-0.8	-0.6716	0.3	0	-0.0009	0.0764	0.939	0.956	-0.0013	0.0872	0.914	0.924
-0.6	-0.8	-0.6716	0.3	0.3	-0.0010	0.0790	0.945	0.955	0.0829	0.1440	0.829	0.768
-0.6	0	-0.3795	0	0	-0.0023	0.0764	0.951	0.958	-0.0027	0.0763	0.944	0.954
-0.6	0	-0.3795	0	0.3	-0.0017	0.0794	0.949	0.951	-0.0031	0.0933	0.922	0.938
-0.6	0	-0.3795	0.3	0	-0.0031	0.0849	0.95	0.959	-0.0048	0.0943	0.915	0.93
-0.6	0	-0.3795	0.3	0.3	-0.0034	0.0892	0.952	0.959	-0.0051	0.1154	0.911	0.933
-0.6	0.5	-0.1969	0	0	-0.0032	0.0808	0.955	0.962	-0.0039	0.0807	0.943	0.957
-0.6	0.5	-0.1969	0	0.3	-0.0027	0.0830	0.951	0.953	-0.0575	0.1115	0.812	0.892
-0.6	0.5	-0.1969	0.3	0	-0.0044	0.0903	0.943	0.964	-0.0069	0.0999	0.914	0.929
-0.6	0.5	-0.1969	0.3	0.3	-0.0048	0.0938	0.949	0.962	-0.0600	0.1329	0.828	0.926
-0.6	0.8	-0.0874	0	0	-0.0038	0.0837	0.958	0.96	-0.0046	0.0836	0.944	0.957
-0.6	0.8	-0.0874	0	0.3	-0.0035	0.0847	0.956	0.952	-0.0902	0.1332	0.705	0.809
-0.6	0.8	-0.0874	0.3	0	-0.0052	0.0937	0.949	0.963	-0.0082	0.1036	0.916	0.935
-0.6	0.8	-0.0874	0.3	0.3	-0.0057	0.0960	0.948	0.96	-0.0928	0.1535	0.728	0.893
0	-0.8	-0.2921	0	0	-0.0019	0.1125	0.95	0.954	-0.0014	0.1124	0.945	0.946
0	-0.8	-0.2921	0	0.3	-0.0016	0.1137	0.948	0.95	0.0835	0.1513	0.821	0.833
0	-0.8	-0.2921	0.3	0	-0.0020	0.1224	0.952	0.954	-0.0027	0.1389	0.915	0.921
0	-0.8	-0.2921	0.3	0.3	-0.0027	0.1260	0.955	0.957	0.0815	0.1762	0.854	0.871
0	0	0	0	0	-0.0034	0.1120	0.953	0.954	-0.0034	0.1118	0.944	0.947
0	0	0	0	0.3	-0.0027	0.1147	0.949	0.954	-0.0036	0.1237	0.935	0.936
0	0	0	0.3	0	-0.0042	0.1224	0.954	0.957	-0.0061	0.1380	0.917	0.923
0	0	0	0.3	0.3	-0.0050	0.1276	0.95	0.953	-0.0068	0.1526	0.917	0.92
0	0.5	0.1826	0	0	-0.0043	0.1122	0.95	0.956	-0.0046	0.1120	0.945	0.948
0	0.5	0.1826	0	0.3	-0.0037	0.1142	0.949	0.954	-0.0582	0.1375	0.876	0.886

