Journal of Modern Applied Statistical Methods

Volume 21 | Issue 1 Article 1

Robust Trimmed Likelihood Discriminant Analysis for Multivariate Repeated Data

Anita Brobbey
Department of Community Health Sciences, University of Calgary, Canada

Lisa M. Lix
Department of Community Health Sciences, University of Manitoba, Canada

Alberto Nettel-Aguirre
Centre for Health and Social Analytics, National Institute for Applied Statistics Research Australia, University of Wollongong, Australia

Tyler Williamson
Department of Community Health Sciences, University of Calgary, Canada

Samuel Wiebe
Department of Clinical Neurosciences, University of Calgary, Canada
Department of Community Health Sciences, University of Calgary, Canada

Tolulope Sajobi
Department of Community Health Sciences, University of Calgary, Canada
Department of Clinical Neurosciences, University of Calgary, Canada

Recommended Citation

Anita Brobbey, Lisa M. Lix, Alberto Nettel-Aguirre, Tyler Williamson, Samuel Wiebe, Tolulope Sajobi (2022). Robust Trimmed Likelihood Discriminant Analysis for Multivariate Repeated Data. Journal of Modern Applied Statistical Methods, 21(1), https://doi.org/10.56801/Jmasm.V21.i1.1

Journal of Modern Applied Statistical Methods
2022, Vol. 21, No. 1,
Doi: 10.56801/Jmasm.V21.i1.1

ISSN 1538 – 9472

Robust Trimmed Likelihood Discriminant Analysis for Multivariate Repeated Data

Anita Brobbey
Department of Community Health Sciences, University of Calgary, Canada

Lisa M. Lix
Department of Community Health Sciences, University of Manitoba, Canada

Alberto Nettel-Aguirre
National Institute for Applied Statistics Research Australia, University of Wollongong, Australia

Tyler Williamson
Department of Community Health Sciences, University of Calgary, Canada

Samuel Wiebe
Department of Community Health Sciences, University of Calgary, Canada

Tolulope Sajobi
Department of Community Health Sciences, University of Calgary, Canada

Repeated measures discriminant analysis (RMDA) have been developed for distinguishing between two or more independent groups in multivariate repeated measures designs, in which multiple outcomes are repeatedly measured at two or more measurement occasions. However, these models, which are based on structured covariances, rely on the assumption of multivariate normality. Monte Carlo methods were used to compare the accuracy of RMDA procedures based on maximum likelihood estimators and robust maximum trimmed likelihood estimators under a variety of data analytic conditions. RMDA based on robust estimators are recommended for discriminating between population in multivariate repeated measures designs characterized by non-normal distributions.

Keywords: Repeated measures, longitudinal data, robust methods, covariance structure, trimmed estimators, non-normality, Outliers.

1. Introduction

Multivariate repeated measures (MRM) data, in which multiple outcomes are repeatedly measured at two or more occasions, are commonly collected in several disciplines including medicine, ecology, and environmental sciences, where investigators seek to understand changes in multiple correlated outcomes over time or different occasions[1-6]. Multivariate repeated measures data are particularly useful for studying evolutions in subjects' responses over time on multiple characteristics[7].For example, Fieuws and Verbeke[1] reported data on a cohort of patients who having undergone kidney transplant, were longitudinally monitored at irregularly spaced intervals over a 10 year period. The repeated collection of multiple biochemical and physiologicalmarkers, whichconstitute multivariate repeated

measures data, were used to predict 10-year success of graft.Multivariate repeated measures data are inherently challenging to analyze because they are typically characterized by non-Gaussian distributions, and high-dimensional data[8,9]. Classical classification and prediction models developed for data collected in a cross-sectional study are not appropriate to address the complexities observed in multivariate repeated measures data [8, 9] .

Repeated measures discriminant analysis (RMDA), which assume parsimonious mean and covariance structures, have been proposed for discriminating between population groups in MRM data. These procedureshave been primarily developed based on mixed-effects regression models, covariance pattern models, and growth curve models[10-14]. For example, Roy and Khattree developed RMDA procedures based on structured means and Kronecker product variance-covariance matrix of unstructured between-response correlation matrix and compound symmetric (CS) or first-order autoregressive (AR-1) within-response correlation for predicting group membership in MRM data[12,13]. One approach that has been widely used in applied behavioral research is growth curve modeling analysis. Discriminant analysis have been extended to the study of multivariate response curves that can be used to classify a given patient's response curve (example: linear and quadratic shape) to the prognostic group it resembles most[15].However, misspecification of the functional form of the growth curve can potentially lead to biased parameter estimates, misleading conclusions and lower accuracy[16,17]. Similarly, RMDA have been developed based on multivariate linear and non-linear mixed-effects models that assume no structure[1,18] and a Kronecker product structure[6,19-21] for the within-variable and between-variable covariance matrices. RMDA based on mixed-effects models are known to be advantageous in that they can accommodate time-varying and time-invariant covariates in addition to the longitudinally measured outcomes to improve classification accuracy. Generalized linear mixed models have been extended in MRM data studies for different type responses (continuous, counts and binary)[1,2,22]. However, when the number of parameters increases with the sample size and a collection of responses, the random-effect approaches are more likely to be computationally intensive and unstable. Moreover, under misspecification of the random effect parameters commonly assumed as multivariate Gaussian distribution, estimates of the mixed model parameters may become seriously biased[25] and consequently, the performance of the discriminant procedure may also be affected[2].

RMDA procedures assume the data are sampled from a multivariate Gaussian distribution, which may not be tenable. Multivariate repeated measures data are frequently characterized by multivariate skewed or heavy-tailed distributions[23]. So far, there has been limited investigations of RMDA procedures that are robust (i.e., insensitive) to departures from the assumptions of multivariate normality for discriminating between population groups in MRM non-normal data. The lack of these RMDA procedures that are robust to violation of the distributional assumptions of discriminant analysis has limited the application of RMDA in several applied research settings where MRM data are routinely collected (e.g., cancerscreening).

This study develops robust discriminant analysis models forMRM non-normal data. Specifically, we examined the accuracy of RMDA based on maximum trimmed likelihood estimation (MTLE) methods[24,25] under a variety of data-analytic conditions.Multivariate repeated measures data from the Manitoba Inflammatory Bowel Disease (IBD) Cohort Study, was used to demonstrate the application of these procedures.

2. Methodology

2.1 Repeated Measures Data Analysis (RMDA)

Let $\mathbf{y}_{ji} = (\mathbf{y}_{ji\,1}, \mathbf{y}_{ji\,2}, \ldots, \mathbf{y}_{jiq})$ be a $pq \times 1$ vector of $q$outcomes, each repeatedly measured at p occasions for the ith individual in the jth population, sampled from a multivariate normal distribution such that $\mathbf{y}_{ji} \cdot N_{pq}(\boldsymbol{\mu}_j, \boldsymbol{\Omega}_j)$, where$\boldsymbol{\mu}_j$ and$\boldsymbol{\Omega}_j$ is assumed to bepqx 1mean vector and$pqxpq$ positive definite covariance matrix respectively. When $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ are unknown and completely unspecified, a total of $pq+pq(pq+1)/2$ unknown parameters must be estimated. This number increases very rapidly as p and $q$increase. Estimation of so many parameters will require a very large sample, which may not always be feasible. A parsimonious approach to parameter estimation is to assume that $\boldsymbol{\Omega}_j$ has a Kronecker product structure:$\boldsymbol{\Omega}_j = \mathbf{V}_j \otimes \boldsymbol{\Sigma}_j$, where $\mathbf{V}_j$ and $\boldsymbol{\Sigma}_j$ are $p \times p$ and $q \times q$ positive definite matrices respectively, and $\otimes$ denotes the Kronecker product function[12, 13]. The matrix $\mathbf{V}_j$ is the correlation matrix of the repeated measures on a given response variable and it is assumed to be the same for all response variables. The matrix $\boldsymbol{\Sigma}_j$ represents the variance-covariance matrix between the measurements on all response variables at a given time point and this is assumed constant for all time points. Suppose that no structures whatsoever are assumed on $\mathbf{V}$ and $\boldsymbol{\Sigma}$ except that they are positive definite matrices; then, the classifier has the form

$$\lambda(\boldsymbol{y}) = \arg\max_j \ln(\pi_j f_j(\boldsymbol{y})). \tag{1}$$

where

$$f_j(\mathbf{y}) = (2\pi)^{-\frac{pq}{2}} |\mathbf{V}_j|^{-\frac{q}{2}} |\boldsymbol{\Sigma}_j|^{-\frac{q}{2}} \exp\left[-\frac{1}{2} D_j(\mathbf{y}_i)\right] \tag{2}$$

π_j is the prior probability that an observation $\boldsymbol{y}_I$ is from class j, and $D_j(\mathbf{y}_i) = (\mathbf{y}_i - \boldsymbol{\mu}_j)' (\boldsymbol{V}_j^{-1} \otimes \boldsymbol{\Sigma}_j^{-1})(\mathbf{y}_i - \boldsymbol{\mu}_j)$ is the squared Mahalanobis distance between the multiple response vector $\boldsymbol{y}_i$ and the population mean, $\boldsymbol{\mu}_j$. The parameters $\boldsymbol{\mu}_j$, $\mathbf{V}_j$ and $\boldsymbol{\Sigma}_j$ are unknown and should be estimated relying on training samples from the different classes.

