A new diagnostic accuracy measure and cut-point selection criterion

Tuochuan Dong,1 Kristopher Attwood,2 Alan Hutson,1 Song Liu2 and Lili Tian1

Abstract
Most diagnostic accuracy measures and criteria for selecting optimal cut-points are only applicable to diseases with binary or three stages. Currently, there exist two diagnostic measures for diseases with general $k$ stages: the hypervolume under the manifold and the generalized Youden index. While hypervolume under the manifold cannot be used for cut-points selection, generalized Youden index is only defined upon correct classification rates. This paper proposes a new measure named maximum absolute determinant for diseases with $k$ stages ($k \geq 2$). This comprehensive new measure utilizes all the available classification information and serves as a cut-points selection criterion as well. Both the geometric and probabilistic interpretations for the new measure are examined. Power and simulation studies are carried out to investigate its performance as a measure of diagnostic accuracy as well as cut-points selection criterion. A real data set from Alzheimer’s Disease Neuroimaging Initiative is analyzed using the proposed maximum absolute determinant.

Keywords
Maximum absolute determinant, optimal cut-points, volume for the parallelotope, Alzheimer’s Disease

1 Introduction
In diagnostic studies, the most straightforward classification rule is binary, i.e. non-diseased or diseased. The probabilities that a diagnostic biomarker correctly classifies a non-diseased subject or a diseased subject are defined as specificity or sensitivity, respectively. Plotting sensitivity against 1-specificity for all the possible cut-points or decision thresholds from a continuous biomarker leads to what is termed a receiver operating characteristic (ROC) curve. There exist many detailed reviews pertaining to inference in the framework of ROC curve.1–4 For diseases with binary classification, the most popular diagnostic measure of accuracy is the area under the ROC curve (AUC), which has been extensively studied.5–13 Another popular diagnostic measure, the Youden index,14 is defined as

1Department of Biostatistics, University at Buffalo, Buffalo, NY, USA
2Department of Biostatistics and Bioinformatics, Roswell Park Cancer Institute, Buffalo, NY, USA

Corresponding author:
Lili Tian, Department of Biostatistics, 717 Kimball Tower, 3435 Main St. Bldg. 26 Buffalo, NY 14214, USA.
Email: ltian@buffalo.edu
the maximum of specificity + sensitivity – 1. Inference procedures regarding Youden index have also been well studied.\textsuperscript{15–18}

In practice, many disease processes can be naturally classified into more than two stages. For example, in Alzheimer’s disease there exists a transition stage called mild cognitive impairment (MCI), which is the cognitive change between normal aging (non-diseased) and complete memory loss (fully diseased).\textsuperscript{19,20} Therefore, we often have a three-stage classification in Alzheimer’s disease: non-diseased, early diseased and fully diseased. Consequently, the ROC curve for binary classification is extended to ROC surface for three-stage classification and the AUC to the volume under the ROC surface (VUS), which is the most widely used quantitative summary index for diseases with three stages.\textsuperscript{21} There are several recent papers in the statistical literature focused on VUS research.\textsuperscript{22–28} Furthermore, as an extension of the Youden index to diseases with three stages, Nakas et al.\textsuperscript{29} introduced the generalized Youden index (GYI), defined as the maximum of the sum of the correct classification rates (CCRs) for three groups minus 1. There are several recent papers in the statistical literature focused on GYI research.\textsuperscript{30,31} For general cases of multiple-class diagnoses, the concept of hypervolume under the manifold (HUM) was first proposed as an extension of AUC by Scurfield\textsuperscript{32,33} in the field of mathematical psychology. Nakas and Yiannoutsos\textsuperscript{26} first proposed the non-parametric estimator for HUM, and Li and Fine\textsuperscript{34} proposed the corresponding inference procedures and methods for estimating classification probabilities and calculating HUM. Furthermore, the concept of GYI has also been extended to the diseases with general k stages.\textsuperscript{35} A brief review of HUM and GYI is given in Section 2.

The critical step for making diagnoses based on a specific biomarker is to select the optimal cut-points. For diseases with a binary classification, Fluss et al.\textsuperscript{16} discussed the performance of Youden index in terms of selecting optimal cut-point under different assumptions. Perkins and Schisterman\textsuperscript{36} compared two common approaches for optimal cut-point selection: the point on the ROC curve closest to (0,1) and the Youden index. More recently, Liu\textsuperscript{37} developed a method for selecting the optimal cut-point by maximizing the product of specificity and sensitivity. For diseases with three stages, He and Frey\textsuperscript{38} discussed the “optimal” decision making criteria using likelihoods; Nakas et al.\textsuperscript{29} proposed selecting the optimal cut-points using GYI; and Attwood et al.\textsuperscript{39} introduced two new criteria as the closest to perfection/minimum distance (MD) method, and the max volume (MV) method which are the extensions of the corresponding criteria for diseases with binary status. For diseases with general k stages, GYI\textsuperscript{35} can be used as a criterion for the optimal cut-points selection.

In this paper, we propose a new measure of diagnostic accuracy for diseases with \( \geq 2 \) classes. This new measure directly incorporates all possible true and false classification rates. It has an appealing geometric interpretation, which justifies its use as a measure for classification accuracy. Furthermore, the proposed measure can also be used for cut-point selection. The rest of paper is organized as follows: in Section 2, preliminary background information is introduced and some existing diagnostic accuracy measures are reviewed. In Section 3, a new measure, namely the maximum absolute determinant (MADET), is proposed and specific details are discussed. In Section 4, a power study is carried out to assess the performance of the proposed measure in comparison with HUM and GYI in testing the equality of two biomarkers. In Section 5, issues for optimal cut-points selection are discussed and simulation results comparing criteria MADET with other existing criteria are presented. In Section 6, several important biomarkers from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) study are analyzed. In Section 7, summary and discussion are given.
2 Preliminaries

In this section, we present the definitions, notation and a review of existing measures of diagnostic accuracy for diseases with general \( k \) stages.

2.1 Notations

For a disease that can be classified into \( k \) different classes, let \( S \) denote the true disease status for a subject and let \( T \) denote the diagnostic result based on a continuous biomarker. A subject with true disease status \( i \) (\( i = 1, 2, \ldots, k \)) can correctly or incorrectly be classified as having disease state \( j \) (\( j = 1, 2, \ldots, k \)). Define the conditional probability of a classification given the true disease status as

\[
P_{ij} = P(T = j | S = i), \quad \text{where} \quad i = 1, \ldots, k, j = 1, \ldots, k
\]

Hence, \( P_{ij} \) denotes the probability that a subject from class \( i \) is correctly diagnosed and \( P_{ij} (i \neq j) \) is the probability that a subject from class \( i \) is incorrectly diagnosed as class \( j \). For a subject in class \( i \), let \( \mathbf{P}_i \) denote the vector consisting of \( k \) classification rates, i.e.

\[
\mathbf{P}_i = \begin{pmatrix}
P_{i1} \\
\vdots \\
P_{ik}
\end{pmatrix} \quad \text{where} \quad i = 1, \ldots, k
\]

By its definition, it follows

\[
\begin{pmatrix}
\mathbf{1}^T \\
\mathbf{P}_i
\end{pmatrix} = \sum_{j=1}^{k} P_{ij} = 1, \quad \text{for} \quad i = 1, \ldots, k
\]

Hence, at a given \( i \), out of \( k \) components \( P_{i1}, P_{i2}, \ldots, P_{ik} \), only \( k - 1 \) of them are independent.

Now consider \( k \) independent subjects whose true disease statuses are from each of the \( k \) classes, respectively. Then we have vectors \( \mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k \) to fully describe the performance of the biomarkers being considered, where \( \mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k \) are mutually independent. Next define a \( k \times k \) square probability matrix (SPM) as

\[
\mathbf{P} = \begin{pmatrix}
\mathbf{P}_1^T \\
\vdots \\
\mathbf{P}_k^T
\end{pmatrix} = (P_{ij})
\]

The definition of SPM captures both the correct and false classification rates for all \( k \) classes. Overall, there are \( k \) CCRs and \( k(k - 1) \) false classification rates.

The class probabilities \( P_{ij} \)'s for \( i, j = 1, \ldots, k \) need to be assessed by some decision rules. To make a diagnosis for a disease with \( k \) stages based on a continuous biomarker \( X \), cut-points \( c_1, \ldots, c_{k-1} \) are needed. Here \( c_1 < \cdots < c_{k-1} \) is assumed. If \( c_{j-1} < X \leq c_j \), for \( j = 1, \ldots, k \), this subject is classified into stage \( j \). Let \( X_i \) denote the measurement for a subject from the \( i \)th class. Hence, the corresponding conditional probability \( P_{ij} \) can be written as

\[
P_{ij} = P(c_{j-1} < X_i \leq c_j | S = i), \quad \text{where} \quad i = 1, \ldots, k, j = 1, \ldots, k
\]
2.2 Existing diagnostic measures for diseases with \( k \) stages

In this subsection, we review two measures of diagnostic accuracy proposed for diseases with general \( k \) stages, i.e. the HUM\(^{26,34}\) and the GYI.\(^{35}\)

2.2.1 HUM

Let \( t_i \) denote the CCR for the \( i \)th disease class, i.e. \( P_{i,i} \) as defined in equation (1). By plotting the \( t_i \)'s (\( i = 1, \ldots, k \)) in the \( k \)-dimensional space, a manifold is obtained. The corresponding HUM is utilized as a measure of diagnostic accuracy for diseases with general \( k \) stages. A rigorous mathematical definition of HUM is given by Li and Fine\(^{34}\) as

\[
HUM = \int_0^1 \int_0^{f_1(t_1)} \cdots \int_0^{f_{k-2}(t_{k-1}, \ldots, t_{k-2})} f_{k-1}(t_1, \ldots, t_{k-1}) dt_{k-1} \cdots dt_2 dt_1
\]

where \( f_i \)'s are the recursive equations defined as \( t_i = f_{i-1}(t_1, \ldots, t_i-1), i = 2, \ldots, k \). The \( t_i \)'s can be estimated via multinomial logistic regression or by using cut-points. For details, see the paper by Li and Fine.\(^{34}\) Both the AUC and VUS are special cases of HUM for diseases with binary and three stages, respectively.

2.2.2 GYI

Nakas et al.\(^{35}\) generalized the Youden Index to the general \( k \) stages classification problem. The \( k \) stages GYI is defined as

\[
GYI(c_1, c_2, \ldots, c_{k-1}) = \max_{c_1 < \cdots < c_{k-1}} [P_{1,1} + \cdots + P_{k,k} - 1]
\]

Under the condition \( c_1 < \cdots < c_{k-1} \), it was shown that GYI is the sum of the Youden indexes for the adjacent classes. As defined, GYI is not only a direct measure of overall CCR, but also a tool for optimal cut-points selection.