True parameters					$\hat{r}_b$				$\hat{r}_b^H$			
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*
0	0.5	0.1826	0.3	0	-0.0055	0.1223	0.955	0.96	-0.0081	0.1384	0.92	0.927
0	0.5	0.1826	0.3	0.3	-0.0065	0.1269	0.949	0.955	-0.0618	0.1661	0.874	0.884
0	0.8	0.2921	0	0	-0.0049	0.1125	0.952	0.959	-0.0053	0.1124	0.944	0.951
0	0.8	0.2921	0	0.3	-0.0045	0.1134	0.951	0.956	-0.0910	0.1555	0.802	0.808
0	0.8	0.2921	0.3	0	-0.0063	0.1221	0.953	0.96	-0.0093	0.1389	0.924	0.928
0	0.8	0.2921	0.3	0.3	-0.0074	0.1257	0.957	0.964	-0.0947	0.1831	0.825	0.835
0.3	-0.8	-0.1024	0	0	-0.0018	0.1072	0.951	0.954	-0.0012	0.1071	0.943	0.945
0.3	-0.8	-0.1024	0	0.3	-0.0015	0.1084	0.947	0.947	0.0840	0.1464	0.793	0.834
0.3	-0.8	-0.1024	0.3	0	-0.0020	0.1180	0.948	0.954	-0.0024	0.1327	0.914	0.922
0.3	-0.8	-0.1024	0.3	0.3	-0.0028	0.1212	0.956	0.954	0.0820	0.1695	0.824	0.877
0.3	0	0.1897	0	0	-0.0032	0.1035	0.95	0.954	-0.0030	0.1034	0.945	0.946
0.3	0	0.1897	0	0.3	-0.0026	0.1063	0.944	0.95	-0.0032	0.1166	0.933	0.938
0.3	0	0.1897	0.3	0	-0.0040	0.1136	0.955	0.958	-0.0055	0.1279	0.921	0.929
0.3	0	0.1897	0.3	0.3	-0.0048	0.1185	0.953	0.956	-0.0065	0.1438	0.916	0.92
0.3	0.5	0.3723	0	0	-0.0041	0.1017	0.952	0.955	-0.0042	0.1016	0.945	0.946
0.3	0.5	0.3723	0	0.3	-0.0035	0.1038	0.944	0.952	-0.0578	0.1306	0.884	0.871
0.3	0.5	0.3723	0.3	0	-0.0051	0.1108	0.955	0.956	-0.0074	0.1258	0.927	0.929
0.3	0.5	0.3723	0.3	0.3	-0.0060	0.1152	0.954	0.954	-0.0616	0.1573	0.879	0.874
0.3	0.8	0.4819	0	0	-0.0046	0.1009	0.951	0.954	-0.0049	0.1008	0.945	0.946
0.3	0.8	0.4819	0	0.3	-0.0042	0.1018	0.949	0.954	-0.0906	0.1491	0.805	0.789
0.3	0.8	0.4819	0.3	0	-0.0058	0.1090	0.953	0.96	-0.0086	0.1249	0.925	0.932
0.3	0.8	0.4819	0.3	0.3	-0.0068	0.1124	0.952	0.959	-0.0945	0.1748	0.826	0.82
0.6	-0.8	0.0874	0	0	-0.0010	0.0847	0.948	0.952	-0.0001	0.0845	0.943	0.945
0.6	-0.8	0.0874	0	0.3	-0.0007	0.0859	0.951	0.947	0.0853	0.1313	0.726	0.834
0.6	-0.8	0.0874	0.3	0	-0.0012	0.0952	0.949	0.95	-0.0006	0.1055	0.923	0.926
0.6	-0.8	0.0874	0.3	0.3	-0.0019	0.0976	0.947	0.956	0.0840	0.1493	0.761	0.912
0.6	0	0.3795	0	0	-0.0022	0.0773	0.947	0.955	-0.0018	0.0772	0.941	0.947
0.6	0	0.3795	0	0.3	-0.0017	0.0802	0.945	0.944	-0.0019	0.0949	0.928	0.936
0.6	0	0.3795	0.3	0	-0.0028	0.0861	0.951	0.956	-0.0035	0.0960	0.919	0.929