Specifically, DA rule for two groups allocates $\boldsymbol{y}_i$ to population 1 if $\lambda_{12}\, \boldsymbol{y}i \leq 0$ where

$$\hat{\lambda}_{12}(\mathbf{y}_i) = D_1^*(\mathbf{y}_i) - D_2^*(\mathbf{y}_i) + 2\log\frac{\hat{\pi}_2}{\hat{\pi}_1} \tag{3}$$

With $D_j^*(\mathbf{y}_i) = (\mathbf{y}_i - \hat{\boldsymbol{\mu}}_j)' \hat{\boldsymbol{\Omega}}_j^{-1} (\mathbf{y}_i - \hat{\boldsymbol{\mu}}_j) + \log|\hat{\boldsymbol{\Omega}}_j|$ are π_1 and π_2 the a priori probabilities that observations belong to populations 1 and 2, respectively. The parameters $\boldsymbol{\mu}_j$, $\mathbf{V}_j$ and $\boldsymbol{\Sigma}_j$ are estimated using MLE. Based on the choice of covariance structures, estimates of the Mahalanobis distance and classification rule can be derived11, 12. The homoscedastic model is obtained when the variance components are homogeneous, that is, $\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_2 = \boldsymbol{\Omega}$, the pooled covariance matrix for j = 1,2. The above classifier implies classification of a unit with multiple response vector $\mathbf{y}_i$ in the first group, if and only if

$$\left(\mathbf{y}_i - \frac{\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2}{2}\right)' (\mathbf{V}^{-1} \otimes \boldsymbol{\Sigma}^{-1})(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) > \log \frac{\pi_2}{\pi_1} \tag{4}$$

which is the linear discriminant analysis (LDA) function, and quadratic discriminant analysis (QDA) function when $\mathbf{V}1 \neq \mathbf{V}2$as

$$(\mathbf{y}_i - \boldsymbol{\mu}_2)' (\mathbf{V}_2^{-1} \otimes \boldsymbol{\Sigma}_2^{-1})(\mathbf{y}_i - \boldsymbol{\mu}_2) - (\mathbf{y}_i - \boldsymbol{\mu}_1)' (\mathbf{V}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1})(\mathbf{y}_i - \boldsymbol{\mu}_1) > \log \left|\frac{\boldsymbol{\Omega}_1}{\boldsymbol{\Omega}_2}\right| + 2 \log \frac{\pi_2}{\pi_1} \tag{5}$$

However, conventional RMDA procedures rely on the assumption of multivariate normal distribution, which may not betenable in multivariate repeated measures data, which are usually characterized by non-normal distributions[2,26].

2.2 Robust RMDA

An alternate approach to overcome these limitations involves the development of robust RMDA procedures based on maximum trimmed likelihood estimators (MTLE) ofmean, $\boldsymbol{\mu}_j$ and covariance components, $\mathbf{V}_j$ and $\boldsymbol{\Sigma}_j$. In MTLE,the contribution each $\boldsymbol{yi}$to the (log)likelihood function scores ($\ell\ \boldsymbol{\theta}$; $\boldsymbol{yi}$ are ranked from the smallest to the highest. Loglikelihood function scores at the extreme tails are assigned smaller or no weights. Depending on the weights assigned to observations at the tails, different robust estimators could be derived. Specifically, in this study we developed robust RMDA based on minimum covariance determinant (MCD) and minimum volume ellipsoid (MVE) estimators, which are special cases ofMTLE[24, 25].

For any given value of$\boldsymbol{\theta}$:

$$\ell(\boldsymbol{\theta}; \mathbf{y}_1) \geq \ell(\boldsymbol{\theta}; \mathbf{y}_2) \geq \cdots \geq \ell(\boldsymbol{\theta}; \mathbf{y}_n) \tag{6}$$

where $\ell\ (\boldsymbol{\theta}; \boldsymbol{y}_i) = \ln f(\boldsymbol{y}_i; \boldsymbol{\theta})$ is the contribution of the ith observation to the log-likelihood function.

Note, the original indices of the observations may not satisfy the likelihood ordering in (6) for all values of $\boldsymbol{\theta}$. If, for a given value of $\boldsymbol{\theta}$, the above ordering is not satisfied, the indices of the observations can be changed so that (6) is satisfied[24,25]. The ordering of the observations may be different for different values of $\boldsymbol{\theta}$. The trimmed log likelihood function is given as

$$\sum_{i=1}^{h} \boldsymbol{w}_i \ell(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{7}$$

where h is the trimming parameterand $w_i \geq 0$ are weights. The MTLE $\boldsymbol{\theta}$ (h, w) is obtained by maximizing the trimmed log likelihood function. The key idea is to trim the $nj - h$ points that are the most unlikely from the estimation of the likelihood function. Special cases of MTLE includes MLE, MCD, and MVE when win = 1. The parameter h has to be set manually.A recommendable and common choice for h that yields maximum breakdown is: h=(n_j + pq + 1)/2, which asymptotically reaches half of the data, but any integer h within the interval [(n_j + pq + 1)/2, nj] can be chosen[27].Some researchers have investigated 5%-25% trimming[41-43]. Even though one research argue 20% trimming[42], another recommend no more than 10% trimming to achieve optimal results[43].The farther h is from n_j, the more robust but the less efficient are the estimators.When $h = n_j$, we obtain the MLE of $\boldsymbol{\theta}$ and$d$ for $h < n_j$, the MCD and MVE estimators are MTLEs of $\boldsymbol{\theta} = (\boldsymbol{\mu}_h, \boldsymbol{\Omega}_h)$ for h observations yielding the desired robust estimates[24,28]. The MCD approachis similar to the MVE in that it searches for a portion of the data that minimizes the impact of outlying observations on estimation of means and covariance parameters. However, whereas MVE seeks to minimize the volume of an ellipsoid created by the retained data, whiles MCD minimizes the determinant of the variance-covariance matrix. For example; the location estimate of MVE is the center of the minimum volume ellipsoid covering (at least) h of the data whiles for MCD, the location estimate is the mean of h of the data for which the determinant of the covariance matrix is minimal.

3. Simulation Study

A Monte Carlo simulation study was conducted to examine the accuracy of robustRMDA procedures in comparison to the conventional RMDA based on MLE estimators. Specifically we investigated the following procedures: (a) RMDA that assumes structured means and Kronecker correlation matrix of unstructured between- responses and within-response AR-1 correlation matrices (st-UNAR), (b) RMDA that assumes unstructured means and Kronecker correlation matrix of unstructured between-responses and within-response AR-1 correlation matrices(un-UNAR), (c) RMDA that assumes structured means and Kronecker correlation matrix of structured between-responses and within-response CS correlation matrices (st-UNCS) and (d) RMDA that assumes unstructured means and Kronecker correlation matrix of between-responses and within-response CS correlation matrices(un-UNCS). The parameters of eachRMDA procedure were estimated using MLE, MVE and MCD estimators. Moreover, repeated measures LDA was used for classification when group covariances were homogeneous, while repeated measures QDA was adopted when group covariances were heterogeneous.

The following simulation conditions were investigated: (a) number of different outcomes (q), (b)number of repeated occassions(p), (c) total sample size (n), (d) group sizes(n_1, n_2), (e) Covariance pattern and magnitude of correlation among the repeated measurements(ρ), (f) mean configuration, (g) Covariance heterogeneity, and (h) population distribution. All procedures were investigated for two independent groups.