2.2.3 Overview of HUM and GYI and justification for the new measure

Both HUM and GYI have certain advantages and disadvantages. HUM is a global measure in nature, i.e. it gauges the diagnostic accuracy of a biomarker across all the possible cut-points. Consequently, HUM cannot be used to determine the optimal cut-points. Hence, Li and Fine\(^{34}\) proposed to use multinomial logistic regression for estimating class probabilities. On the other hand, GYI can be used to determine the optimal cut-points, which yield the maximum overall correct classification. However, GYI considers only the \( k \) CCRs and ignores the \( k(k-1) \) false classification rates, which are also the critical and integral components for describing diagnostic accuracy. As the number of classes \( k \) increases, less information is utilized by the GYI method. Hence, GYI becomes less optimal in this setting.

Therefore, there is a need for a comprehensive diagnostic measure, which is able to take into account all the information related to a classification rule, i.e. the information contained in the SPM \( \mathcal{P} \), and which can also serve as a tool for optimal cut-point selection. The goal of this paper is to fill this gap.

3 New diagnostic accuracy measure

In this section, we propose a new measure of diagnostic accuracy, which directly utilizes all the classification rates contained in the SPM \( \mathcal{P} \) in equation (2).
We term this new measure as the maximum (supremum) absolute determinant (MADET) of the SPM. It is defined as

$$\text{MADET} = \sup \ |\det(P)|$$  \hspace{1cm} (8)

If the class probabilities are estimated using cut-points as discussed in Section 2.1, MADET can be expressed as

$$\text{MADET} = \max_{c_1, \ldots, c_{k-1}} |\det(P)(c_1, \ldots, c_{k-1})|$$  \hspace{1cm} (9)

where $c_1, \ldots, c_{k-1}$ need to be determined by some criteria.

### 3.1 Geometric interpretation of MADET

In the following section, we investigate the geometric interpretation of MADET and subsequently justify its use as a diagnostic accuracy measure. For this purpose, several definitions from advanced geometry need to be restated.

**Definition of parallelotope:** Given a set of independent vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ of dimension $k$ in the Euclidean space, a parallelotope is the set of vectors

$$\mathcal{V}(\tilde{a}_1, \ldots, \tilde{a}_k) = \left\{ \sum_i x_i \tilde{a}_i : 0 \leq x_i \leq 1 \right\}$$

for $1 \leq i \leq k$.

Given an orthonormal basis $\tilde{e}_1, \ldots, \tilde{e}_k$, the vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ can be expressed as

$$(\tilde{a}_1, \ldots, \tilde{a}_k)^T = A(\tilde{e}_1, \ldots, \tilde{e}_k)^T$$

where $A$ is a $k \times k$ matrix. The base of the $k$-dimensional parallelotope on vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ is defined as the (k-1)-dimensional parallelotope $\mathcal{V}(\tilde{a}_1, \ldots, \tilde{a}_{k-1})$ and its height is defined as the orthogonal component of $\tilde{a}_k$ with respect to the sub-space spanned by $\tilde{a}_1, \ldots, \tilde{a}_{k-1}$, denoted as $\text{ort}(\tilde{a}_1, \ldots, \tilde{a}_{k-1})\tilde{a}_k$. The volume of a $k$-dimensional parallelotope ($k > 1$) is the product of the volume of its base and its height. The volume of a one-dimensional parallelotope $\mathcal{V}(\tilde{a}_1)$ is the length of the vector $\tilde{a}_1$. The volume of parallelotope on vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ can be easily obtained as stated in the following theorem:

**Theorem 1:** Assume that vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ are expressed via an orthonormal basis $\tilde{e}_1, \ldots, \tilde{e}_k$ with matrix $A$. The volume of the $k$-dimensional parallelotope $\mathcal{V}(\tilde{a}_1, \ldots, \tilde{a}_k)$ is equal to the determinant of matrix $A$, i.e.

$$\text{vol}(\mathcal{V}(\tilde{a}_1, \ldots, \tilde{a}_k)) = |\det(A)|$$  \hspace{1cm} (10)

The detailed proof can be found in many text books in advanced algebra, e.g., Vinberg.\cite{40}

The volume of a $k$-dimensional parallelotope ($k > 1$), $\mathcal{V}(\tilde{a}_1, \ldots, \tilde{a}_k)$, is a measure of how divergent the vectors $\tilde{a}_1, \ldots, \tilde{a}_k$ are in the $k$-dimensional space. When $k=2$, the absolute value of a $2 \times 2$
matrix determinant is the area of a corresponding parallelogram with the two row vectors as sides. When $k = 3$, the absolute value of a $3 \times 3$ matrix determinant is the volume of a corresponding parallelepiped with the three row vectors as sides. Generally, the determinant of a matrix is an intuitive way to picture the volume of $k$-dimensional parallelepiped spanned by its row vectors.

Consider the specific settings in medical diagnosis for diseases with $k$ stages. For the $i$th ($i = 1, \ldots, k$) disease class, the classification rate vector $\mathbf{P}_i$ (as defined in equation (2)) corresponds to $\mathbf{a}_i$ and the SPM $\mathbf{P}$ in equation (4) corresponds to matrix $A$. According to the theorem stated above, MADET as defined in equation (9), is equal to the largest volume of the $k$-dimensional parallelepiped spanned by $\mathbf{P}_i$’$s$ ($i = 1, \ldots, k$). This directly measures how far apart the $\mathbf{P}_i$’$s$ are in the $k$-dimensional space. The further apart the $\mathbf{P}_i$’$s$, the larger the difference between the diagnostic results for the $k$ disease classes and hence the better the diagnostic accuracy of the test. Therefore, a larger MADET indicates a better diagnostic test. When MADET is equal to 1, the classification rate vectors $\mathbf{P}_i$’$s$ ($i = 1, \ldots, k$) coincide with the corresponding orthonormal bases $\mathbf{e}_i$’$s$ ($i = 1, \ldots, k$) and $A$ becomes a $k \times k$ identity matrix. For such cases, $P_{i,i} = 1$ and $P_{i,j} = 0$ for $i \neq j$ indicating a perfect biomarker which can correctly classify subjects in each class. On the other hand, MADET is 0 when the at least one vector $\mathbf{P}_i$’$s$ ($i = 1, \ldots, k$) lies in the hyperplane spanned by the others, thus the set of vectors $\mathbf{P}_i$’$s$ ($i = 1, \ldots, k$) are dependent. For example, for $k = 3$, MADET will be 0 when 1) all three $\mathbf{P}_i$’$s$ ($i = 1, \ldots, 3$) coincide; 2) any two of the $\mathbf{P}_i$’$s$ ($i = 1, \ldots, 3$) coincide; 3) all three $\mathbf{P}_i$’$s$ ($i = 1, \ldots, 3$) fall on the same two-dimensional plane. In other words, from a geometrical perspective, MADET is equal to 0 when the $\mathbf{P}_i$’$s$ ($i = 1, \ldots, 3$) do not form a parallelepiped together. Hence, as the number of stages $k$ is fixed, MADET equal to 0 means the biomarker being considered has no diagnostic ability to distinguish among the $k$ stages.

Note that equation (10) shows the defined volume $\text{vol} \mathbf{V}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$, i.e. $\det |A|$ can be negative. The sign of the value of the determinant provides some information on the orientation of the parallelepiped. The sign of the $\det(\mathbf{P})$ becomes opposite when two columns of $\mathbf{P}$ are switched; i.e. when the order of disease classes are switched. Since this paper focuses on diseases with ordinal stages, we use the absolute value of the determinants in the definition of MADET, i.e. $|\det(\mathbf{P})|$, in equation (9).

From the definition of MADET, the geometric interpretation to justify its use as a diagnostic accuracy is very straightforward. The probabilistic interpretation of MADET for diseases with binary and three stages classification are also obtainable. However, for $k > 3$, the determinant of the matrix $\mathbf{P}$ becomes very complicated, and it becomes a hard task to elucidate the probabilistic interpretation of MADET for diseases with general $k$ stages. In the following, we discuss detailed probabilistic and geometric interpretation for binary and three stages classification.

### 3.2 Special case 1: binary classification

Consider diseases with binary status (1 for non-diseased, and 2 for diseased). The classification rate vector $\mathbf{P}_i$’$s$ defined in equation (2) equals to $(\text{spe}, 1 - \text{spe})^T$ for $i = 1$ and $(1 - \text{sen}, \text{sen})^T$ for $i = 2$ (where $\text{spe}$ stands for the specificity and $\text{sen}$ for the sensitivity). Therefore, the $2 \times 2$ SPM is

$$
\mathbf{P} = \begin{pmatrix}
\text{spe} & 1 - \text{spe} \\
1 - \text{sen} & \text{sen}
\end{pmatrix}
$$

and the newly proposed diagnostic measure maximum absolute determinant (MADET) is

$$
\text{MADET} = \max_{\mathbf{e}_i} (\text{spec} + \text{sen} - 1)
$$
It is obvious that MADET is the same as the Youden index.

As stated in the Introduction, the Youden index is a diagnostic measure as well as a criterion for optimal cut-point selection. The fact that Youden index is the same as MADET offers a new geometric interpretation for this well-known index, as shown in Figure 1. The classification rate vector for a randomly chosen non-diseased subject, \( \mathbf{P}_1 = (P_{1,1}, P_{1,2}) \), is represented by the vector \( \overrightarrow{OA} \), and the one for a diseased subject, \( \mathbf{P}_2 = (P_{2,1}, P_{2,2}) \), is denoted by the vector \( \overrightarrow{OB} \). Due to the fact \( P_{i,1} + P_{i,2} = 1 \), for \( i = 1, 2 \) as stated in equation (3), the ends of two vectors \( \overrightarrow{OA} \) and \( \overrightarrow{OB} \) always fall on the line \( P(T = 1|S) + P(T = 2|S) = 1 \). As stated in Section 3, MADET, i.e. the Youden index for diseases with binary classification, is equal to the area of the corresponding parallelogram with the two row vectors \( \overrightarrow{OA} \) and \( \overrightarrow{OB} \) as sides, indicated by the yellow shaded area: \( \triangle OABC \). Meanwhile, one half of the MADET, is the area of the triangular \( \triangle OAB \). Since \( OD \), the height of the triangular, is fixed as \( \sqrt{2} \), MADET simply measures the length of the line segment \( AB \), i.e. how far apart the two vectors \( \overrightarrow{OA} \) and \( \overrightarrow{OB} \) fall. Hence, for \( k = 2 \), MADET or Youden index serves as a direct measure of diagnostic accuracy; i.e. a measure of the difference between the classification rate vectors \( \mathbf{P}_1 \) for a random non-diseased subject and \( \mathbf{P}_2 \) for a random diseased subject. If \( \overrightarrow{OA} \) rests at \( (1, 0) \) and \( \overrightarrow{OB} \) at \( (0, 1) \), the corresponding MADET reaches the maximum 1. This means the biomarker has the perfect sensitivity and specificity. On the other side, if the diagnostic biomarker has no power for diagnosis, the classification rate vectors for a non-diseased subject and a diseased subject will completely overlap, i.e. \( \overrightarrow{OA} \) and \( \overrightarrow{OB} \) overlap, yielding MADET as 0.