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

True parameters					$\hat{r}_b$			$\hat{r}_b^H$				
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*
0.6	0	0.3795	0.3	0.3	-0.0033	0.0904	0.95	0.957	-0.0046	0.1168	0.917	0.928
0.6	0.5	0.5620	0	0	-0.0029	0.0733	0.945	0.953	-0.0029	0.0733	0.944	0.948
0.6	0.5	0.5620	0	0.3	-0.0024	0.0755	0.944	0.946	-0.0565	0.1105	0.862	0.829
0.6	0.5	0.5620	0.3	0	-0.0037	0.0806	0.955	0.959	-0.0052	0.0911	0.925	0.929
0.6	0.5	0.5620	0.3	0.3	-0.0042	0.0843	0.951	0.957	-0.0597	0.1315	0.873	0.832
0.6	0.8	0.6716	0	0	-0.0034	0.0713	0.95	0.954	-0.0035	0.0712	0.943	0.947
0.6	0.8	0.6716	0	0.3	-0.0030	0.0723	0.949	0.951	-0.0893	0.1312	0.737	0.686
0.6	0.8	0.6716	0.3	0	-0.0042	0.0772	0.954	0.96	-0.0062	0.0886	0.928	0.931
0.6	0.8	0.6716	0.3	0.3	-0.0048	0.0798	0.955	0.954	-0.0927	0.1511	0.78	0.73

Coverage*: coverage under Fisher's Z scale.

Table 2

Simulation results on estimating the subject-level BCC (r_w) using our proposed estimator ($\hat{r}_w$), the meta analysis approach ($\hat{r}_w^{Meta}$) and the Bland–Altman “correlation within subject” estimator ($\hat{r}_w^{BA}$) across scenarios.

True parameters							$\hat{r}_w$			$\hat{r}_w^{Meta}$			$\hat{r}_w^{BA}$		
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage
-0.6	-0.8	-0.6716	0	0	-0.0002	0.0183	0.937	0.937	-0.0473	0.0511	0.11	0.134	-0.0002	0.0000	0.134
-0.6	-0.8	-0.6716	0	0.3	-0.0004	0.0215	0.941	0.937	-0.0575	0.0654	0.307	0.392	0.0000	0.0000	0.392
-0.6	-0.8	-0.6716	0.3	0	0.0001	0.0216	0.94	0.939	-0.0468	0.0521	0.23	0.275	0.0001	0.0001	0.275
-0.6	-0.8	-0.6716	0.3	0.3	0.0001	0.0254	0.94	0.951	-0.0569	0.0676	0.432	0.545	0.0004	0.0004	0.545
-0.6	0	-0.3795	0	0	-0.0009	0.0510	0.939	0.94	-0.0010	0.0706	0.951	0.954	-0.0009	0.0009	0.954
-0.6	0	-0.3795	0	0.3	-0.0011	0.0635	0.935	0.936	0.0001	0.1201	0.951	0.954	-0.0010	0.0010	0.954
-0.6	0	-0.3795	0.3	0	-0.0001	0.0598	0.942	0.942	-0.0002	0.0837	0.952	0.954	-0.0002	0.0002	0.954
-0.6	0	-0.3795	0.3	0.3	0.0009	0.0745	0.945	0.946	0.0022	0.1407	0.956	0.96	-0.0001	0.0001	0.96
-0.6	0.5	-0.1969	0	0	-0.0008	0.0384	0.935	0.94	0.0633	0.0795	0.532	0.571	-0.0008	0.0008	0.571
-0.6	0.5	-0.1969	0	0.3	-0.0011	0.0471	0.935	0.936	0.0793	0.1118	0.662	0.713	-0.0013	0.0013	0.713
-0.6	0.5	-0.1969	0.3	0	-0.0003	0.0449	0.941	0.943	0.0634	0.0852	0.647	0.699	-0.0005	0.0005	0.699
-0.6	0.5	-0.1969	0.3	0.3	0.0003	0.0550	0.937	0.942	0.0795	0.1230	0.739	0.803	-0.0006	0.0006	0.803
-0.6	0.8	-0.0874	0	0	-0.0005	0.0185	0.933	0.935	0.0468	0.0507	0.117	0.149	-0.0004	0.0004	0.149
-0.6	0.8	-0.0874	0	0.3	-0.0004	0.0219	0.936	0.937	0.0577	0.0655	0.3	0.384	-0.0008	0.0008	0.384
-0.6	0.8	-0.0874	0.3	0	-0.0002	0.0216	0.936	0.942	0.0467	0.0522	0.22	0.287	-0.0003	0.0003	0.287
-0.6	0.8	-0.0874	0.3	0.3	0.0000	0.0257	0.939	0.939	0.0577	0.0686	0.388	0.517	-0.0005	0.0005	0.517
0	-0.8	-0.2921	0	0	-0.0002	0.0183	0.937	0.937	-0.0473	0.0511	0.11	0.134	-0.0002	0.0002	0.134
0	-0.8	-0.2921	0	0.3	-0.0004	0.0215	0.941	0.938	-0.0575	0.0654	0.307	0.392	0.0000	0.0000	0.392
0	-0.8	-0.2921	0.3	0	0.0001	0.0216	0.94	0.935	-0.0468	0.0521	0.23	0.275	0.0001	0.0001	0.275
0	-0.8	-0.2921	0.3	0.3	0.0001	0.0255	0.941	0.948	-0.0569	0.0676	0.432	0.545	0.0004	0.0004	0.545
0	0	0	0	0	-0.0009	0.0510	0.939	0.94	-0.0010	0.0706	0.951	0.954	-0.0009	0.0009	0.954
0	0	0	0	0.3	-0.0011	0.0638	0.935	0.935	0.0001	0.1201	0.951	0.954	-0.0010	0.0010	0.954
0	0	0	0.3	0	-0.0001	0.0600	0.94	0.94	-0.0002	0.0837	0.952	0.954	-0.0002	0.0002	0.954
0	0	0	0.3	0.3	0.0007	0.0750	0.942	0.944	0.0022	0.1407	0.956	0.96	-0.0001	0.0001	0.96
0	0.5	0.1826	0	0	-0.0008	0.0384	0.935	0.94	0.0633	0.0795	0.532	0.571	-0.0008	0.0008	0.571