The number of repeated measurements was set at p = 3, and 5 whilst the number of different outcomes was set at q = 3, and 7. Previous studies about RMDA procedures have considered p ranging from 3 to 10, an increase in classification accuracy was quite significant when p increases from three to five[11,12]. Total sample sizes of n = 100, 140 and 200 were investigated. This is consistent with previous simulation studies that examined the accuracy of RMDA based on parsimonious covariance structures with $n$ranging between 60 and 500.Moreover, consistent with previous studies, we examined the impact of equal and unequal group sizes[11,12,29,30,31,32]. For n = 100, we set $n1$, $n2$ = 50, 50 , and (40, 60).Similar equal (1:1) and unequal (2:3) group size ratios were investigated when n = 140 and 200considering small-sampled studies for multivariate RMDA procedures[12,13,33-35]. A variety of mean configurationsof different forms have been previously investigated in the development of RMDA procedures[11,12]. In this study, four mean configuration structures for population 1 ($\boldsymbol{\mu}1$) were selected for each pair ofp and q. (see Table 1) and mean configuration structures for population 2($\boldsymbol{\mu}2$) was the null vector for all conditions[11,12,36].

Table 1. Four mean configuration structures assumed for population 1 ($\boldsymbol{\mu}1$) in the Monte Carlo Study

	p	*q*=3	*q*=7
Configuration			
I		$\mathbf{1}_p \otimes$ (30,30,30)	$\mathbf{1}_p \otimes$ (30,30,30,30,30,30,30)
II	3 or 5	$\mathbf{1}_p \otimes$ (27,29,31)	$\mathbf{1}_p \otimes$ (30,31,32,33,34,35,36)
III		$\mathbf{1}_p \otimes$ (30,25,30)	$\mathbf{1}_p \otimes$ (30,27,24,21,24,27,30)
IV	3	(1,1.1,1.2) $\otimes$ (30,25,30)	(1,1.1,1.2)$\otimes$ (30,27,24,21,24,27,30)
	5	(1,1.1,1.2,1.3,1.4)$\otimes$ (30,25,30)	(1,1.1,1.2,1.3,1.4)$\otimes$ (30,27,24,21,24,27,30)

Note: For population 2, $\boldsymbol{\mu}2 = \mathbf{1}$p $\otimes$ 25$\mathbf{1}$q for all conditions; q=number of outcome variables; p=number of repeated occasions

The descriptions of the four configurations for $\boldsymbol{\mu}1$ in Table 1 are as follows; configuration I- III had no change in mean pattern over time ($\mathbf{1}p$) for constant,

monotonic increasing and quadratic mean patterns among repeated outcomes respectively, and configuration IV was assumed to have monotonic increasing mean pattern over time (for example; (1,1.1,1.2)) for non-constant means among the repeated measurements. Furthermore, the accuracy of RMDA procedures is known to be influenced by the magnitude and pattern of within- and between-variable correlations[37]. Therefore, we investigated the following components of the assumed Kronecker variance-covariance matrices: $\boldsymbol{\Omega}j = \mathbf{V}j \otimes \boldsymbol{\Sigma}j$, j = 1, 2where $\boldsymbol{\Sigma}j$ was assumed to be a q × q unstructured variance-covariance matrix with a common variance of 60 among outcome variables and the p × p correlation matrix $\mathbf{V}j = \mathbf{V}j$ (ρj) was assumed to follow a AR-1 or CS structure with ρj chosen as 0.3 and 0.7, representing moderate to strong autocorrelation in the data[11,12] (See Table 2 for more details).

Table 2. Configuration of unstructured between-responses covariance matrix Σ_1 given within-response correlation coefficient (ρ) for the Monte Carlo Study

Within-response correlation coefficient *(ρ)*	0.3	0.7
***q*=3**	$\boldsymbol{\Sigma}_1 = 60\begin{pmatrix} 1 & 0.15 & 0.3 \\ 0.15 & 1 & 0.45 \\ 0.3 & 0.45 & 1 \end{pmatrix}$	$\boldsymbol{\Sigma}_1 = 60\begin{pmatrix} 1 & 0.65 & 0.66 \\ 0.65 & 1 & 0.7 \\ 0.66 & 0.7 & 1 \end{pmatrix}$
***q*=7**	$\boldsymbol{\Sigma}_1 = 60\begin{pmatrix} 1 & 0.35 & 0.2 & 0.2 & 0.3 & 0.25 & 0.3 \\ 0.35 & 1 & 0.3 & 0.3 & 0.37 & 0.32 & 0.34 \\ 0.2 & 0.3 & 1 & 0.31 & 0.35 & 0.32 & 0.34 \\ 0.2 & 0.3 & 0.31 & 1 & 0.3 & 0.25 & 0.33 \\ 0.3 & 0.37 & 0.35 & 0.3 & 1 & 0.28 & 0.34 \\ 0.25 & 0.32 & 0.32 & 0.25 & 0.28 & 1 & 0.35 \\ 0.3 & 0.34 & 0.34 & 0.33 & 0.34 & 0.35 & 1 \end{pmatrix}$	$\boldsymbol{\Sigma}_1 = 60\begin{pmatrix} 1 & 0.65 & 0.66 & 0.7 & 0.72 & 0.65 & 0.75 \\ 0.65 & 1 & 0.7 & 0.7 & 0.67 & 0.72 & 0.74 \\ 0.66 & 0.7 & 1 & 0.71 & 0.75 & 0.72 & 0.74 \\ 0.7 & 0.7 & 0.71 & 1 & 0.7 & 0.65 & 0.73 \\ 0.72 & 0.67 & 0.75 & 0.7 & 1 & 0.68 & 0.74 \\ 0.65 & 0.72 & 0.63 & 0.65 & 0.68 & 1 & 0.75 \\ 0.75 & 0.74 & 0.74 & 0.73 & 0.74 & 0.75 & 1 \end{pmatrix}$

Note: $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ or $\boldsymbol{\Sigma}_1 = 3\boldsymbol{\Sigma}_2$

In order to assess the performance of the discriminant function, random samples were generated from both multivariate normal and multivariate non-normal distributions. With specified mean, $\boldsymbol{\mu}_j$ and covariance matrix $\boldsymbol{\Omega}_j$, pq-variate normal distribution populations were generated using the mvrnorm function from the MASS R package[38], multivariate lognormal distribution were generated using the rlnorm function from the compositions R package[39], multivariate t distribution were generated using the rmvt function from the mvnfast R package[40], and multivariate Cauchy distribution were generated using the rmsc function from the sn R package[40].

For robust estimators, the proportion of trimmed data was fixed at 10% symmetric trimming. The FASTMCD and FASTMVE algorithms[44-46] were used to define a subsample of observations for the trimmed means and covariances. More specifically, robust estimates of the LDA and QDA procedures that assumed unstructured or structured means and structured covariances were derived by maximizing the likelihood of the 90% best subsample of original observations using the fast algorithms. These means and covariance have high robustness properties[44, 47].

Fixed-effects analysis of variance (ANOVA) model was used to assess the relative importance of different simulation factors on the variations in the average classification accuracy for each procedure[48,49]. The percentage of explained variance attributable to each main effect and interactions were evaluated using η^2, an R^2 equivalent in regression analysis[50]. The classification was performed on the generated samples from each of the two populations. Some of the previous classification research have employed the accuracy or the error rate (1-accuracy) metric to discriminate between two or groups[11,12,29]. Thus, the overall classification accuracy (correctly classified / total sample) was used as performance metric in this study and the standard errors were also calculated. For each procedure and each method of estimation, a total of 1440 combination of simulation factors was investigated. There were 1000 replications for each combination. The Monte Carlo study was conducted using R version 3.6.3.

4. Results

4.1 Simulation Study Results

Table 3 describes the relative contribution of each of simulation conditions on the overall accuracy of the RMDA procedures. Specifically, population distribution, mean configurations, group covariance ratio, and estimation method accounted for more than 70% of the total variation in the classification accuraciesof the RMDA procedures. Therefore, our description of the simulation results will focus on these important simulation factors.

Tables 4 and 5 describe the average predictive accuracies and standard errors of repeated measures LDA and repeated measures QDA based on MLE and MVE, respectively, by population distribution and number of outcomes. The average classification accuracy of the RMDA procedures were highest when the data were sampled from a multivariate normal distribution and lowest when the data were sampled from extremely heavy-tailed distribution, regardless of the type of estimation method adopted, number of outcome variables, or mean configuration. In particular, there were negligible differences in the average classification accuracy of RMDA procedures based on MLE and those based on robust estimators when the data were sampled from a multivariate normal distribution, or a moderately

multivariate heavy-tailed distribution. However, the robust estimators procedures were more accurate than MLE when the data were sample from a multivariate heavy-tailed distribution. For example, the average classification accuracy of st-UNCS based on MLE and MVE were 0.86 and 0.85 when q = 7 and data were sampled from a multivariate normal distribution with outcome variables. Whereas, the average accuracy of the former and latter procedures were 0.54 and 0.70 when q = 7 and the data were sampled from a multivariate Cauchy distribution, respectively (Table 4). Similar patterns were observed in repeated measures quadratic discriminant analysis procedures (Table 5).