**Figure 1.** Illustration of MADET for the disease with binary classification. The red line denotes \( P(T = 1|S) + P(T = 2|S) = 1 \), and the area of the parallelogram \( \triangle OABC \) represents MADET.
3.3 Special case 2: Diseases with three stages

Consider diseases with three ordinal stages. Let $i = 1, 2, 3$ denote the non-diseased stage, the early diseased stage and the fully diseased stage respectively. The corresponding SPM is

$$
P = \begin{pmatrix}
P_{1,1} & P_{1,2} & P_{1,3} \\
P_{2,1} & P_{2,2} & P_{2,3} \\
P_{3,1} & P_{3,2} & P_{3,3}
\end{pmatrix}
$$

where $P_{i,j}$ stands for the probability of a subject from $i$th class being classified as $j$th class. We know that

$$
P_{i,1} + P_{i,2} + P_{i,3} = 1, \quad \text{for } i = 1, 2, 3
$$

(11)

Simple manipulations lead to

$$
\text{MADET} = \max_{c_1, c_2} \det \left( \begin{array}{ccc}
P_{1,1} & P_{1,2} & P_{1,3} \\
P_{2,1} & P_{2,2} & P_{2,3} \\
P_{3,1} & P_{3,2} & P_{3,3}
\end{array} \right)

= \max_{c_1, c_2} |P_{1,1}P_{2,2}P_{3,3} + P_{1,2}P_{2,3}P_{3,1} + P_{1,3}P_{3,2}P_{2,1} \\
- P_{1,1}P_{2,3}P_{3,2} - P_{1,2}P_{2,1}P_{3,3} - P_{1,3}P_{3,1}P_{2,2}|

= \max_{c_1, c_2} |1 - \left[ P_{1,1}(1 - P_{2,2}) + P_{2,2}(1 - P_{3,3}) + P_{3,3}(1 - P_{1,1}) \right] \\
- (P_{1,2}P_{2,1} + P_{1,3}P_{3,1} + P_{2,3}P_{3,2})|.
$$

(12)

Unlike the binary case, MADET for diseases with three stages is quite different from GYI, defined as $\max_{c_1, c_2} (P_{1,1} + P_{2,2} + P_{3,3} - 1)$. Instead, MADET is a rather complicated function of the correct and false classification rates, due to the fact that it simultaneously takes all the correct and false classification rates into consideration.

From equation (12), MADET has the following probabilistic interpretation. Consider two subjects 1 and 2 from two different disease groups $(i, j; i, j \in \{1, 2, 3\}, i \neq j)$ respectively, let $S_l$ denote the true disease status for subject $l$, and $T_l$ denote the corresponding diagnostic result, for $l = 1, 2$. MADET can be further written as

$$
\text{MADET} = \max_{c_1, c_2} |1 - P(T_1 = i, T_2 \neq j | S_1 = i, S_2 = j; i, j \in \{1, 2, 3\}, i \neq j) \\
- P(T_1 \neq i, T_2 \neq j | S_1 = i, S_2 = j; T_1, T_2 \in \{i, j\}; i, j \in \{1, 2, 3\}, i \neq j)|
$$

where $P(T_1 = i, T_2 \neq j | S_1 = i, S_2 = j; i, j \in \{1, 2, 3\}, i \neq j)$ stands for the probability that the first subject is correctly classified but the second is not, and $P(T_1 \neq i, T_2 \neq j | S_1 = i, S_2 = j; T_1, T_2 \in \{i, j\}; i, j \in \{1, 2, 3\}, i \neq j)$ stands for the conditional probability that two subjects are classified into the wrong classes given the fact that the two disease classes for the two subjects are known, while which subject belongs to which class is unknown. In other words, MADET measures the total probability 1 with the penalization of the probability that the diagnostic biomarker correctly diagnoses the first subject but fails for the second, and the conditional probability that the diagnostic biomarker falsely classifies both subjects given the two groups $(i, j)$. 
The geometric interpretation of MADET for the disease with three stages is as follows. In Figure 2, the classification rate vectors for three disease classes, i.e. $P_1$, $P_2$ and $P_3$ for randomly chosen non-diseased, early stage and diseased subjects are represented by vectors $\vec{OA}$, $\vec{OB}$ and $\vec{OC}$ respectively. Based on equation (11), the ends of $\vec{OA}$, $\vec{OB}$ and $\vec{OC}$ fall on the yellow plane defined as $P(T = 1|S) + P(T = 2|S) + P(T = 3|S) = 1$. According to the aforementioned theorem in Section 3, MADET is equal to the volume of the parallelepiped spanned by $\vec{OA}$, $\vec{OB}$ and $\vec{OC}$, and one sixth of the MADET denotes the volume of the tetrahedron $OABC$, i.e. the red shaded polyhedron. Similar to the two-dimensional case, the height of the tetrahedron, from point $O$ to the plane $P(T = 1|S) + P(T = 2|S) + P(T = 3|S) = 1$, is fixed as $\frac{\sqrt{3}}{6}$. Therefore MADET measures the area of the triangular $\triangle ABC$ on the yellow plane $P(T = 1|S) + P(T = 2|S) + P(T = 3|S) = 1$. The larger the MADET, the more diverse $\vec{P_1}$, $\vec{P_2}$ and $\vec{P_3}$ are. Hence, larger value of MADET indicates larger difference among the classification rates vectors $\vec{P_1}$, $\vec{P_2}$ and $\vec{P_3}$, and therefore better diagnostic ability of the biomarker. For a perfect diagnostic biomarker, $\vec{P_1}$, $\vec{P_2}$ and $\vec{P_3}$ will be $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, and the MADET will reach the maximum value 1. For a diagnostic biomarker with no diagnostic ability for distinguishing among the three disease classes, the three vectors $\vec{OA}$, $\vec{OB}$ and $\vec{OC}$ fall on the same plane, including scenarios such as all three vectors or any two of them overlap, yielding MADET as 0.
In order to assess the performance of the newly proposed diagnostic measure MADET, a power study was conducted comparing our new methodology with two existing measures: GYI and VUS, in testing for differences between biomarker levels across disease stages. We considered diseases with three stages; i.e. $k = 3$ under a variety of underlying distributions. The settings are presented in the Table 1. The true values for each measure under $H_0$ and $H_a$ are shown in Table 2. Under normal, Gamma and exponential respectively, the first two scenarios have null hypothesis for which the three distributions are identical. Such scenarios are for testing the hypothesis if the biomarkers have any diagnostic accuracy at all. The last 3 scenarios are for testing if the biomarkers have some diagnostic ability. For scenario 3, $H_0$ and $H_a$ are the same. Hence, scenario 3 is used for some level of checking type-I error control.

For each scenario, the sample sizes for three classes $(n_1, n_2, n_3)$ are assumed as equal and set as 10, 20, 40. Random samples are simulated under $H_0$ for 2000 runs and 95% percentiles for the values of the statistics (VUS in equation (6), GYI in equation (7) and MADET in equation (8)) are estimated from the smoothed kernel distribution functions. Under $H_a$, $H_0$ is rejected if the estimated statistics are greater than their corresponding 95% percentiles from $H_0$. The power for each measure is calculated as the proportion that $H_0$ is rejected out of the 1000 replications. The results are presented in Table 3.

As Table 3 shows, for scenario 3, all the three measures generally maintain the nominal level. For all other scenarios, as the sample sizes go up from 10 to 40, the powers increase. There is no clear winner in term of power. For the scenarios 1 and 2, the tests based on MADET have lower or comparable power than those based on GYI or VUS. Explanations can also be obtained based on Table 2. From null hypothesis consisting of three identical distributions to alternative hypothesis in the neighborhood, MADET increase more slowly than GYI or VUS, yielding a lower power for

---

**Table 1.** Distributional scenarios for the power study.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Distributions under $H_0$</th>
<th>Distributions under $H_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-diseased</td>
<td>Early diseased</td>
</tr>
<tr>
<td>Normal 1</td>
<td>N(0, 1)</td>
<td>N(0, 1)</td>
</tr>
<tr>
<td>Normal 2</td>
<td>N(0, 1)</td>
<td>N(0, 1)</td>
</tr>
<tr>
<td>Normal 3</td>
<td>N(0, 1)</td>
<td>N(5.3, 1)</td>
</tr>
<tr>
<td>Normal 4</td>
<td>N(0, 1)</td>
<td>N(5.3, 1)</td>
</tr>
<tr>
<td>Gamma 1</td>
<td>G(2, 6)</td>
<td>G(2, 6)</td>
</tr>
<tr>
<td>Gamma 2</td>
<td>G(2, 6)</td>
<td>G(2, 6)</td>
</tr>
<tr>
<td>Gamma 3</td>
<td>G(2, 1)</td>
<td>G(3.2, 4)</td>
</tr>
<tr>
<td>Gamma 4</td>
<td>G(2, 1)</td>
<td>G(3.2, 4)</td>
</tr>
<tr>
<td>Exponential 1</td>
<td>E(1)</td>
<td>E(1)</td>
</tr>
<tr>
<td>Exponential 2</td>
<td>E(1)</td>
<td>E(1)</td>
</tr>
<tr>
<td>Exponential 3</td>
<td>E(1)</td>
<td>E(0.96)</td>
</tr>
<tr>
<td>Exponential 4</td>
<td>E(1)</td>
<td>E(0.96)</td>
</tr>
</tbody>
</table>
For scenarios 4 and 5, the tests based on MADET have higher power than those based on GYI or VUS. For example, for the scenario “Gamma 4”, when the sample size is 40, the test based on MADET achieves a power of 0.992, while the one based on GYI is merely at 0.05 nominal level and the one for VUS is only 0.13.