True parameters										$\hat{r}_w$				$\hat{r}_w^{Meta}$				$\hat{r}_w^{BA}$	
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*		Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*	Bias	
0	0.5	0.1826	0	0.3	-0.0011	0.0472	0.934	0.938		0.0793	0.1118	0.662	0.713	-0.0013				-0.0013	
0	0.5	0.1826	0.3	0	-0.0004	0.0451	0.939	0.94		0.0634	0.0852	0.647	0.699	-0.0005				-0.0005	
0	0.5	0.1826	0.3	0.3	0.0002	0.0553	0.939	0.942		0.0795	0.1230	0.739	0.803	-0.0006				-0.0006	
0	0.8	0.2921	0	0	-0.0005	0.0185	0.933	0.935		0.0468	0.0507	0.117	0.149	-0.0004				-0.0004	
0	0.8	0.2921	0	0.3	-0.0004	0.0219	0.937	0.934		0.0577	0.0655	0.3	0.384	-0.0008				-0.0008	
0	0.8	0.2921	0.3	0	-0.0003	0.0217	0.939	0.942		0.0467	0.0522	0.22	0.287	-0.0003				-0.0003	
0	0.8	0.2921	0.3	0.3	0.0000	0.0257	0.942	0.94		0.0577	0.0686	0.388	0.517	-0.0005				-0.0005	
0.3	-0.8	-0.1024	0	0	-0.0002	0.0183	0.937	0.937		-0.0473	0.0511	0.11	0.134	-0.0002				-0.0002	
0.3	-0.8	-0.1024	0	0.3	-0.0003	0.0215	0.941	0.939		-0.0575	0.0654	0.307	0.392	0.0000				0.0000	
0.3	-0.8	-0.1024	0.3	0	0.0001	0.0216	0.939	0.937		-0.0468	0.0521	0.23	0.275	0.0001				0.0001	
0.3	-0.8	-0.1024	0.3	0.3	0.0001	0.0254	0.941	0.95		-0.0569	0.0676	0.432	0.545	0.0004				0.0004	
0.3	0	0.1897	0	0	-0.0009	0.0510	0.939	0.94		-0.0010	0.0706	0.951	0.954	-0.0010				-0.0010	
0.3	0	0.1897	0	0.3	-0.0011	0.0638	0.936	0.937		0.0001	0.1201	0.951	0.954	-0.0010				-0.0010	
0.3	0	0.1897	0.3	0	-0.0001	0.0600	0.938	0.939		-0.0002	0.0837	0.952	0.954	-0.0002				-0.0002	
0.3	0	0.1897	0.3	0.3	0.0007	0.0749	0.943	0.944		0.0022	0.1407	0.956	0.96	-0.0001				-0.0001	
0.3	0.5	0.3723	0	0	-0.0008	0.0384	0.935	0.94		0.0633	0.0795	0.532	0.571	-0.0008				-0.0008	
0.3	0.5	0.3723	0	0.3	-0.0011	0.0471	0.932	0.935		0.0793	0.1118	0.662	0.713	-0.0013				-0.0013	
0.3	0.5	0.3723	0.3	0	-0.0004	0.0452	0.938	0.938		0.0634	0.0852	0.647	0.699	-0.0005				-0.0005	
0.3	0.5	0.3723	0.3	0.3	0.0002	0.0553	0.939	0.943		0.0795	0.1230	0.739	0.803	-0.0006				-0.0006	
0.3	0.8	0.4819	0	0	-0.0005	0.0185	0.933	0.935		0.0468	0.0507	0.117	0.149	-0.0004				-0.0004	
0.3	0.8	0.4819	0	0.3	-0.0004	0.0219	0.937	0.934		0.0577	0.0655	0.3	0.384	-0.0008				-0.0008	
0.3	0.8	0.4819	0.3	0	-0.0003	0.0217	0.939	0.942		0.0467	0.0522	0.22	0.287	-0.0003				-0.0003	
0.3	0.8	0.4819	0.3	0.3	0.0000	0.0257	0.942	0.938		0.0577	0.0686	0.388	0.517	-0.0005				-0.0005	
0.6	-0.8	0.0874	0	0	-0.0002	0.0183	0.937	0.937		-0.0473	0.0511	0.11	0.134	-0.0002				-0.0002	
0.6	-0.8	0.0874	0	0.3	-0.0003	0.0215	0.941	0.94		-0.0575	0.0654	0.307	0.392	0.0000				0.0000	
0.6	-0.8	0.0874	0.3	0	0.0001	0.0215	0.94	0.938		-0.0468	0.0521	0.23	0.275	0.0001				0.0001	
0.6	-0.8	0.0874	0.3	0.3	0.0001	0.0254	0.942	0.949		-0.0569	0.0676	0.432	0.545	0.0004				0.0004	
0.6	0	0.3795	0	0	-0.0009	0.0510	0.939	0.94		-0.0010	0.0706	0.951	0.954	-0.0010				-0.0010	
0.6	0	0.3795	0	0.3	-0.0010	0.0637	0.936	0.936		0.0001	0.1201	0.951	0.954	-0.0010				-0.0010	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