Table 3. Proportion of Variance in Overall Classification Accuracy Explained by the Simulation Factors

Simulation Factor	un-UNAR	un-UNCS	st-UNAR	st-UNCS
p	*	0.3	*	0.2
q	0.6	0.6	1.1	0.6
Covariance structure	*	*	*	*
ρ	2.0	1.2	1.2	1.4
n	*	*	*	*
Mean Configuration	10.5	11.0	9.2	10.5
Population Distribution	51.4	49.8	56.0	50.2
Covariance ratio (QDA vs LDA)	0.9	*	3.3	1.1
Estimation (MLE vs MVE)	0.6	1.7	1.5	1.5
Population Distribution x Covariance ratio	5.1	4.9	3.9	5.0
Population Distribution x Mean Configuration	5.0	5.2	4.1	5.0

Note: *= Estimated percentage of variation explained close to zero; C.I = confidence interval; p= number of repeated occasions; p = number of responses; ρ= coefficient of correlation; n =sample size; un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid; LDA= Linear Discriminant Analysis; QDA= Quadratic Discriminant Analysis

Table 4. Overall Mean Accuracy of Repeated Measures LDA procedures based on MLE and Robust Estimator, Estimator, MVE (standard error) by population distribution, Number of Outcomes for equal group covariance

Distribution	q	MLE				MVE			
		un-UNAR	un-UNCS	st-UNAR	st-UNCS	un-UNAR	un-UNCS	st-UNAR	st-UNCS
Normal	3	0.77(0.03)	0.77(0.03)	0.77(0.03)	0.78(0.03)	0.77(0.03)	0.77(0.03)	0.76(0.03)	0.78(0.03)
	7	0.84(0.03)	0.85(0.03)	0.84(0.03)	0.86(0.03)	0.84(0.03)	0.85(0.03)	0.84(0.03)	0.85(0.03)
T	3	0.77(0.03)	0.77(0.03)	0.77(0.03)	0.78(0.03)	0.77(0.03)	0.77(0.03)	0.76(0.03)	0.77(0.03)
	7	0.84(0.03)	0.85(0.03)	0.84(0.03)	0.86(0.03)	0.84(0.03)	0.85(0.03)	0.84(0.03)	0.85(0.03)
Lognormal	3	0.54(0.03)	0.53(0.03)	0.54(0.03)	0.53(0.03)	0.56(0.04)	0.56(0.04)	0.56(0.04)	0.56(0.04)
	7	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.57(0.04)	0.57(0.04)	0.57(0.04)	0.57(0.04)
Cauchy	3	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.70(0.04)	0.70(0.04)	0.70(0.04)	0.70(0.04)
	7	0.56(0.06)	0.56(0.06)	0.56(0.06)	0.56(0.07)	0.73(0.05)	0.73(0.05)	0.73(0.05)	0.73(0.05)

Note: un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid; LDA= Linear Discriminant Analysis

Table 5. Overall Mean Accuracy of Repeated Measures QDA procedures based on MLE and Robust Estimator, MVE (standard error) by population distribution, Number of outcomes for unequal group covariance

Distribution	q	MLE				MVE			
		un-UNAR	un-UNCS	st-UNAR	st-UNCS	un-UNAR	un-UNCS	st-UNAR	st-UNCS
Normal	3	0.83(0.04)	0.83(0.03)	0.85(0.03)	0.83(0.03)	0.82(0.04)	0.83(0.03)	0.84(0.03)	0.83(0.03)
	7	0.80(0.04)	0.81(0.03)	0.86(0.03)	0.81(0.03)	0.79(0.04)	0.81(0.03)	0.86(0.03)	0.80(0.03)
T	3	0.83(0.04)	0.83(0.03)	0.85(0.03)	0.83(0.03)	0.82(0.04)	0.83(0.03)	0.84(0.03)	0.83(0.03)
	7	0.80(0.04)	0.81(0.03)	0.86(0.03)	0.81(0.03)	0.79(0.04)	0.81(0.03)	0.86(0.03)	0.80(0.03)
Lognormal	3	0.65(0.11)	0.69(0.04)	0.68(0.05)	0.69(0.04)	0.68(0.07)	0.68(0.04)	0.67(0.05)	0.68(0.04)
	7	0.70(0.04)	0.70(0.04)	0.70(0.04)	0.70(0.04)	0.69(0.04)	0.69(0.04)	0.69(0.04)	0.69(0.04)
Cauchy	3	0.53(0.08)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.66(0.05)	0.66(0.05)	0.66(0.05)	0.66(0.05)
	7	0.55(0.06)	0.56(0.07)	0.56(0.06)	0.56(0.07)	0.67(0.06)	0.67(0.06)	0.67(0.06)	0.67(0.06)

Note: un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid; QDA= Quadratic Discriminant Analysis

Furthermore, when the data were sampled from a multivariate normal distribution, the average accuracy of each repeated measures LDA procedures increased as the number of outcomes increased, regardless of the method or estimation. However, the increase in classification accuracy as q increased was smaller when the data were

sampled from a multivariate non-normal distribution. For example, when the data were sampled from a multivariate normal distribution, the average increase in classification accuracy of the un-UNAR procedure based on MLE and MVE was about 0.07 as q increased from 3 to 7. But when data were sampled from a multivariate lognormal distribution, there were negligible differences in the classification accuracies for these procedures as q increased. In contrast, the average classification accuracy of the repeated measures QDA procedures decreased as q increased for almost all the investigated population data distributions, except for the multivariate lognormal and Cauchy distributions (Table 5).

However, for multivariate lognormal and Cauchy distributions, smaller to no change in average classification accuracy was observed for all procedures for MLE. For the un-UNAR procedure when data were sampled from multivariate lognormal, 0.54 average classification accuracy was observed when number of responses were both three and seven, whereas for data sampled from multivariate Cauchy distribution average classification accuracy were 0.54 and 0.56, respectively for MLE (Table 4). In contrast, the decreased in average classification accuracy for non-normal distributions were much lower for RMDA based on MLE compared to the robust methods. Moreover, higher average classification accuracies were observed in these non-normal distributions under robust methods. For the un-UNAR procedure under MVE, when data were sampled from multivariate lognormal, 0.56 average accuracy was observed when number of responses, $q = 3$ compared to 0.57 accuracy when $q = 7$, whereas for data sampled from multivariate Cauchy distribution average accuracies were 0.70 and 0.73, respectively (Table 4). Again, similar observations were seen for all procedures in robust methods. For both MLE and robust methods average accuracies, smaller to no change was observed for structured and unstructured mean for all procedures, irrespective of population distributions and number of responses.

For Table 5, when QDA classifier was adopted for unequal group covariance, opposite effect of the number of responses seen in Table 4 was observed, that is $q = 3$ had a higher average accuracythan $q = 7$ in normal distribution and t-distribution for un-UNAR, un-UNCS and st-UNCS except st-UNAR for both MLE and robust methods. For example: for the un-UNAR procedure under MVE, the average accuracy for $q = 3$ was 0.82, whereas for $q = 7$, it was 79, whiles for the st-UNAR procedure, the average accuracy for $q = 3$ was 0.84, whereas for $q = 7$, it was 0.86. Also, higher average accuracies were observed for the st-UNAR procedure compared to the other procedures. In addition, we observed higher increase in classification accuracy for increase number of responses in Table 5 for multivariate lognormal distribution under MLE compared to Table 4. For example, for the un-UNAR procedure under MLE, when data were sampled from multivariate lognormal, 0.65 accuracy was observed when number of responses, $q = 3$ compared to 0.70 accuracy rate when $q = 7$. With regards to MLE and robust methods, higher average accuracies were observed for Cauchy distribution based on robust methods compared to MLE, but smaller to no increase average accuracies were observed in other population distributions (Table 4).