Again, some explanations can be obtained using the results presented in Table 2. Under the scenario

| Table 2. MADET, GYI and VUS under $H_0$ and $H_a$ in Table 1. |
|----------------|----------------|----------------|----------------|----------------|----------------|
|                | Normal 1       | Normal 2       | Normal 3       | Normal 4       | Normal 5       |
| Hypothesis     | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| $H_0$          | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
| $H_a$          | 0.0119         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| $H_0$          | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
| $H_a$          | 0.0119         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| $H_0$          | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
| $H_a$          | 0.0119         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| $H_0$          | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
| $H_a$          | 0.0119         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| $H_0$          | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |
| $H_a$          | 0.0119         | 0.0000         | 0.0000         | 0.0000         | 0.0000         | 0.0000         |

| Table 3. Simulated power for MADET, GYI and VUS. |
|----------------|----------------|----------------|----------------|----------------|----------------|
|                | Normal 1       | Normal 2       | Normal 3       | Normal 4       | Normal 5       |
| Sample size    | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| 10             | 0.245          | 0.507          | 0.601          | 0.282          | 0.102          | 0.043          |
| 20             | 0.448          | 0.793          | 0.864          | 0.340          | 0.067          | 0.059          |
| 40             | 0.701          | 0.972          | 0.993          | 0.725          | 0.054          | 0.052          |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| 10             | 0.0113         | 0.0181         | 0.5534         | 0.0480         | 0.5773         | 0.4259         |
| 20             | 0.1134         | 0.8181         | 0.6534         | 0.0480         | 0.5773         | 0.4259         |
| 40             | 0.2130         | 0.9720         | 0.9930         | 0.7250         | 0.0540         | 0.0520         |
|                | MADET          | GYI            | VUS            | MADET          | GYI            | VUS            |
| 10             | 0.1134         | 0.0181         | 0.5534         | 0.0480         | 0.5773         | 0.4259         |
| 20             | 0.2130         | 0.9720         | 0.9930         | 0.7250         | 0.0540         | 0.0520         |
| 40             | 0.4556         | 0.9720         | 0.9930         | 0.7250         | 0.0540         | 0.0520         |
Gamma 4, MADET is 0.2960 under $H_a$, a big increase from 0.0516 under $H_0$, while GYI and VUS only increase very slightly. This shows that MADET can capture some differences between null and alternative hypotheses in classification rates that GYI and VUS cannot.

5 Selection of optimal cut-point(s)

In diagnostic studies, it is critical to determine the optimal cut-points of a continuous biomarker for the purpose of making diagnosis. Not all the diagnostic measures of accuracy that we have discussed can be utilized as criteria for optimal cut-point selection. For example, HUM cannot be used to select cut-points, since it is defined over all the possible cut-points. The proposed new measure, MADET in equation (8), not only serves as an overall diagnostic measure, as explained in Section 3, but also a criterion for the optimal cut-points selection. For cases with general $k$ stages, the set of cut-points $\hat{c} = (c_1, c_2, \ldots, c_{k-1})$ that maximizes the absolute value of the determinant $det(\mathcal{P})$, is the optimal cut-points determined by MADET.

In the following, we address the issues regarding cut-point selections using the proposed criterion MADET in comparison with several existing methods. A simulation study was performed to assess their relative performance.

5.1 Selection criteria

There are three popular methods for obtaining the optimal cut-off point in the binary outcome disease case: Youden index ($YI$),\textsuperscript{16,17} the closest-to-(0, 1) criterion (MD),\textsuperscript{36} and the maximum area method (MV).\textsuperscript{37} Note that the proposed MADET equals to Youden index for the binary case. Their corresponding objective functions, $YI(c_1)$ for the Youden index, $MD(c_1)$ for the closest-to-(0, 1) criterion, and $MV(c_1)$ for the maximum area approach, are as follows:

\[
YI(c_1) = P_{1,1} + P_{2,2} - 1, \\
MV(c_1) = P_{1,1} \times P_{2,2}, \\
MD(c_1) = \sqrt{(1 - P_{1,1})^2 + (1 - P_{2,2})^2},
\]

The optimal cut-off point is obtained by maximizing $YI(c_1)$, $MV(c_1)$ or minimizing $MD(c_1)$ for $c_1$ respectively. Liu\textsuperscript{37} also discussed the comparisons of optimal cut points selected by the three criteria.

For the 3-class disease case, we have as follows

\[
\text{MADET}(c_1, c_2) = \max_{c_1, c_2} |1 - P_{1,1}(1 - P_{2,2}) + P_{2,2}(1 - P_{3,3}) + P_{3,3}(1 - P_{1,1})| \\
- (P_{1,2}P_{2,1} + P_{1,3}P_{3,1} + P_{2,3}P_{3,2}), \\
GYI(c_1, c_2) = P_{1,1} + P_{2,2} + P_{3,3} - 1, \\
MV(c_1, c_2) = P_{1,1} \times P_{2,2} \times P_{3,3}, \\
MD(c_1, c_2) = \sqrt{(1 - P_{1,1})^2 + (1 - P_{2,2})^2 + (1 - P_{3,3})^2}
\]

The $MV(c_1, c_2)$, now called the max volume criterion, and $MD(c_1, c_2)$, the closest to perfection method, have been proposed by Attwood et al.\textsuperscript{39} Similarly, the optimal cut-off points are obtained by maximizing $GYI(c_1, c_2)$, $MV(c_1, c_2)$ or minimizing $MD(c_1, c_2)$ for $c_1$, $c_2$ respectively.
For general case with \( k \) stages, the Youden index, the closest-to-(0, 1) criterion, and the maximum volume (MV) approach can be easily extended to any \( k \) stages diseases as follows

\[
\text{MADET}(c_1, c_2, \ldots, c_{k-1}) = |\det(P)(c_1, \ldots, c_{k-1})|,
\]

\[
\text{GYI}(c_1, c_2, \ldots, c_{k-1}) = P_{1,1} + P_{2,2} + \cdots + P_{k,k} - 1,
\]

\[
\text{MV}(c_1, c_2, \ldots, c_{k-1}) = P_{1,1} * P_{2,2} * \cdots * P_{k,k},
\]

\[
\text{MD}(c_1, c_2, \ldots, c_{k-1}) = \sqrt{(1 - P_{1,1})^2 + (1 - P_{2,2})^2 + \cdots + (1 - P_{k,k})^2}
\]

### 5.2 Indexes to evaluate the selection criteria

Before we start to compare the different selection criteria, it is critical to establish a few judging rules or indexes. The performance indexes concerning the accuracy of the estimation of the cut-points include relative bias and root mean square error (RMSE). The relative bias for a statistic \( \hat{T} \) based on \( N \) rounds of iterations is

\[
\text{Relative bias} (\hat{T}) = \left[ \frac{1}{N} \sum_{i=1}^{N} \hat{T}_i - T \right]/T
\]

and its variance is

\[
\text{Variance} (\hat{T}) = \frac{1}{N} \sum_{i=1}^{N} (\hat{T}_i - T)^2/(N - 1)
\]

The RMSE is defined as \( \sqrt{\text{Bias}^2 + \text{Variance} (\hat{T})} \).

Another index for evaluating the selection criterion \( Q \) (MADET/GYI/MV/MD) is the total correct classification rate (TCCR) at the optimal cut-points \( \bar{c}_Q \), which is defined as

\[
\text{TCCR}_Q = \sum_{i=1}^{k} P_{i,i}(\bar{c}_Q).
\]

The maximum of TCCR, i.e. \( \text{TCCR}_{GYI} \), is achieved at the optimal cut-points \( \bar{c}_{GYI} \). Therefore, the loss of TCCR of MADET is defined as

\[
\text{Loss} = \left[ (\text{TCCR}_{GYI} - \text{TCCR}_Q)/\text{TCCR}_{GYI} \right] \times 100 \%
\]

The final consideration for cut-points selection is the balance of CCRs among classes. For example, for \( k = 2 \), Youden index 0.4 corresponds to many scenarios, e.g. 1) 0.7 sensitivity and 0.7 specificity, 2) 0.5 sensitivity and 0.9 specificity, 3) 0.9 sensitivity and 0.5 specificity. While the preference depends on the specific diseases, generally speaking, the balance between sensitivity and specificity is more desirable and one may favor scenario 1) instead of the other two. To measure the balance, we define a measure called the maximum minimum difference (MMDIF) as

\[
\text{MMDIF} = \left[ \max(P_{1,1}, \ldots, P_{k,k}) - \min(P_{1,1}, \ldots, P_{k,k}) \right]/\min(P_{1,1}, \ldots, P_{k,k})
\]

The smaller the MMDIF, the better the balance.
5.3 Simulations

Simulations were performed for the three-stage and four-stage diseases to evaluate the four selection criteria: MADET, GYI, MV and MD.

5.3.1 Three-stage diseases

Simulation was performed to compare MADET, GYI, MV and MD in selecting the optimal cut-off points for the three-class diseases. Two distributional settings from \(^{29}\) and four new ones are employed and listed in Table 4. The sample sizes were set as \((20, 20, 20)\), \((50, 50, 50)\), \((100, 100, 100)\) and separately. For a total of \(R = 10,000\) rounds of replications, the \(P_{ij}\)'s are estimated with the smoothed kernel distribution functions proposed by \(^{41}\). The empirical approach for the estimation was also attempted, but its performance was much worse than the kernel one. Hence, it was removed from consideration.

Tables 5 and 6 contain the simulation results for \(c_1\) and \(c_2\) given by MADET, GYI, MV and MD under the six different distributional settings. Generally estimates of \(c_1\) and \(c_2\) from MV and MD have the smallest RMSE, implying the estimates given by these two criteria are more accurate to the true cut-off points given by each criterion respectively. MADET has larger RMSE but is comparable or slightly better than GYI in terms of estimation accuracy.

Table 7 presents the loss of the TCCRs and the balance of the estimated CCRs. MADET has a smaller loss of total CCR for scenario 2, 4 and 6 than MD or MV. Moreover, MADET achieves better balance in scenario 1 and 3, especially for the larger or unbalanced sample sizes.

5.3.2 Four-stage diseases

The performance of MADET, GYI, MV and MD in the four-class disease scenario for determining optimal cut-points was also compared via a simulation study. Three scenarios from normal, gamma and exponential distributions respectively were selected and presented in Table 8. The sample sizes were set as \((20, 20, 20, 20)\), \((50, 50, 50, 50)\), and \((100, 100, 100, 100)\) separately. The \(P_{ij}\)'s are estimated with the smoothed kernel distribution functions, and \(c_1\), \(c_2\) and \(c_3\) are selected under the MADET, GYI, MV and MD criteria for \(R = 10,000\) rounds of replications. Relative Bias, RMSE, the Loss of total CCR and MMDIF are used as the key indexes to judge the performance of each criterion. The results are presented in Tables 9 to 12.