True parameters										$\hat{r}_w$				$\hat{r}_w^{Meta}$				$\hat{r}_w^{BA}$			
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*		Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*	Bias	RMSE	Coverage	Coverage*
0.6	0	0.3795	0.3	0	-0.0001	0.0599	0.935	0.936		-0.0002	0.0837	0.952	0.954	-0.0002	0.0837	0.952	0.954	-0.0002	0.0837	0.952	0.954
0.6	0	0.3795	0.3	0.3	0.0007	0.0747	0.941	0.942		0.0022	0.1407	0.956	0.96	0.0022	0.1407	0.956	0.96	0.0001	0.1407	0.956	0.96
0.6	0.5	0.5620	0	0	-0.0008	0.0384	0.935	0.94		0.0633	0.0795	0.532	0.571	0.0633	0.0795	0.532	0.571	-0.0008	0.0795	0.532	0.571
0.6	0.5	0.5620	0	0.3	-0.0010	0.0471	0.932	0.933		0.0793	0.1118	0.662	0.713	0.0793	0.1118	0.662	0.713	-0.0013	0.1118	0.662	0.713
0.6	0.5	0.5620	0.3	0	-0.0004	0.0451	0.938	0.939		0.0634	0.0852	0.647	0.699	0.0634	0.0852	0.647	0.699	-0.0005	0.0852	0.647	0.699
0.6	0.5	0.5620	0.3	0.3	0.0001	0.0551	0.942	0.943		0.0795	0.1230	0.739	0.803	0.0795	0.1230	0.739	0.803	-0.0006	0.1230	0.739	0.803
0.6	0.8	0.6716	0	0	-0.0005	0.0185	0.933	0.935		0.0468	0.0507	0.117	0.149	0.0468	0.0507	0.117	0.149	-0.0004	0.0507	0.117	0.149
0.6	0.8	0.6716	0	0.3	-0.0004	0.0219	0.937	0.931		0.0577	0.0655	0.3	0.384	0.0577	0.0655	0.3	0.384	-0.0008	0.0655	0.3	0.384
0.6	0.8	0.6716	0.3	0	-0.0003	0.0218	0.94	0.94		0.0467	0.0522	0.22	0.287	0.0467	0.0522	0.22	0.287	-0.0003	0.0522	0.22	0.287
0.6	0.8	0.6716	0.3	0.3	0.0000	0.0257	0.942	0.936		0.0577	0.0686	0.388	0.517	0.0577	0.0686	0.388	0.517	-0.0005	0.0686	0.388	0.517