Tables 6 and 7 describe the overall average classification accuracy of the repeated measures LDA and QDA procedures by population distribution and mean configuration, respectively. Figures 2 and 3 show the overall average accuracy of robust RMDA procedures for non-normal population distributions from Tables 6 and 7 respectively. There were negligible differences in the accuracy of the RMDA based on MLE and robust estimators when the data were sampled from multivariate normal or multivariate t distribution. However, the robust procedures were significantly more accurate than MLE when the data were sampled from multivariate Cauchy distribution. For example, the average classification accuracy of st-UNAR based on MLE and robust estimators were 0.72 and 0.71, when the data were sample from a multivariate t-distributions with mean configuration I, respectively. Whereas the average accuracy of the former and latter procedures were 0.53 and 0.64 when the data were sampled from a multivariate Cauchy distribution with the same mean configuration I. On the other hand, the impact of choice of mean configuration on the accuracy of the repeated measures LDA models was confounded by the population distribution. Specifically, when the data were sampled from a multivariate normal distribution, the average classification accuracy of the repeated measures LDA procedures were lowest under mean configuration I, which assumed no change in mean pattern over time for constant mean among repeated outcomes, but highest under mean configuration IV, which assumed unstructured means among the repeated measurements regardless of the estimation methods. However, there were negligible differences in the classification accuracy of the procedures based on MLE estimators across all the mean configurations when the data were sampled from a multivariate log-normal or Cauchy distribution. In contrast, accuracy of the procedures based on robust estimators varied across mean configurations when the data were sampled from a multivariate lognormal or multivariate Cauchy distribution. For example, the average accuracy of the st-UNAR procedure based on MLE increased by 0.16 (were 0.72 and 0.88) across the mean configurations when the data were sampled from multivariate normal distribution, whereas there was negligible change in average accuracy of this procedure across the mean configurations when the data were generated from a multivariate log-normal distribution. In contrast, the change in average classification accuracy for st-UNAR procedure based on robust estimators across the mean configurations were 0.16 and 0.13 when the data were sampled from multivariate normal and multivariate Cauchy distributions, respectively (Table 6). Thus, the average accuracy of the procedures based on robust estimators increased across the mean configurations when the data were sampled from multivariate Cauchy distribution compared to procedures based on MLE (Table 6 and Figure 2). Similar results were obtained for repeated measures quadratic discriminant analysis (Table 7 and Figure 3). Results for RMDA based MCD and MVE were similar, hence we reported results for MVE estimation to avoid repetition.

Table 6. Overall Mean Accuracy of repeated Measures LDA procedures based on MLE and Robust Estimator, MVE (standard error) by population distribution, mean configuration for equal group covariance

Distribution	Mean Configuration	MLE				MVE			
		un-UNAR	un-UNCS	st-UNAR	st-UNCS	un-UNAR	un-UNCS	st-UNAR	st-UNCS
Normal	I	0.72(0.03)	0.72(0.04)	0.72(0.03)	0.73(0.04)	0.71(0.04)	0.72(0.04)	0.71(0.03)	0.72(0.04)
	II	0.76(0.03)	0.78(0.03)	0.76(0.03)	0.79(0.03)	0.76(0.03)	0.78(0.03)	0.76(0.03)	0.78(0.03)
	III	0.79(0.03)	0.77(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)
	IV	0.88(0.02)	0.89(0.02)	0.88(0.02)	0.89(0.02)	0.88(0.03)	0.89(0.03)	0.87(0.03)	0.89(0.03)
T	I	0.72(0.03)	0.72(0.04)	0.72(0.03)	0.73(0.04)	0.71(0.04)	0.72(0.04)	0.71(0.03)	0.72(0.04)
	II	0.76(0.03)	0.78(0.03)	0.76(0.03)	0.79(0.03)	0.76(0.03)	0.79(0.03)	0.76(0.03)	0.79(0.03)
	III	0.79(0.03)	0.77(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)	0.79(0.03)
	IV	0.88(0.02)	0.89(0.02)	0.88(0.03)	0.89(0.02)	0.88(0.03)	0.89(0.03)	0.87(0.03)	0.89(0.03)
Lognormal	I	0.54(0.03)	0.53(0.03)	0.54(0.03)	0.54(0.03)	0.56(0.04)	0.56(0.04)	0.56(0.04)	0.56(0.04)
	II	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.56(0.04)	0.56(0.04)	0.56(0.04)	0.56(0.04)
	III	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.54(0.03)	0.56(0.04)	0.56(0.04)	0.56(0.04)	0.56(0.04)
	IV	0.54(0.04)	0.54(0.03)	0.54(0.04)	0.54(0.03)	0.57(0.04)	0.57(0.04)	0.57(0.04)	0.57(0.04)
Cauchy	I	0.53(0.04)	0.53(0.04)	0.53(0.04)	0.53(0.04)	0.64(0.04)	0.64(0.04)	0.64(0.04)	0.64(0.05)
	II	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.69(0.05)	0.69(0.05)	0.68(0.05)	0.69(0.05)
	III	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.69(0.05)	0.69(0.05)	0.69(0.05)	0.69(0.05)
	IV	0.57(0.04)	0.57(0.03)	0.57(0.04)	0.57(0.03)	0.78(0.04)	0.78(0.04)	0.77(0.04)	0.78(0.04)

Note: un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid; MVE= minimum volume ellipsoid

Table 7. Overall mean Accuracy of Repeated Measures QDA procedures based on MLE and Robust Estimator, MVE (standard error) by population distribution, mean configuration for unequal group covariance

Distribution	Mean Configuration	MLE				MVE			
		un-UNAR	un-UNCS	st-UNAR	st-UNCS	un-UNAR	un-UNCS	st-UNAR	st-UNCS
Normal	I	0.69(0.03)	0.70(0.02)	0.76(0.03)	0.70(0.02)	0.69(0.03)	0.69(0.03)	0.75(0.03)	0.69(0.03)
	II	0.78(0.04)	0.77(0.04)	0.82(0.03)	0.77(0.04)	0.77(0.04)	0.76(0.04)	0.81(0.03)	0.77(0.04)
	III	0.81(0.04)	0.80(0.04)	0.81(0.03)	0.80(0.04)	0.80(0.04)	0.80(0.04)	0.81(0.03)	0.80(0.04)
	IV	0.90(0.04)	0.92(0.03)	0.94(0.02)	0.92(0.03)	0.90(0.05)	0.92(0.03)	0.94(0.02)	0.92(0.03)
T	I	0.69(0.03)	0.70(0.02)	0.76(0.03)	0.70(0.03)	0.69(0.03)	0.69(0.03)	0.75(0.03)	0.69(0.03)
	II	0.78(0.04)	0.77(0.04)	0.82(0.03)	0.77(0.04)	0.77(0.04)	0.76(0.04)	0.81(0.04)	0.77(0.04)
	III	0.81(0.04)	0.80(0.04)	0.81(0.03)	0.80(0.04)	0.80(0.04)	0.80(0.04)	0.81(0.04)	0.80(0.04)
	IV	0.90(0.04)	0.92(0.03)	0.94(0.02)	0.92(0.03)	0.90(0.05)	0.92(0.03)	0.94(0.02)	0.92(0.03)
Lognormal	I	0.68(0.08)	0.69(0.04)	0.69(0.04)	0.69(0.04)	0.68(0.05)	0.68(0.04)	0.68(0.04)	0.68(0.04)
	II	0.68(0.07)	0.70(0.04)	0.69(0.04)	0.70(0.04)	0.69(0.05)	0.68(0.04)	0.68(0.04)	0.68(0.04)
	III	0.67(0.07)	0.68(0.04)	0.68(0.04)	0.68(0.04)	0.67(0.05)	0.67(0.04)	0.67(0.04)	0.67(0.04)
	IV	0.68(0.08)	0.70(0.04)	0.70(0.05)	0.70(0.04)	0.69(0.05)	0.69(0.04)	0.68(0.04)	0.69(0.04)
Cauchy	I	0.53(0.06)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.62(0.05)	0.62(0.05)	0.63(0.05)	0.62(0.05)
	II	0.54(0.07)	0.54(0.05)	0.55(0.05)	0.55(0.05)	0.64(0.05)	0.64(0.05)	0.65(0.05)	0.64(0.05)
	III	0.54(0.07)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.65(0.05)	0.65(0.05)	0.64(0.05)	0.65(0.05)
	IV	0.55(0.08)	0.56(0.04)	0.56(0.05)	0.56(0.03)	0.70(0.05)	0.71(0.04)	0.71(0.04)	0.71(0.04)

Note: un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid; QDA= Quadratic Discriminant Analysis