From Table 9 to 11, we can see that MD and MV provide more accurate estimate of cut-points than MADET and GYI. Table 12 shows that MADET can provide more balanced CCRs, i.e. small MMDIF with relatively higher percentage of loss of total CCR.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Non-diseased</th>
<th>Early diseased</th>
<th>Fully diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N(0, 1.0)</td>
<td>N(0.5, 1.0)</td>
<td>N(1, 1.0)</td>
</tr>
<tr>
<td>2</td>
<td>N(0, 1.2)</td>
<td>N(0.5, 0.8)</td>
<td>N(1, 1.4)</td>
</tr>
<tr>
<td>3</td>
<td>Gamma(2, 1)</td>
<td>Gamma(3, 1.0)</td>
<td>Gamma(4, 1.0)</td>
</tr>
<tr>
<td>4</td>
<td>Gamma(2, 1)</td>
<td>Gamma(7, 0.3)</td>
<td>Gamma(12, 0.3)</td>
</tr>
<tr>
<td>5</td>
<td>Exp(1)</td>
<td>Exp(0.9)</td>
<td>Exp(0.035)</td>
</tr>
<tr>
<td>6</td>
<td>Exp(1)</td>
<td>Exp(0.5)</td>
<td>Exp(0.035)</td>
</tr>
</tbody>
</table>
Alzheimer's disease (AD) is an irreversible neurodegenerative disease that results in a loss of mental function due to the deterioration of brain tissue. It is the most common cause of dementia among people over the age of 65, affecting an estimated 5.3 million Americans, but no prevention methods or cures have been discovered yet. The ADNI is a global research effort that aims to track the progression of the disease using biomarkers to assess the brain structure and function. It is a longitudinal study that assesses clinical, imaging, genetic and biospecimen biomarkers at 58 sites in the United States and Canada. The study had three phases: ADNI1, ADNIGO and ADNI2 and participants were followed and reassessed over time to track the pathology of the disease as it progresses.

The severity of dementia is staged by the clinical dementia rating (CDR) and a global CDR is derived from individual ratings in multiple domains by an experienced clinician. Here CDR 0 indicates no dementia and CDR 0.5, 1, 2, and 3 represent very mild, mild, moderate, and severe dementia, respectively. As patients with large CDR such as 2 or 3 are rarely available, we combined patients with CDR greater than or equal to 1 as the fully diseased group, while CDR 0 and 0.5 refer to the non-diseased and early diseased groups respectively. Data such as MRI and PET images, Table 5. Relative bias and root mean squared error of the $c_1$ estimates for the three-class disease.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>True $c_1$</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MADET:</td>
<td>$-0.229$</td>
<td>$-0.319$</td>
<td>$0.536$</td>
<td>$-0.3492$</td>
<td>$0.5179$</td>
<td>$-0.3469$</td>
<td>$0.5044$</td>
</tr>
<tr>
<td>Norm</td>
<td>1</td>
<td>GYI:</td>
<td>$0.250$</td>
<td>$-0.9106$</td>
<td>$0.8889$</td>
<td>$-0.4587$</td>
<td>$0.5027$</td>
<td>$-0.2353$</td>
<td>$0.3642$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$-0.035$</td>
<td>$0.8135$</td>
<td>$0.3250$</td>
<td>$0.6146$</td>
<td>$0.2100$</td>
<td>$0.4146$</td>
<td>$0.1545$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$0.046$</td>
<td>$0.3354$</td>
<td>$0.2800$</td>
<td>$0.3525$</td>
<td>$0.1929$</td>
<td>$0.2778$</td>
<td>$0.1442$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>MADET:</td>
<td>$-0.258$</td>
<td>$-0.0249$</td>
<td>$0.4517$</td>
<td>$0.0362$</td>
<td>$0.3166$</td>
<td>$0.0466$</td>
<td>$0.2402$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>$-0.238$</td>
<td>$0.0136$</td>
<td>$0.5243$</td>
<td>$0.0722$</td>
<td>$0.3227$</td>
<td>$0.0852$</td>
<td>$0.2427$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$-0.060$</td>
<td>$0.2558$</td>
<td>$0.2740$</td>
<td>$0.2659$</td>
<td>$0.1800$</td>
<td>$0.2391$</td>
<td>$0.1350$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$-0.066$</td>
<td>$0.1318$</td>
<td>$0.2612$</td>
<td>$0.2116$</td>
<td>$0.1746$</td>
<td>$0.2034$</td>
<td>$0.1318$</td>
</tr>
<tr>
<td>Gamma</td>
<td>3</td>
<td>MADET:</td>
<td>$1.382$</td>
<td>$0.4038$</td>
<td>$0.8079$</td>
<td>$0.3561$</td>
<td>$0.7193$</td>
<td>$0.2855$</td>
<td>$0.6196$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>$2.000$</td>
<td>$0.0891$</td>
<td>$0.7083$</td>
<td>$0.0817$</td>
<td>$0.4734$</td>
<td>$0.0680$</td>
<td>$0.3881$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$1.727$</td>
<td>$0.0678$</td>
<td>$0.3285$</td>
<td>$0.0454$</td>
<td>$0.2183$</td>
<td>$0.0321$</td>
<td>$0.1615$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$1.721$</td>
<td>$0.0699$</td>
<td>$0.3299$</td>
<td>$0.0462$</td>
<td>$0.2220$</td>
<td>$0.0322$</td>
<td>$0.1636$</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>MADET:</td>
<td>$7.048$</td>
<td>$-0.1004$</td>
<td>$1.5364$</td>
<td>$-0.0543$</td>
<td>$1.0757$</td>
<td>$-0.0380$</td>
<td>$0.8156$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>$7.269$</td>
<td>$-0.1169$</td>
<td>$1.6291$</td>
<td>$-0.0554$</td>
<td>$1.1499$</td>
<td>$-0.0344$</td>
<td>$0.8770$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$5.734$</td>
<td>$-0.0698$</td>
<td>$1.1294$</td>
<td>$-0.0548$</td>
<td>$0.7763$</td>
<td>$-0.0445$</td>
<td>$0.5848$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$7.098$</td>
<td>$-0.1081$</td>
<td>$1.5444$</td>
<td>$-0.0551$</td>
<td>$1.0854$</td>
<td>$-0.0377$</td>
<td>$0.8255$</td>
</tr>
<tr>
<td>Exp</td>
<td>5</td>
<td>MADET:</td>
<td>$1.030$</td>
<td>$0.1397$</td>
<td>$0.6499$</td>
<td>$0.1564$</td>
<td>$0.6410$</td>
<td>$0.1570$</td>
<td>$0.6398$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>$1.054$</td>
<td>$0.3155$</td>
<td>$1.3432$</td>
<td>$0.2821$</td>
<td>$1.2368$</td>
<td>$0.2536$</td>
<td>$1.0966$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$0.735$</td>
<td>$0.1281$</td>
<td>$0.2695$</td>
<td>$0.0867$</td>
<td>$0.1613$</td>
<td>$0.0633$</td>
<td>$0.1141$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$0.727$</td>
<td>$0.1486$</td>
<td>$0.2669$</td>
<td>$0.1003$</td>
<td>$0.1652$</td>
<td>$0.0728$</td>
<td>$0.1187$</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>MAD:</td>
<td>$1.341$</td>
<td>$0.1454$</td>
<td>$0.6779$</td>
<td>$0.1266$</td>
<td>$0.5328$</td>
<td>$0.0845$</td>
<td>$0.4076$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>$1.386$</td>
<td>$0.1748$</td>
<td>$0.7847$</td>
<td>$0.1179$</td>
<td>$0.5436$</td>
<td>$0.0729$</td>
<td>$0.4065$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>$1.026$</td>
<td>$0.1166$</td>
<td>$0.2800$</td>
<td>$0.0872$</td>
<td>$0.1802$</td>
<td>$0.0636$</td>
<td>$0.1333$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>$1.054$</td>
<td>$0.1446$</td>
<td>$0.3400$</td>
<td>$0.1021$</td>
<td>$0.2191$</td>
<td>$0.0732$</td>
<td>$0.1625$</td>
</tr>
</tbody>
</table>

R Bias: relative bias; RMSE: root mean square error.

6 ADNI example

Alzheimer’s disease (AD) is an irreversible neurodegenerative disease that results in a loss of mental function due to the deterioration of brain tissue. It is the most common cause of dementia among people over the age of 65, affecting an estimated 5.3 million Americans, but no prevention methods or cures have been discovered yet. The ADNI is a global research effort that aims to track the progression of the disease using biomarkers to assess the brain structure and function. It is a longitudinal study that assesses clinical, imaging, genetic and biospecimen biomarkers at 58 sites in the United States and Canada. The study had three phases: ADNI1, ADNIGO and ADNI2 and participants were followed and reassessed over time to track the pathology of the disease as it progresses.

The severity of dementia is staged by the clinical dementia rating (CDR) and a global CDR is derived from individual ratings in multiple domains by an experienced clinician. Here CDR 0 indicates no dementia and CDR 0.5, 1, 2, and 3 represent very mild, mild, moderate, and severe dementia, respectively. As patients with large CDR such as 2 or 3 are rarely available, we combined patients with CDR greater than or equal to 1 as the fully diseased group, while CDR 0 and 0.5 refer to the non-diseased and early diseased groups respectively. Data such as MRI and PET images,
genetics, cognitive tests, cerebrospinal fluid (CSF) and blood biomarkers are collected. Neuroimaging measures include brain volume, ventricular volume, and bilateral hippocampal volumes. Cognitive measures represent five domains respectively: memory, language, executive function, spatial ability, and attention. The CSF variables include T-tau, Aβ42, p-tau181, the ratio of the first two variables, and the ratio of the last two variables. Genetic and other lab tests are not considered in our case. According to hypotheses proposed by Heneka et al.42 and Mudher and Lovestone, 43 the biomarkers of interest include hippocampus volume (Hippocampus), brain volume (WholeBrain), T-tau (TAU), Aβ42 (ABETA142), and p-tau181 (PTAU181P).

For the purpose of this paper, only the data at the 24th month visit is included and the disease status defined at that time point is adopted. There are 194, 290 and 183 subjects for the non-diseased, the early diseased, and the fully diseased groups respectively. However, due to missing values, the actual sample sizes for each variable are smaller. We use three diagnostic accuracy measures MADET, GYI and VUS, as well as the cut-off points selection criteria MD and MV discussed in the paper to evaluate the five biomarkers of interests: hippocampus volume (Hippocampus), brain volume (WholeBrain), T-tau (TAU), Aβ42 (ABETA142), and p-tau181 (PTAU181P). The corresponding cut-points are determined by MADET, GYI, MD and MV, respectively.