Coverage*: coverage under Fisher's Z scale.

Table 3

Simulation results on estimating the overall BCC (r_o) using our proposed estimator (r_o) and the naive Pearson correlation coefficient (r_p) across scenarios.

True parameters						r_o			r_p	
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias	
-0.6	-0.8	-0.6716	0	0	-0.0016	0.0417	0.948	0.953	-0.0018	
-0.6	-0.8	-0.6716	0	0.3	-0.0015	0.0425	0.949	0.954	-0.0012	
-0.6	-0.8	-0.6716	0.3	0	-0.0016	0.0447	0.94	0.95	-0.0022	
-0.6	-0.8	-0.6716	0.3	0.3	-0.0015	0.0465	0.945	0.951	-0.0020	
-0.6	0	-0.3795	0	0	-0.0007	0.0590	0.951	0.952	-0.0001	
-0.6	0	-0.3795	0	0.3	-0.0005	0.0613	0.953	0.955	0.0003	
-0.6	0	-0.3795	0.3	0	-0.0008	0.0639	0.944	0.954	-0.0004	
-0.6	0	-0.3795	0.3	0.3	-0.0007	0.0676	0.948	0.947	0.0003	
-0.6	0.5	-0.1969	0	0	-0.0001	0.0692	0.951	0.949	0.0010	
-0.6	0.5	-0.1969	0	0.3	0.0001	0.0710	0.947	0.947	0.0012	
-0.6	0.5	-0.1969	0.3	0	-0.0001	0.0736	0.947	0.949	0.0008	
-0.6	0.5	-0.1969	0.3	0.3	-0.0004	0.0771	0.947	0.951	0.0016	
-0.6	0.8	-0.0874	0	0	0.0004	0.0749	0.947	0.949	0.0017	
-0.6	0.8	-0.0874	0	0.3	0.0004	0.0758	0.942	0.943	0.0018	
-0.6	0.8	-0.0874	0.3	0	0.0002	0.0783	0.944	0.948	0.0016	
-0.6	0.8	-0.0874	0.3	0.3	-0.0004	0.0813	0.942	0.944	0.0023	
0	-0.8	-0.2921	0	0	-0.0038	0.0713	0.949	0.952	-0.0045	
0	-0.8	-0.2921	0	0.3	-0.0036	0.0721	0.949	0.956	-0.0034	
0	-0.8	-0.2921	0.3	0	-0.0039	0.0752	0.943	0.956	-0.0056	
0	-0.8	-0.2921	0.3	0.3	-0.0041	0.0783	0.95	0.958	-0.0055	
0	0	0	0	0	-0.0025	0.0717	0.945	0.947	-0.0025	
0	0	0	0	0.3	-0.0022	0.0740	0.953	0.955	-0.0017	
0	0	0	0.3	0	-0.0028	0.0778	0.955	0.955	-0.0033	
0	0	0	0.3	0.3	-0.0029	0.0817	0.951	0.952	-0.0028	
0	0.5	0.1826	0	0	-0.0016	0.0716	0.951	0.952	-0.0011	
0	0.5	0.1826	0	0.3	-0.0014	0.0733	0.95	0.949	-0.0006	
0	0.5	0.1826	0.3	0	-0.0019	0.0769	0.953	0.954	-0.0018	