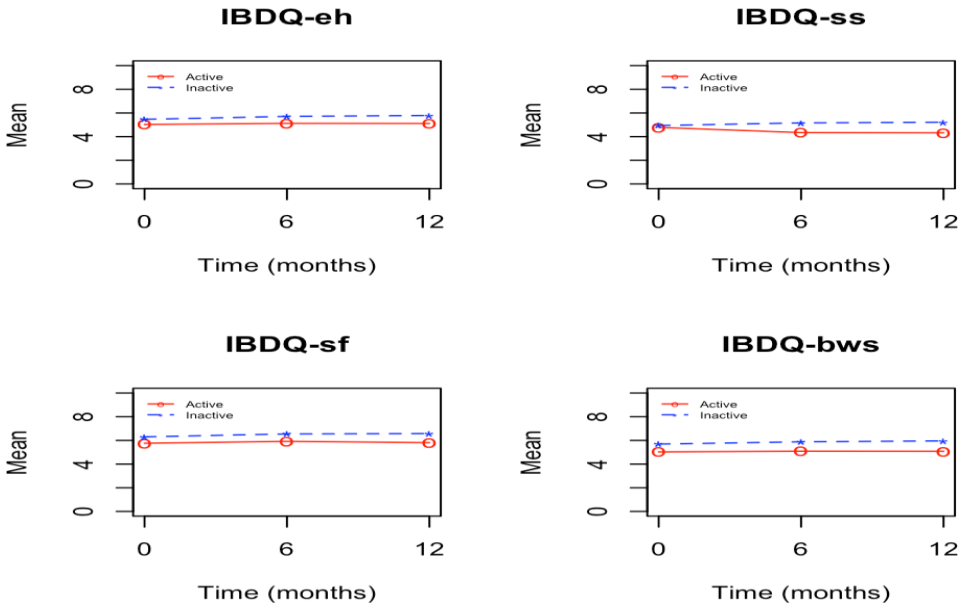

Figure 1. Observed mean longitudinal profiles of an indicator of whether a participant had active (Red) or inactive (Blue) IBD in each of the four IBDQ domains: emotional health (IBDQ-eh), systematic symptoms (IBDQ-ss), social function (IBDQ-sf) and bowel symptoms (IBDQ-bws)

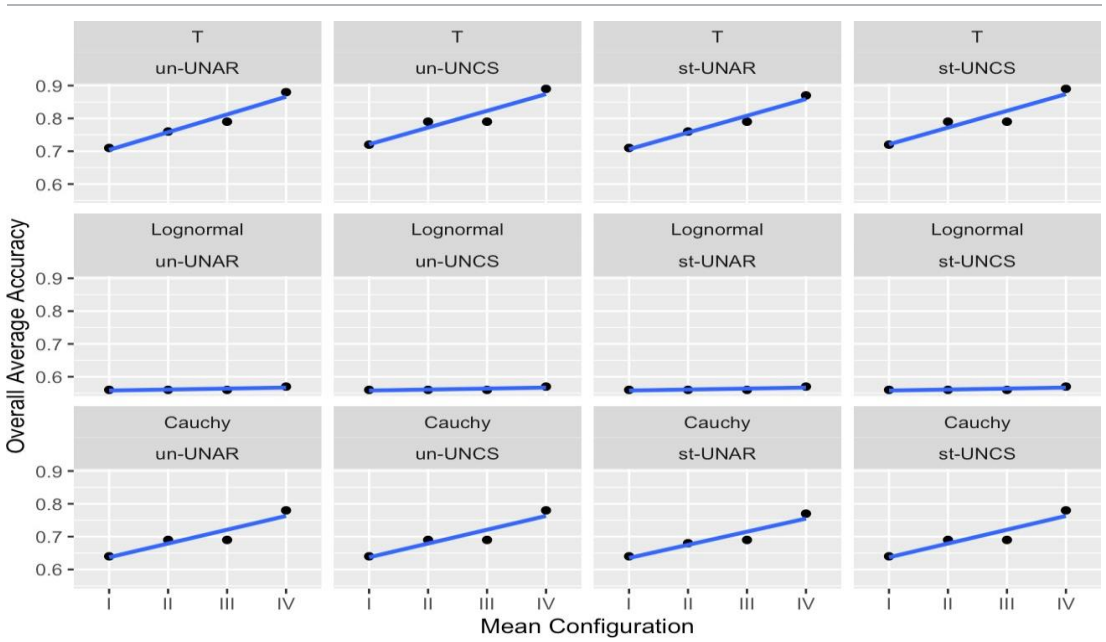

Figure 2: Overall Average Accuracy of Robust repeated measures LDA procedures by non-normal population distributions, mean configuration for equal group covariance

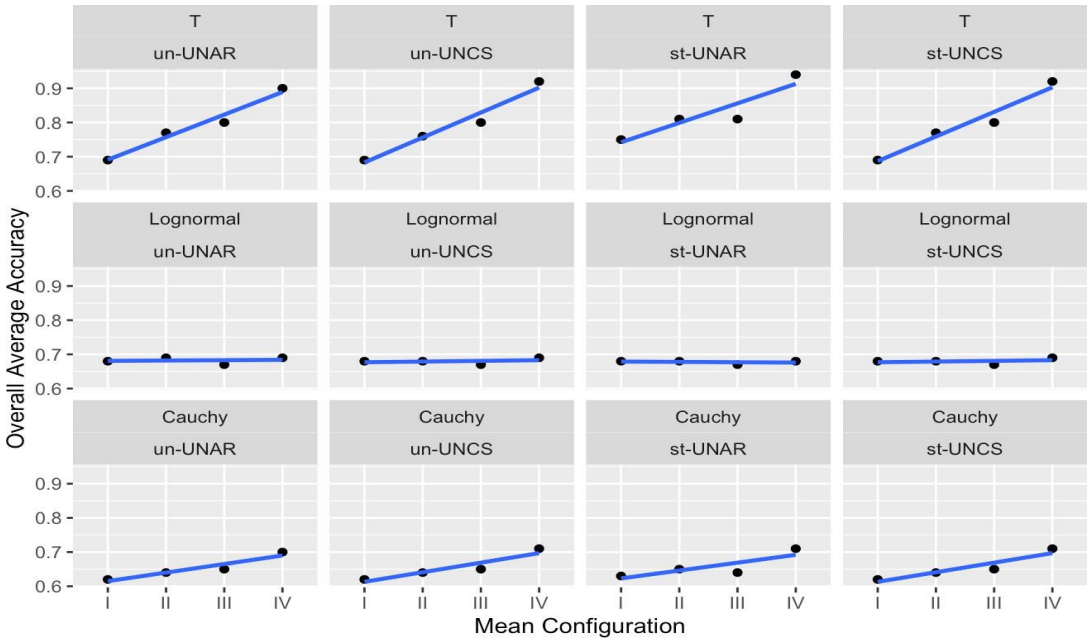

Figure 3: Overall Average Accuracy of Robust repeated measures QDA procedures by non-normal population distributions, mean configuration for unequal group covariance

In addition, we investigated the influence of class imbalance on the accuracy of these proposed models, for which additional simulation condition results are provided in the Appendix. Tables 10 and 11 contain class-specific accuracies of repeated measures LDA procedures based on MLE and robust MVE by number of outcomes and sample sizes for normal and Cauchy distributions respectively. Conclusions and observations from the additional simulation results remained the same as the initial simulations. Thus, class imbalance did not influence the proposed repeated measures models.

4.2 Manitoba Inflammatory Bowel Disease Study

Multivariate repeated measures data from the Manitoba Inflammatory Bowel Disease (IBD) Cohort Study, a prospective longitudinal cohort study to investigate the determinants of disease outcomes in community dwelling individuals living with Crohn's disease or ulcerative colitis, were used to demonstrate the application of these methods. Data were collected at six-month intervals, after baseline, using self-report instruments. Study participants were rated as having active (n1 = 214) or inactive (n2 = 127) disease based on self-reported IBD symptoms at study entry. Details about the Manitoba IBD Cohort Study have been previously published elsewhere[51,37]. Differences between active and inactive disease groups on a disease-specific measure of quality of life, the IBD questionnaire (IBDQ)[29], were

investigated in the first year of the study (i.e., three measurement occasions, at baseline (0 months), 6 months, and 12 months, p=3). The primary research question is to be able to identify active and inactive disease groups within one year of diagnosis using their longitudinal profiles of quality of life. Multivariate repeated measures data collected on the four IBDQ domains (q=4) namely emotional health (IBDQ-eh), systematic symptoms (IBDQ-ss), social function (IBDQ-sf) and bowel symptoms (IBDQ-bws) over the one-year period were used to discriminate between both groups of participants.