### Table 6. Relative bias and root mean squared error of the c2 estimates for the three-class disease.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>True c2</th>
<th>n = m = k = 20</th>
<th>n = m = k = 50</th>
<th>n = m = k = 100</th>
<th>n = 50, m = 30, k = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm</td>
<td>1</td>
<td>MADET:</td>
<td>1.229</td>
<td>−0.0134</td>
<td>0.5385</td>
<td>−0.0395</td>
<td>0.5221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>0.750</td>
<td>0.4384</td>
<td>1.0547</td>
<td>0.1902</td>
<td>0.5286</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>1.035</td>
<td>0.0302</td>
<td>0.3180</td>
<td>0.0165</td>
<td>0.2082</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>1.046</td>
<td>0.0191</td>
<td>0.2793</td>
<td>0.0127</td>
<td>0.1925</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>MADET:</td>
<td>1.356</td>
<td>0.0240</td>
<td>0.4248</td>
<td>0.0256</td>
<td>0.2921</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>1.373</td>
<td>0.0383</td>
<td>0.4546</td>
<td>0.0319</td>
<td>0.2913</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>1.171</td>
<td>0.0246</td>
<td>0.2992</td>
<td>0.0214</td>
<td>0.1922</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>1.171</td>
<td>0.0174</td>
<td>0.2729</td>
<td>0.0188</td>
<td>0.1836</td>
</tr>
<tr>
<td>Gamma</td>
<td>3</td>
<td>MADET:</td>
<td>3.618</td>
<td>0.1469</td>
<td>1.1534</td>
<td>0.1223</td>
<td>1.0332</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>3.000</td>
<td>0.1934</td>
<td>1.3685</td>
<td>0.1200</td>
<td>0.8397</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>3.551</td>
<td>0.0553</td>
<td>0.5976</td>
<td>0.0375</td>
<td>0.3960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>3.559</td>
<td>0.0532</td>
<td>0.5590</td>
<td>0.0369</td>
<td>0.3809</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>MADET:</td>
<td>29.660</td>
<td>0.0211</td>
<td>2.9509</td>
<td>0.0143</td>
<td>2.0614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>29.624</td>
<td>0.0190</td>
<td>2.9355</td>
<td>0.0128</td>
<td>2.0470</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>29.989</td>
<td>0.0077</td>
<td>2.0260</td>
<td>0.0052</td>
<td>1.3158</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>29.820</td>
<td>0.0137</td>
<td>2.3685</td>
<td>0.0095</td>
<td>1.6081</td>
</tr>
<tr>
<td>Exp</td>
<td>5</td>
<td>MADET:</td>
<td>5.950</td>
<td>−0.3002</td>
<td>2.2638</td>
<td>−0.2763</td>
<td>2.1732</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>3.754</td>
<td>0.1196</td>
<td>1.2012</td>
<td>0.1551</td>
<td>0.9935</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>5.045</td>
<td>−0.1109</td>
<td>1.3829</td>
<td>−0.0010</td>
<td>1.0497</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>4.378</td>
<td>−0.0110</td>
<td>1.2128</td>
<td>0.0727</td>
<td>0.9868</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>MADET:</td>
<td>7.857</td>
<td>−0.1201</td>
<td>2.3486</td>
<td>−0.0178</td>
<td>1.7272</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>5.719</td>
<td>0.1106</td>
<td>1.7490</td>
<td>0.1211</td>
<td>1.3364</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>7.145</td>
<td>−0.0082</td>
<td>1.8531</td>
<td>0.0455</td>
<td>1.3349</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>6.508</td>
<td>0.0444</td>
<td>1.8112</td>
<td>0.0761</td>
<td>1.3311</td>
</tr>
</tbody>
</table>

R Bias: relative bias; RMSE: root mean square error.
Table 7. Correct classification rate of each disease group, total CCR and MMDIF for the three-class disease.

<table>
<thead>
<tr>
<th>Sample size Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>$P_{1,1}$</th>
<th>$P_{2,2}$</th>
<th>$P_{3,3}$</th>
<th>Total CCR</th>
<th>Loss of CCR(%)</th>
<th>MMDIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = m = k = 20$ Norm</td>
<td>1</td>
<td>MADET</td>
<td>0.4615</td>
<td>0.4616</td>
<td>0.4412</td>
<td>1.3643</td>
<td>6.1400</td>
<td>0.0462</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.5622</td>
<td>0.3537</td>
<td>0.5377</td>
<td>1.4536</td>
<td>0.0000</td>
<td>0.5895</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.4945</td>
<td>0.4265</td>
<td>0.4929</td>
<td>1.4138</td>
<td>2.7300</td>
<td>0.1594</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.4914</td>
<td>0.4251</td>
<td>0.4894</td>
<td>1.4059</td>
<td>3.2800</td>
<td>0.1560</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>MADET</td>
<td>0.4429</td>
<td>0.6646</td>
<td>0.4150</td>
<td>1.5224</td>
<td>1.1100</td>
<td>0.6014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.4541</td>
<td>0.6748</td>
<td>0.4106</td>
<td>1.5395</td>
<td>0.0000</td>
<td>0.6434</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.4884</td>
<td>0.5658</td>
<td>0.4584</td>
<td>1.5126</td>
<td>1.7400</td>
<td>0.2343</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.4883</td>
<td>0.5629</td>
<td>0.4595</td>
<td>1.5107</td>
<td>1.8700</td>
<td>0.2250</td>
</tr>
<tr>
<td>Gamma 3</td>
<td></td>
<td>MADET</td>
<td>0.5588</td>
<td>0.4598</td>
<td>0.4462</td>
<td>1.4648</td>
<td>4.7600</td>
<td>0.2524</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6367</td>
<td>0.3252</td>
<td>0.5761</td>
<td>1.5381</td>
<td>0.0000</td>
<td>0.9579</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.5425</td>
<td>0.4410</td>
<td>0.5149</td>
<td>1.4984</td>
<td>2.5800</td>
<td>0.2302</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.5415</td>
<td>0.4415</td>
<td>0.5124</td>
<td>1.4954</td>
<td>2.7700</td>
<td>0.2265</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>MADET</td>
<td>0.9968</td>
<td>0.7845</td>
<td>0.7901</td>
<td>2.5715</td>
<td>0.0100</td>
<td>0.2706</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.9980</td>
<td>0.7808</td>
<td>0.7928</td>
<td>2.5716</td>
<td>0.0000</td>
<td>0.2782</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.9766</td>
<td>0.7853</td>
<td>0.7896</td>
<td>2.5514</td>
<td>0.7900</td>
<td>0.2436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.9970</td>
<td>0.7821</td>
<td>0.7909</td>
<td>2.5700</td>
<td>0.0600</td>
<td>0.2748</td>
</tr>
<tr>
<td>Exp 5</td>
<td></td>
<td>MADET</td>
<td>0.6303</td>
<td>0.3898</td>
<td>0.8316</td>
<td>1.8517</td>
<td>4.5500</td>
<td>1.1334</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6666</td>
<td>0.4421</td>
<td>0.8313</td>
<td>1.9399</td>
<td>0.0000</td>
<td>0.8803</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.5433</td>
<td>0.5202</td>
<td>0.8259</td>
<td>1.8894</td>
<td>2.6000</td>
<td>0.5877</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.5464</td>
<td>0.5145</td>
<td>0.8288</td>
<td>1.8897</td>
<td>2.5900</td>
<td>0.6109</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>MADET</td>
<td>0.7652</td>
<td>0.4970</td>
<td>0.7794</td>
<td>2.0416</td>
<td>0.8000</td>
<td>0.5682</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.7874</td>
<td>0.4792</td>
<td>0.7916</td>
<td>2.0581</td>
<td>0.0000</td>
<td>0.6519</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.6699</td>
<td>0.5832</td>
<td>0.7771</td>
<td>2.0302</td>
<td>1.3600</td>
<td>0.3325</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.6908</td>
<td>0.5657</td>
<td>0.7826</td>
<td>2.0391</td>
<td>0.9200</td>
<td>0.3834</td>
</tr>
<tr>
<td>$n = m = k = 50$ Norm</td>
<td>1</td>
<td>MADET</td>
<td>0.4598</td>
<td>0.4528</td>
<td>0.4501</td>
<td>1.3628</td>
<td>4.0900</td>
<td>0.0216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.5709</td>
<td>0.2870</td>
<td>0.5629</td>
<td>1.4208</td>
<td>0.0000</td>
<td>0.9892</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.4870</td>
<td>0.4140</td>
<td>0.4894</td>
<td>1.3904</td>
<td>2.1400</td>
<td>0.1821</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.4834</td>
<td>0.4177</td>
<td>0.4853</td>
<td>1.3865</td>
<td>2.4200</td>
<td>0.1618</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>MADET</td>
<td>0.4275</td>
<td>0.6782</td>
<td>0.4041</td>
<td>1.5098</td>
<td>0.2600</td>
<td>0.6783</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.4329</td>
<td>0.6827</td>
<td>0.3982</td>
<td>1.5138</td>
<td>0.0000</td>
<td>0.7145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.4813</td>
<td>0.5607</td>
<td>0.4518</td>
<td>1.4939</td>
<td>1.3200</td>
<td>0.2410</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.4802</td>
<td>0.5609</td>
<td>0.4523</td>
<td>1.4934</td>
<td>1.3500</td>
<td>0.2401</td>
</tr>
<tr>
<td>Gamma 3</td>
<td></td>
<td>MADET</td>
<td>0.5431</td>
<td>0.4640</td>
<td>0.4579</td>
<td>1.4650</td>
<td>2.9300</td>
<td>0.1861</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6284</td>
<td>0.2877</td>
<td>0.5931</td>
<td>1.5092</td>
<td>0.0000</td>
<td>1.1842</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.5309</td>
<td>0.4354</td>
<td>0.5138</td>
<td>1.4801</td>
<td>1.9300</td>
<td>0.2193</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.5293</td>
<td>0.4377</td>
<td>0.5118</td>
<td>1.4788</td>
<td>2.0200</td>
<td>0.2093</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>MADET</td>
<td>0.9950</td>
<td>0.7789</td>
<td>0.7941</td>
<td>2.5680</td>
<td>0.0000</td>
<td>0.2774</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.9967</td>
<td>0.7751</td>
<td>0.7963</td>
<td>2.5681</td>
<td>0.0000</td>
<td>0.2859</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.9740</td>
<td>0.7851</td>
<td>0.7913</td>
<td>2.5503</td>
<td>0.6900</td>
<td>0.2406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.9954</td>
<td>0.7787</td>
<td>0.7932</td>
<td>2.5673</td>
<td>0.0300</td>
<td>0.2783</td>
</tr>
<tr>
<td>Exp 5</td>
<td></td>
<td>MADET</td>
<td>0.6387</td>
<td>0.3765</td>
<td>0.8328</td>
<td>1.8480</td>
<td>2.9200</td>
<td>1.2120</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6539</td>
<td>0.4166</td>
<td>0.8331</td>
<td>1.9036</td>
<td>0.0000</td>
<td>0.9998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.5325</td>
<td>0.5125</td>
<td>0.8194</td>
<td>1.8644</td>
<td>2.0600</td>
<td>0.5988</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.5331</td>
<td>0.5083</td>
<td>0.8262</td>
<td>1.8676</td>
<td>1.8900</td>
<td>0.6254</td>
</tr>
</tbody>
</table>