True parameters										$\hat{r}_o$			$\hat{r}_p$		
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias						
0	0.5	0.1826	0.3	0.3	-0.0023	0.0804	0.953	0.955	-0.0013						
0	0.8	0.2921	0	0	-0.0010	0.0715	0.953	0.952	-0.0002						
0	0.8	0.2921	0	0.3	-0.0008	0.0723	0.947	0.944	0.0000						
0	0.8	0.2921	0.3	0	-0.0013	0.0753	0.95	0.951	-0.0008						
0	0.8	0.2921	0.3	0.3	-0.0020	0.0783	0.953	0.951	-0.0003						
0.3	-0.8	-0.1024	0	0	-0.0046	0.0763	0.955	0.956	-0.0057						
0.3	-0.8	-0.1024	0	0.3	-0.0044	0.0771	0.951	0.95	-0.0044						
0.3	-0.8	-0.1024	0.3	0	-0.0047	0.0801	0.947	0.952	-0.0069						
0.3	-0.8	-0.1024	0.3	0.3	-0.0050	0.0834	0.954	0.955	-0.0069						
0.3	0	0.1897	0	0	-0.0031	0.0687	0.95	0.953	-0.0034						
0.3	0	0.1897	0	0.3	-0.0027	0.0709	0.956	0.955	-0.0025						
0.3	0	0.1897	0.3	0	-0.0034	0.0744	0.953	0.953	-0.0043						
0.3	0	0.1897	0.3	0.3	-0.0035	0.0782	0.955	0.958	-0.0039						
0.3	0.5	0.3723	0	0	-0.0020	0.0633	0.951	0.954	-0.0018						
0.3	0.5	0.3723	0	0.3	-0.0017	0.0650	0.951	0.949	-0.0013						
0.3	0.5	0.3723	0.3	0	-0.0023	0.0684	0.953	0.953	-0.0026						
0.3	0.5	0.3723	0.3	0.3	-0.0027	0.0714	0.954	0.954	-0.0022						
0.3	0.8	0.4819	0	0	-0.0013	0.0602	0.954	0.954	-0.0008						
0.3	0.8	0.4819	0	0.3	-0.0011	0.0610	0.949	0.945	-0.0005						
0.3	0.8	0.4819	0.3	0	-0.0016	0.0638	0.956	0.952	-0.0015						
0.3	0.8	0.4819	0.3	0.3	-0.0022	0.0663	0.958	0.956	-0.0011						
0.6	-0.8	0.0874	0	0	-0.0052	0.0753	0.953	0.953	-0.0065						
0.6	-0.8	0.0874	0	0.3	-0.0050	0.0759	0.951	0.951	-0.0052						
0.6	-0.8	0.0874	0.3	0	-0.0053	0.0786	0.95	0.951	-0.0079						
0.6	-0.8	0.0874	0.3	0.3	-0.0056	0.0817	0.955	0.956	-0.0079						
0.6	0	0.3795	0	0	-0.0034	0.0594	0.947	0.949	-0.0040						
0.6	0	0.3795	0	0.3	-0.0030	0.0616	0.955	0.952	-0.0031						
0.6	0	0.3795	0.3	0	-0.0037	0.0644	0.952	0.951	-0.0049						
0.6	0	0.3795	0.3	0.3	-0.0037	0.0677	0.955	0.957	-0.0046						
0.6	0.5	0.5620	0	0	-0.0021	0.0485	0.953	0.95	-0.0022						