Of the 389 participants who provided data at baseline (month 0), 213 had complete IBDQ domains at the end of the first year. Among the 213 participants, 133 were participants with active IBD. Table 8 and Figure 1 describe the differences on each domains for active and inactive participants in the Manitoba IBD Cohort Study. Participants with inactive disease had higher quality of life scores on all four domains than participants with active disease (Figure 1). The group means and descriptive measures of multivariate skewness and kurtosis for the IBDQ data are reported in Table 8. The expected Mardia's multivariate skewness is 0 and kurtosis is 24 for a multivariate normal distribution of 4 variables[52]. P-value smaller than 0.05 indicated significant skewness or kurtosis. At least one of these tests was significant, thus the underlying joint population was non-normal. Overall, the multivariate skewness and kurtosis suggested a moderate departure from the assumption of a normal distribution in the active disease group when compared with the inactive group (Table 8). A non-constant trend was observed in the group means for both the active and inactive disease group (Table 8) and moderate difference was observed in group covariances. Hence, we used RMQDA assuming Kronecker product covariance. The advantage of imposing the Kronecker product structure on the data is that it reduces the number of parameters to estimate, which results in greater precision of the estimates since n/pq is small for both active (~11) and inactive (~7) disease groups.

Table 8. Descriptive Statistics of IBDQ Domains in active and inactive IBD participants in Manitoba IBD Cohort Study

Measurement Occasion	IBDQ Domains	Active n=133			Inactive n=80		
		Mean (SD)	*Skewness (p-value)*	*Kurtosis (p-value)*	*Mean (SD)*	*Skewness (p-value)*	*Kurtosis (p-value)*
0-month	Emotional health	5.03(1.12)			5.46(1.15)		
	Systemic symptoms	4.78(1.30)	2.35* (<0.001)	24.9* (<0.001)	4.94(1.23)	6.07* (<0.001)	29.6* (<0.001)
	Social function	5.76(1.35)			6.30(0.87)		
	Bowel symptoms	5.03(1.11)			5.69(0.88)		
6-month	Emotional health	5.12(1.03)			5.71(0.89)		
	Systemic symptoms	4.34(1.21)	2.72* (<0.001)	25.45 (0.09)	5.16(1.15)	14.01* (<0.001)	39.74* (<0.001)
	Social function	5.92(1.19)			6.55(0.76)		
	Bowel symptoms	5.09(1.04)			5.88(0.90)		
12-month	Emotional health	5.11(1.11)			5.79(0.91)		
	Systemic symptoms	4.32(1.28)	4.59* (<0.001)	31.93* (<0.001)	5.22(1.14)	16.28* (<0.001)	38.33* (<0.001)
	Social function	5.81(1.30)			6.58(1.00)		
	Bowel symptoms	5.08(1.11)			5.96(0.90)		

Note: *p-value < 0.05, the joint distribution of the variables has significant skewness or kurtosis

In estimating parameters for the proposed robust RMQDA method (MVE), the symmetric trimming parameter was chosen to be 10%, and compared to RMQDA based on MLE. Results of these approaches were reported in Table 9. Overall, we observed a 1% to 3% increase in all robust methods compared to MLE with 10% trimming. As observed from the simulation, these robust procedures may not always be more efficient than RMDA based on MLE for moderate departures from a multivariate normal distribution.

Table 9. Overall Classification Accuracy of Conventional and Robust QDA procedures for IBD data

	un-UNAR	un-UNCS	st-UNAR	st-UNCS
MLE	0.50	0.60	0.64	0.51
MVE (10%)	0.53	0.63	0.65	0.52

Note: un-UNAR = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; un-UNCS = unstructured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; st-UNAR = structured means-Kronecker product of unstructured between-outcomes and within-outcome AR-1 correlation matrices; st-UNCS = structured means-Kronecker product of unstructured between-outcomes and within-outcome CS correlation matrices; MLE=Maximum likelihood estimator; MVE= minimum volume ellipsoid

5. Discussion

This study investigated RMDA procedures based on structured and unstructured means with Kronecker covariances based on maximum trimmed estimators for discriminating between population groups. As expected, the accuracy of the RMDA procedures were highest in multivariate normal distributions but lowest when the data were sampled from a multivariate Cauchy distribution. RMDA procedures based on MTLE were more accurate than the conventional RMDA based on MLEs when data were sampled from multivariate lognormal and Cauchy distributions[24,36]. However, there were negligible differences in the average accuracy rates of the former and latter procedures under multivariate normal and moderately heavy-tailed distributions. Furthermore, our results also showed that the impact of group mean separation i.e., distance and data dimensions on the accuracy of the conventional RMDA procedures could be masked by the departure from the assumption of multivariate normality. In contrast, the impact of both group mean separation and data dimension on the accuracy of the RMDA procedures based on MTLE was not confounded by departure from the assumption of multivariate normality. A common criticism of trimmed estimators is that they are less powerful in small-sampled studies under multivariate normal distributions[53]. However, our simulation study showed negligible differences in the accuracy of RMDA procedures based on MTLE and those based on MLE in small-sampled conditions[36].

Of note is the finding that we observed similar in the accuracy of the investigated RMDA procedures based on parsimonious means/or covariance matrices, regardless of the method of estimation. This can be attributed to the fact that all these procedures were investigated in scenarios where the underlying means and covariance structures were correctly specified. It is most likely that the predictive performances of these procedures might vary especially in multivariate repeated measures data in which group means and covariances are unstructured where the assumption of parsimony (i.e. Kronecker product assumption for group means and/or covariance) are violated. While previous research studies have suggested that RMDA procedures often result in decreased predictive accuracy when the means and covariances are mis-specified[36,54], there is limited investigation of the robustness of the RMDA based on MTLE to model mis-specification. Future research investigations will examine the robustness of RMDA procedures based on MTLE to mis-specification of group means and covariance structure.

This study has some limitations. Our simulation only investigated the classification performance of the investigated models in multivariate normal and multivariate heavy-tailed distributions but not in multivariate skewed distribution. Previous investigations have shown that trimmed estimators are particularly more efficient in data with moderate to significant heavy-tailed distributions[36]. Second, the assumption of complete MRM data in which there is no missing data on all outcomes and at all measurement occasions might not be realistic in multivariate longitudinal data often encountered in applied research. In clinical settings, missing data often

occur in multivariate longitudinal studies because patients miss some of their regular appointments or because some variables may not be measured at particular visits. RMDA based on mixed-effects models have been proposed for incomplete multivariate longitudinal data but the misspecification of the common assumption of the random effects parameter as multivariate normal distribution may seriously affect the accuracy of discriminant analysis classification rules. [25]Pattern mixture and selection models have been proposed to adjust for potential bias in models when it cannot be assumed that the mechanism of missingness is ignorable[55,56,57]. Further research could investigate the development of RMDA procedures based on these models with imputations and further developments in which mixed-effects models can be extended to these robust trimmed methods for classification. In addition, the procedures developed in this study are based on two-group multivariate repeated designs. Nevertheless, our conclusions can be extended and generalized to multi-group designs. In addition, this study rely on the assumption of Kronecker product structure covariance to capture relationship among multivariate repeated measures. Various researches have used Kronecker product covariance structures to address sample size and computational issues in multivariate repeated measures[12,13,33-35]. While Kronecker structures provide a parsimonious model approach to parameter estimation, the accuracy of the resulting RMDA procedures may be reduced when the means and/or covariance structure of the data is mis-specified[58]. It is important that the choice of these RMDA procedures be guided first by a preliminary examination of the appropriate means and/or covariance structure in the multivariate repeated measures data[21,59]. For example, several procedures have been developed for testing hypotheses Kronecker product covariance structures in multivariate repeated measures for such purposes[13,35,60,61].

In summary, this study proposes a new class of RMDA procedures based on MTLE, which overcome the inherently restrictive distributional assumption of multivariate normality when discriminating between populations groups in multivariate repeated measures data characterized by multivariate non-normal distributions. These procedures are useful for developing classification models for both short-term and long-term outcomes in complex data.

References

Fieuws S, Verbeke G, Maes B, Vanrenterghem Y. Predicting renal graft failure using multivariate longitudinal profiles. Biostatistics. 2007;9(3):419-31.

Hughes DM, Komárek A, Czanner G, Garcia-Fiñana M. Dynamic longitudinal discriminant analysis using multiple longitudinal markers of different types. Statistical methods in medical research. 2018;27(7):2060-80.

Inoue LYT, Etzioni R, Morrell C, Müller P. Modeling disease progression with longitudinal markers. Journal of the American Statistical Association. 2008;103(481):259-70.

Li Y, Wang Y, Wu G, Shi F, Zhou L, Lin W, et al. Discriminant analysis of longitudinal cortical thickness changes in alzheimer's disease using dynamic and network features. Neurobiology of aging. 2012;33(2):427. e15-. e30.

Marshall G, Barón AE. Linear discriminant models for unbalanced longitudinal data. Statistics in medicine. 2000;19(15):1969-81.