(continued)
Table 7. Continued

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>$P_{1,1}$</th>
<th>$P_{2,2}$</th>
<th>$P_{3,3}$</th>
<th>Total CCR</th>
<th>Loss of CCR(%)</th>
<th>MMDIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>MAD</td>
<td></td>
<td></td>
<td>0.7638</td>
<td>0.4949</td>
<td>0.7651</td>
<td>2.0238</td>
<td>0.4900</td>
<td>0.5460</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.7729</td>
<td>0.4686</td>
<td>0.7923</td>
<td>2.0338</td>
<td>0.4949</td>
<td>0.6908</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.6599</td>
<td>0.5807</td>
<td>0.7708</td>
<td>2.0115</td>
<td>1.1000</td>
<td>0.3274</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.6760</td>
<td>0.5631</td>
<td>0.7803</td>
<td>2.0194</td>
<td>0.7100</td>
<td>0.3857</td>
</tr>
<tr>
<td>2</td>
<td>MADET</td>
<td>1</td>
<td>Norm</td>
<td>0.4573</td>
<td>0.4556</td>
<td>0.4479</td>
<td>1.3607</td>
<td>3.3500</td>
<td>0.0210</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.5828</td>
<td>0.2465</td>
<td>0.5785</td>
<td>1.4079</td>
<td>0.0000</td>
<td>1.3643</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.4853</td>
<td>0.4093</td>
<td>0.4869</td>
<td>1.3816</td>
<td>1.8700</td>
<td>0.1896</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.4810</td>
<td>0.4152</td>
<td>0.4827</td>
<td>1.3789</td>
<td>2.0600</td>
<td>0.1626</td>
</tr>
<tr>
<td>4</td>
<td>MADET</td>
<td>2</td>
<td>Gamma</td>
<td>0.9936</td>
<td>0.7773</td>
<td>0.7973</td>
<td>2.5683</td>
<td>0.0100</td>
<td>0.2783</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.9954</td>
<td>0.7737</td>
<td>0.7993</td>
<td>2.5685</td>
<td>0.0000</td>
<td>0.2865</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.9737</td>
<td>0.7856</td>
<td>0.7932</td>
<td>2.5524</td>
<td>0.6300</td>
<td>0.2394</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.9940</td>
<td>0.7782</td>
<td>0.7956</td>
<td>2.5679</td>
<td>0.0200</td>
<td>0.2773</td>
</tr>
<tr>
<td>5</td>
<td>MADET</td>
<td>3</td>
<td>Exp</td>
<td>0.6414</td>
<td>0.3729</td>
<td>0.8351</td>
<td>1.8494</td>
<td>2.2000</td>
<td>1.2395</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.6548</td>
<td>0.3973</td>
<td>0.8390</td>
<td>1.8911</td>
<td>0.0000</td>
<td>1.1118</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.5284</td>
<td>0.5102</td>
<td>0.8201</td>
<td>1.8588</td>
<td>1.7100</td>
<td>0.6074</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.5276</td>
<td>0.5064</td>
<td>0.8299</td>
<td>1.8639</td>
<td>1.4400</td>
<td>0.6388</td>
</tr>
<tr>
<td>6</td>
<td>MADET</td>
<td>4</td>
<td>Norm</td>
<td>0.7548</td>
<td>0.4970</td>
<td>0.7627</td>
<td>2.0145</td>
<td>0.5700</td>
<td>0.3546</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.7629</td>
<td>0.4668</td>
<td>0.7964</td>
<td>2.0261</td>
<td>0.0000</td>
<td>0.7061</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.6538</td>
<td>0.5794</td>
<td>0.7729</td>
<td>2.0661</td>
<td>0.9900</td>
<td>0.3340</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.6678</td>
<td>0.5621</td>
<td>0.7835</td>
<td>2.0134</td>
<td>0.6300</td>
<td>0.3939</td>
</tr>
<tr>
<td>2</td>
<td>MADET</td>
<td>5</td>
<td>Gamma</td>
<td>0.4464</td>
<td>0.4544</td>
<td>0.4633</td>
<td>1.3641</td>
<td>5.1800</td>
<td>0.0379</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.5634</td>
<td>0.3209</td>
<td>0.5543</td>
<td>1.4386</td>
<td>0.0000</td>
<td>0.7557</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.4922</td>
<td>0.4193</td>
<td>0.4910</td>
<td>1.4026</td>
<td>2.5000</td>
<td>0.1739</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.4871</td>
<td>0.4211</td>
<td>0.4888</td>
<td>1.3970</td>
<td>2.8900</td>
<td>0.1608</td>
</tr>
<tr>
<td>4</td>
<td>MADET</td>
<td>6</td>
<td>Exp</td>
<td>0.5312</td>
<td>0.4616</td>
<td>0.4761</td>
<td>1.4688</td>
<td>3.7200</td>
<td>0.1508</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.6280</td>
<td>0.3039</td>
<td>0.5936</td>
<td>1.5255</td>
<td>0.0000</td>
<td>1.0665</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.5352</td>
<td>0.4396</td>
<td>0.5168</td>
<td>1.4916</td>
<td>2.2200</td>
<td>0.2175</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.5323</td>
<td>0.4414</td>
<td>0.5156</td>
<td>1.4894</td>
<td>2.3700</td>
<td>0.2059</td>
</tr>
<tr>
<td>5</td>
<td>MADET</td>
<td>3</td>
<td>Gamma</td>
<td>0.9942</td>
<td>0.7804</td>
<td>0.7929</td>
<td>2.5675</td>
<td>0.0100</td>
<td>0.2740</td>
</tr>
<tr>
<td></td>
<td>GYI</td>
<td></td>
<td></td>
<td>0.9961</td>
<td>0.7761</td>
<td>0.7954</td>
<td>2.5677</td>
<td>0.0000</td>
<td>0.2835</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td></td>
<td></td>
<td>0.9734</td>
<td>0.7851</td>
<td>0.7896</td>
<td>2.5480</td>
<td>0.7700</td>
<td>0.2398</td>
</tr>
<tr>
<td></td>
<td>MV</td>
<td></td>
<td></td>
<td>0.9946</td>
<td>0.7802</td>
<td>0.7916</td>
<td>2.5664</td>
<td>0.0500</td>
<td>0.2748</td>
</tr>
</tbody>
</table>
Table 9. Relative bias and root mean squared error of the $c_1$ estimates for the four-class disease.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>$P_{1,1}$</th>
<th>$P_{2,2}$</th>
<th>$P_{3,3}$</th>
<th>Total CCR</th>
<th>Loss of CCR(%)</th>
<th>MMDIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>N(0, 1)</td>
<td>Exp</td>
<td>MADET</td>
<td>0.5681</td>
<td>0.4261</td>
<td>0.8346</td>
<td>1.8287</td>
<td>4.2600</td>
<td>0.9587</td>
</tr>
<tr>
<td></td>
<td>N(0.5, 1)</td>
<td></td>
<td>GYI</td>
<td>0.6279</td>
<td>0.4531</td>
<td>0.8291</td>
<td>1.9101</td>
<td>0.0000</td>
<td>0.8298</td>
</tr>
<tr>
<td></td>
<td>N(1, 1)</td>
<td></td>
<td>MD</td>
<td>0.5339</td>
<td>0.5166</td>
<td>0.8201</td>
<td>1.8706</td>
<td>2.0700</td>
<td>0.5875</td>
</tr>
<tr>
<td></td>
<td>N(1.5, 1)</td>
<td></td>
<td>MV</td>
<td>0.5352</td>
<td>0.5128</td>
<td>0.8243</td>
<td>1.8723</td>
<td>1.9800</td>
<td>0.6074</td>
</tr>
<tr>
<td>50</td>
<td>Gamma(2, 1)</td>
<td>Exp</td>
<td>MADET</td>
<td>0.7439</td>
<td>0.5074</td>
<td>0.7698</td>
<td>2.0210</td>
<td>0.8900</td>
<td>0.5171</td>
</tr>
<tr>
<td></td>
<td>Gamma(3, 1)</td>
<td></td>
<td>GYI</td>
<td>0.7750</td>
<td>0.4751</td>
<td>0.7891</td>
<td>2.0392</td>
<td>0.0000</td>
<td>0.6609</td>
</tr>
<tr>
<td></td>
<td>Gamma(4, 1)</td>
<td></td>
<td>MD</td>
<td>0.6619</td>
<td>0.5842</td>
<td>0.7712</td>
<td>2.0172</td>
<td>1.0800</td>
<td>0.3201</td>
</tr>
<tr>
<td></td>
<td>Gamma(5, 1)</td>
<td></td>
<td>MV</td>
<td>0.6787</td>
<td>0.5682</td>
<td>0.7777</td>
<td>2.0246</td>
<td>0.7200</td>
<td>0.3687</td>
</tr>
<tr>
<td>100</td>
<td>Exp(1)</td>
<td>Exp</td>
<td>MADET</td>
<td>0.7439</td>
<td>0.5074</td>
<td>0.7698</td>
<td>2.0210</td>
<td>0.8900</td>
<td>0.5171</td>
</tr>
<tr>
<td></td>
<td>Exp(0.9)</td>
<td></td>
<td>GYI</td>
<td>0.7750</td>
<td>0.4751</td>
<td>0.7891</td>
<td>2.0392</td>
<td>0.0000</td>
<td>0.6609</td>
</tr>
<tr>
<td></td>
<td>Exp(0.035)</td>
<td></td>
<td>MD</td>
<td>0.6619</td>
<td>0.5842</td>
<td>0.7712</td>
<td>2.0172</td>
<td>1.0800</td>
<td>0.3201</td>
</tr>
<tr>
<td></td>
<td>Exp(0.02)</td>
<td></td>
<td>MV</td>
<td>0.6787</td>
<td>0.5682</td>
<td>0.7777</td>
<td>2.0246</td>
<td>0.7200</td>
<td>0.3687</td>
</tr>
</tbody>
</table>

MMDIF is defined as: $[\max(P_{1,1}, P_{2,2}, P_{3,3}) - \min(P_{1,1}, P_{2,2}, P_{3,3})]/\min(P_{1,1}, P_{2,2}, P_{3,3})$.