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

True parameters						$\hat{r}_o$			$\hat{r}_p$		
r_b	r_w	r_o	%single	%miss	Bias	RMSE	Coverage	Coverage*	Bias		
0.6	0.5	0.5620	0	0.3	-0.0018	0.0501	0.953	0.953	-0.0017		
0.6	0.5	0.5620	0.3	0	-0.0024	0.0527	0.959	0.952	-0.0029		
0.6	0.5	0.5620	0.3	0.3	-0.0025	0.0552	0.956	0.956	-0.0026		
0.6	0.8	0.6716	0	0	-0.0013	0.0420	0.954	0.953	-0.0011		
0.6	0.8	0.6716	0	0.3	-0.0011	0.0428	0.955	0.95	-0.0008		
0.6	0.8	0.6716	0.3	0	-0.0015	0.0450	0.954	0.951	-0.0016		
0.6	0.8	0.6716	0.3	0.3	-0.0018	0.0468	0.961	0.955	-0.0014		

Coverage*: coverage under Fisher's Z scale.

Table 4

(a–c) Estimations of the within-cluster subject-level BCCs (r_w), cluster-level BCCs (r_b), and overall BCC (r_o) between each pair of the six AD markers with each pair separately fitted by a BLME adjusting for EAO. In the parentheses are 95% Wald CI.

Markers	CSF_tau	Mixed	Paper	PIB_TMP	PIB_PRECUNEUS
(a) r_w					
CSF_Aβ42	-0.2 (-0.56, 0.16)	-0.24 (-0.58, 0.09)	0.18 (-0.2, 0.55)	-0.56* (-0.84, -0.29)	-0.7* (-0.9, -0.5)
CSF_tau		-0.03 (-0.4, 0.35)	0.12 (-0.29, 0.53)	0.18 (-0.21, 0.56)	0.18 (-0.18, 0.55)
Mixed			-0.05 (-0.39, 0.28)	0.27 (-0.11, 0.64)	0.18 (-0.22, 0.57)
Paper				-0.31 (-0.7, 0.07)	-0.32 (-0.69, 0.05)
PIB_TMP					0.9* (0.83, 0.98)
(b) r_b					
CSF_Aβ42	-0.34 (-0.83, 0.15)	-0.29 (-0.96, 0.38)	-0.03 (-0.65, 0.59)	-0.58* (-1.0, -0.02)	-0.53 (-1.0, 0.05)
CSF_tau		-0.34 (-1, 0.32)	-0.48 (-1.24, 0.28)	0.79* (0.13, 1.0)	0.62 (-0.06, 1.31)
Mixed			0.95* (0.5, 1.0)	-0.6 (-1.0, 0.19)	-0.73 (-1.0, 0.13)
Paper				-0.04 (-0.86, 0.77)	-0.17 (-0.92, 0.59)
PIB_TMP					0.85* (0.6, 1.0)
(c) r_o					
CSF_Aβ42	-0.28* (-0.54, -0.02)	-0.25 (-0.52, 0.02)	0.07 (-0.23, 0.36)	-0.55* (-0.75, -0.36)	-0.63* (-0.79, -0.46)
CSF_tau		-0.17 (-0.45, 0.1)	-0.2 (-0.48, 0.09)	0.39* (0.16, 0.61)	0.34* (0.1, 0.57)
Mixed			0.39* (0.17, 0.6)	-0.04 (-0.3, 0.23)	-0.12 (-0.38, 0.14)
Paper				-0.19 (-0.46, 0.07)	-0.25 (-0.51, 0.01)
PIB_TMP					0.89* (0.83, 0.95)

* A significantly nonzero BCC based on the CI and the Wald test.