Morrell CH, Brant LJ, Sheng S, Metter EJ. Screening for prostate cancer using multivariate mixed-effects models. Journal of applied statistics. 2012;39(6):1151-75.

Diggle P. Analysis of longitudinal data: Oxford University Press; 2002.

Verbeke G, Fieuws S, Molenberghs G, Davidian M. The analysis of multivariate longitudinal data: A review. Statistical methods in medical research. 2014;23(1):42-59.

Galecki AT. General class of covariance structures for two or more repeated factors in longitudinal data analysis. Communications in Statistics-Theory and Methods. 1994;23(11):3105- 19.

Lix L, Sajobi T. Discriminant analysis for repeated measures data: A review. Frontiers in Psychology. 2010;1:146.

Roy A, Khattree R. Discrimination and classification with repeated measures data under different covariance structures. Communications in Statistics—Simulation and Computation®. 2005;34(1):167-78.

Roy A, Khattree R. On discrimination and classification with multivariate repeated measures data. Journal of Statistical Planning and Inference. 2005;134(2):462-85.

Roy A, Khattree R. On implementation of a test for kronecker product covariance structure for multivariate repeated measures data. Statistical Methodology. 2005;2(4):297-306.

Tomasko L, Helms RW, Snapinn SM. A discriminant analysis extension to mixed models. Statistics in medicine. 1999;18(10):1249-60.

Albert A. Discriminant analysis based on multivariate response curves: A descriptive approach to dynamic allocation. Statistics in medicine. 1983;2(1):95-106.

Usami S, Murayama K. Time-specific errors in growth curve modeling: Type-1 error inflation and a possible solution with mixed-effects models. Multivariate Behavioral Research. 2018;53(6):876-97.

Wu W, West SG, Taylor AB. Evaluating model fit for growth curve models: Integration of fit indices from sem and mlm frameworks. Psychological methods. 2009;14(3):183.

Komárek A, Hansen BE, Kuiper EM, van Buuren HR, Lesaffre E. Discriminant analysis using a multivariate linear mixed model with a normal mixture in the random effects distribution. Statistics in medicine. 2010;29(30):3267-83.

Marshall G, De la Cruz- Mesía R, Barón AE, Rutledge JH, Zerbe GO. Non-linear random effects model for multivariate responses with missing data. Statistics in medicine. 2006;25(16):2817-30.

Marshall G, De la Cruz- Mesía R, Quintana FA, Barón AE. Discriminant analysis for longitudinal data with multiple continuous responses and possibly missing data. Biometrics. 2009;65(1):69-80.

Roy A. A new classification rule for incomplete doubly multivariate data using mixed effects model with performance comparisons on the imputed data. Statistics in medicine. 2006;25(10):1715-28.

Fieuws S, Verbeke G, Molenberghs G. Random-effects models for multivariate repeated measures. Statistical methods in medical research. 2007;16(5):387-97.

Hughes DM, Komárek A, Czanner G, Garcia-Finana M. Dynamic longitudinal discriminant analysis using multiple longitudinal markers of different types. Statistical methods in medical research. 2018;27(7):2060-80.

Hadi AS, Luceño A. Maximum trimmed likelihood estimators: A unified approach, examples, and algorithms. Computational Statistics & Data Analysis. 1997;25(3):251-72.

Maronna RA. Robust m-estimators of multivariate location and scatter. The Annals of Statistics. 1976;4(1):51-67.

Yan F, Lin X, Huang X. Dynamic prediction of disease progression for leukemia patients by functional principal component analysis of longitudinal expression levels of an oncogene. The annals of applied statistics. 2017;11(3):1649-70.

Rousseuw PJ, Leroy AM. Robust regression and outlier detection. Wiley, New York; 1987.

Cheng T-C, Biswas A. Maximum trimmed likelihood estimator for multivariate mixed continuous and categorical data. Computational Statistics & Data Analysis. 2008;52(4):2042-65.

Barön AE. Misclassification among methods used for multiple group discrimination- the effects of distributional properties. Statistics in medicine. 1991;10(5):757-66.

He X, Fung WK. High breakdown estimation for multiple populations with applications to discriminant analysis. Journal of Multivariate Analysis. 2000;72(2):151-62.

Williams BK, Titus K. Assessment of sampling stability in ecological applications of discriminant analysis. Ecology. 1988;69(4):1275-85.

Fitzmaurice G, Laird N, Ware J. Modelling the mean: Parametric curves. Applied Longitudinal Analysis Hoboken, New Jersey, USA: John Wiley & Sons Inc. 2004:141-7.

Naik DN, Rao SS. Analysis of multivariate repeated measures data with a kronecker product structured covariance matrix. Journal of applied statistics. 2001;28(1):91-105.

Krzyśko M, Skorzybut M. Discriminant analysis of multivariate repeated measures data with a kronecker product structured covariance matrices. Statistical Papers. 2009;50(4):817-35.

Lu N, Zimmerman DL. The likelihood ratio test for a separable covariance matrix. Statistics & probability letters. 2005;73(4):449-57.

Sajobi TT, Lix LM, Dansu BM, Laverty W, Li L. Robust descriptive discriminant analysis for repeated measures data. Computational Statistics & Data Analysis. 2012;56(9):2782-94.

Thomas DR, Zumbo BD. Using a measure of variable importance to investigate the standardization of discriminant coefficients. Journal of Educational and Behavioral Statistics. 1996;21(2):110-30.

Ripley B. Ebooks corporation. Stochastic simulation. Wiley Online Library; 1987.

Aitchison J. The statistical analysis of compositional analysis. Chapman & Hall, London; 1986.

Azzalini A. The skew-normal and related families: Cambridge University Press; 2013.

Stigler SM. Do robust estimators work with real data? The Annals of Statistics. 1977:1055-98.

Wilcox RR. Fundamentals of modern statistical methods: Substantially improving power and accuracy: Springer; 2010.

Ramsey PH, Ramsey PP. Optimal trimming and outlier elimination. Journal of Modern Applied Statistical Methods. 2007;6(2):2.

Rousseeuw PJ, Driessen KV. A fast algorithm for the minimum covariance determinant estimator. Technometrics. 1999;41(3):212-23.

Maronna RA, Martin RD, Yohai VJ. Robust statistics: Theory and methods (with r): John Wiley & Sons; 2019.

Todorov V, Filzmoser P. An object-oriented framework for robust multivariate analysis. 2009.

Rousseeuw PJ, Leroy AM. Robust regression and outlier detection: John wiley & sons; 2005.

Liu Y, Zumbo BD. The impact of outliers on cronbach's coefficient alpha estimate of reliability: Visual analogue scales. Educational and psychological measurement. 2007;67(4):620-34.

Liu Y, Zumbo BD. Impact of outliers arising from unintended and unknowingly included subpopulations on the decisions about the number of factors in exploratory factor analysis. Educational and psychological measurement. 2012;72(3):388-414.

Cohen J. Statistical power analysis for the behavioral sciences, 2nd edn. Á/l. Erbaum Press, Hillsdale, NJ, USA; 1988.

Bernstein CN, Rawsthorne P, Cheang M, Blanchard JF. A population-based case control study of potential risk factors for ibd. American Journal of Gastroenterology. 2006;101(5):993-1002.

Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika. 1970;57(3):519-30.

Srivastava DK, Mudholkar GS. Trimmed t2: A robust analog of hotelling's t2. Journal of Statistical Planning and Inference. 2001;97(2):343-58.

Sajobi TT, Lix LM, Li L, Laverty W. Discriminant analysis for repeated measures data: Effects of mean and covariance misspecification on bias and error in discriminant function coefficients. Journal of Modern Applied Statistical Methods. 2011;10(2):15.

Hedeker D, Gibbons RD. Application of random-effects pattern-mixture models for missing data in longitudinal studies. Psychological methods. 1997;2(1):64.

Little RJ. Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association. 1993;88(421):125-34.

Little RJ. Modeling the drop-out mechanism in repeated-measures studies. Journal of the American Statistical Association. 1995;90(431):1112-21.

Zhou J, Qu A. Informative estimation and selection of correlation structure for longitudinal data. Journal of the American Statistical Association. 2012;107(498):701-10.

Roy A, editor A note on testing of kronecker product covariance structures for doubly multivariate data. Proceedings of the American Statistical Association, statistical computing section; 2007.

Srivastava MS, von Rosen T, Von Rosen D. Models with a kronecker product covariance structure: Estimation and testing. Mathematical Methods of Statistics. 2008;17(4):357-70.

Filipiak K, Klein D, Roy A. A comparison of likelihood ratio tests and rao's score test for three separable covariance matrix structures. Biometrical Journal. 2017;59(1):192-215.