Table 8. Simulation distributional scenarios for the four-class disease.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Non-diseased</th>
<th>Early diseased</th>
<th>Fully diseased</th>
<th>Severely diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N(0, 1)</td>
<td>N(0.5, 1)</td>
<td>N(1, 1)</td>
<td>N(1.5, 1)</td>
</tr>
<tr>
<td>2</td>
<td>Gamma(2, 1)</td>
<td>Gamma(3, 1)</td>
<td>Gamma(4, 1)</td>
<td>Gamma(5, 1)</td>
</tr>
<tr>
<td>3</td>
<td>Exp(1)</td>
<td>Exp(0.9)</td>
<td>Exp(0.035)</td>
<td>Exp(0.02)</td>
</tr>
</tbody>
</table>

Table 9. Relative bias and root mean squared error of the $c_1$ estimates for the four-class disease.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>True $c_1$</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm</td>
<td>1</td>
<td>MADET</td>
<td>-0.527</td>
<td>-0.1970</td>
<td>0.4948</td>
<td>-0.2518</td>
<td>0.4726</td>
<td>-0.3078</td>
<td>0.4569</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>0.250</td>
<td>-1.2528</td>
<td>0.9475</td>
<td>-0.5683</td>
<td>0.5067</td>
<td>-0.3090</td>
<td>0.3711</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>-0.129</td>
<td>0.3852</td>
<td>0.3845</td>
<td>0.2579</td>
<td>0.2339</td>
<td>0.2062</td>
<td>0.1776</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>-0.174</td>
<td>0.1590</td>
<td>0.2876</td>
<td>0.1417</td>
<td>0.1987</td>
<td>0.1286</td>
<td>0.1511</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>MADET</td>
<td>1.071</td>
<td>0.5591</td>
<td>0.7773</td>
<td>0.5310</td>
<td>0.6993</td>
<td>0.5016</td>
<td>0.6487</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>2.000</td>
<td>0.0802</td>
<td>0.6702</td>
<td>0.0770</td>
<td>0.4652</td>
<td>0.0683</td>
<td>0.3829</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>1.616</td>
<td>0.0740</td>
<td>0.3324</td>
<td>0.0484</td>
<td>0.2174</td>
<td>0.0367</td>
<td>0.1655</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>1.576</td>
<td>0.0698</td>
<td>0.3079</td>
<td>0.0437</td>
<td>0.2044</td>
<td>0.0328</td>
<td>0.1554</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>MADET</td>
<td>1.017</td>
<td>0.0555</td>
<td>0.7202</td>
<td>0.1099</td>
<td>0.6624</td>
<td>0.1185</td>
<td>0.5942</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI</td>
<td>1.054</td>
<td>0.2985</td>
<td>1.4941</td>
<td>0.2854</td>
<td>1.4199</td>
<td>0.2593</td>
<td>1.2601</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD</td>
<td>0.716</td>
<td>0.1500</td>
<td>0.2762</td>
<td>0.1059</td>
<td>0.1674</td>
<td>0.0808</td>
<td>0.1205</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV</td>
<td>0.716</td>
<td>0.1582</td>
<td>0.2685</td>
<td>0.1119</td>
<td>0.1694</td>
<td>0.0847</td>
<td>0.1236</td>
</tr>
</tbody>
</table>

R Bias: relative bias; RMSE: root mean square error.
The summary statistics of the five biomarkers are presented in Table 13, and optimality statistics, optimal cut-points, CCRs are presented in Table 14. Note that for Hippocampus, WholeBrain and ABETA142, smaller values indicate a more diseased status.

Hippocampus has the highest estimated MADET and VUS among the five biomarkers, while ABETA142 has the highest estimated GYI. Furthermore, the intervals between the two cut-points

Table 10. Relative bias and root mean squared error of the $c_2$ estimates for the four-class disease.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>True $c_1$</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm</td>
<td>1</td>
<td>MADET:</td>
<td>0.750</td>
<td>0.0116</td>
<td>0.4052</td>
<td>0.0066</td>
<td>0.3943</td>
<td>0.0125</td>
<td>0.3898</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>0.750</td>
<td>0.0250</td>
<td>0.4527</td>
<td>0.0190</td>
<td>0.3491</td>
<td>0.0096</td>
<td>0.2936</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>0.750</td>
<td>0.0059</td>
<td>0.3288</td>
<td>-0.0007</td>
<td>0.2370</td>
<td>-0.0020</td>
<td>0.1851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>0.750</td>
<td>0.0041</td>
<td>0.2698</td>
<td>-0.0006</td>
<td>0.1924</td>
<td>-0.0025</td>
<td>0.1479</td>
</tr>
<tr>
<td>Gamma</td>
<td>2</td>
<td>MADET:</td>
<td>2.658</td>
<td>0.2839</td>
<td>1.0222</td>
<td>0.2705</td>
<td>0.9573</td>
<td>0.2548</td>
<td>0.9134</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>3.000</td>
<td>0.0940</td>
<td>0.7913</td>
<td>0.0644</td>
<td>0.6084</td>
<td>0.0494</td>
<td>0.5038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>3.004</td>
<td>0.0728</td>
<td>0.5621</td>
<td>0.0490</td>
<td>0.3906</td>
<td>0.0367</td>
<td>0.3027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>3.008</td>
<td>0.0646</td>
<td>0.4776</td>
<td>0.0440</td>
<td>0.3329</td>
<td>0.0329</td>
<td>0.2564</td>
</tr>
<tr>
<td>Exp</td>
<td>3</td>
<td>MADET:</td>
<td>5.422</td>
<td>-0.2774</td>
<td>2.0902</td>
<td>-0.2519</td>
<td>1.9733</td>
<td>-0.2403</td>
<td>1.9430</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>3.754</td>
<td>0.1335</td>
<td>1.2305</td>
<td>0.1649</td>
<td>1.0274</td>
<td>0.1413</td>
<td>0.8199</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>3.896</td>
<td>0.0776</td>
<td>1.1724</td>
<td>0.1228</td>
<td>0.9569</td>
<td>0.1094</td>
<td>0.7587</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>3.888</td>
<td>0.0789</td>
<td>1.1698</td>
<td>0.1232</td>
<td>0.9527</td>
<td>0.1105</td>
<td>0.7590</td>
</tr>
</tbody>
</table>

R Bias: relative bias; RMSE: root mean square error.

Table 11. Relative bias and root mean squared error of the $c_3$ estimates for the four-class disease.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>True $c_1$</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
<th>R Bias</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm</td>
<td>1</td>
<td>MADET:</td>
<td>2.027</td>
<td>-0.0407</td>
<td>0.4971</td>
<td>-0.0601</td>
<td>0.4718</td>
<td>-0.0725</td>
<td>0.4558</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>1.250</td>
<td>0.3084</td>
<td>1.1176</td>
<td>0.1421</td>
<td>0.5648</td>
<td>0.0750</td>
<td>0.3721</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>1.629</td>
<td>0.0380</td>
<td>0.4080</td>
<td>0.0196</td>
<td>0.2316</td>
<td>0.0142</td>
<td>0.1757</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>1.674</td>
<td>0.0211</td>
<td>0.2858</td>
<td>0.0143</td>
<td>0.1958</td>
<td>0.0114</td>
<td>0.1503</td>
</tr>
<tr>
<td>Gamma</td>
<td>2</td>
<td>MADET:</td>
<td>5.272</td>
<td>0.0673</td>
<td>1.1563</td>
<td>0.0504</td>
<td>1.0453</td>
<td>0.0325</td>
<td>0.9674</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>4.000</td>
<td>0.2049</td>
<td>1.8667</td>
<td>0.1293</td>
<td>1.1143</td>
<td>0.0850</td>
<td>0.8185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>4.839</td>
<td>0.0577</td>
<td>0.8798</td>
<td>0.0387</td>
<td>0.5618</td>
<td>0.0278</td>
<td>0.4115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>4.922</td>
<td>0.0482</td>
<td>0.7011</td>
<td>0.0333</td>
<td>0.4786</td>
<td>0.0252</td>
<td>0.3586</td>
</tr>
<tr>
<td>Exp</td>
<td>3</td>
<td>MADET:</td>
<td>42.996</td>
<td>-0.0940</td>
<td>16.3399</td>
<td>-0.0723</td>
<td>13.0439</td>
<td>-0.0601</td>
<td>11.1184</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GYI:</td>
<td>37.308</td>
<td>0.0383</td>
<td>17.7952</td>
<td>0.0350</td>
<td>13.1900</td>
<td>0.0308</td>
<td>11.0858</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD:</td>
<td>32.623</td>
<td>0.1319</td>
<td>9.7402</td>
<td>0.1013</td>
<td>6.5223</td>
<td>0.0817</td>
<td>4.8652</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MV:</td>
<td>32.791</td>
<td>0.1330</td>
<td>9.7954</td>
<td>0.0998</td>
<td>6.6383</td>
<td>0.0794</td>
<td>4.9482</td>
</tr>
</tbody>
</table>

R Bias: relative bias; RMSE: root mean square error.
by GYI are shorter than those by MADET, implying that the classification criterion based on GYI tends to diagnose fewer subjects into the early diseased group than that based on MADET. Especially, for biomarkers Tau and PTAU181P, the cut-points $c_1$ and $c_2$ estimated by GYI coincide and the corresponding $P_{2,2}$'s are 0, which means none of the subjects will be diagnosed into the early diseased stage by GYI criterion. Such cut-points obviously lead to an inappropriate diagnosis since the results contradict with the three-stage setting for Alzheimer’s disease. The estimates given by MD and MV are also reasonable. Hence, for Tau and PTAU181P, MADET/MD/MV might be better criteria to use for diagnosis of the three-stage Alzheimer’s Disease.

### Table 12. Correct classification rate of each disease group, total CCR and MMDIF for the four-class disease.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Distribution</th>
<th>Scenario</th>
<th>Method</th>
<th>$P_{1,1}$</th>
<th>$P_{1,2}$</th>
<th>$P_{2,2}$</th>
<th>$P_{3,2}$</th>
<th>$P_{4,2}$</th>
<th>Total CCR</th>
<th>Loss of CCR(%)</th>
<th>MMDIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = m = k = l = 20$</td>
<td>Norm</td>
<td>1</td>
<td>MADET</td>
<td>0.3658</td>
<td>0.3917</td>
<td>0.3900</td>
<td>0.3612</td>
<td>1.5087</td>
<td>9.6800</td>
<td>0.0844</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.5370</td>
<td>0.3047</td>
<td>0.3049</td>
<td>0.5238</td>
<td>1.6704</td>
<td>0.0000</td>
<td>0.7624</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.4557</td>
<td>0.3539</td>
<td>0.3543</td>
<td>0.4535</td>
<td>1.6174</td>
<td>3.1700</td>
<td>0.2877</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.4390</td>
<td>0.3566</td>
<td>0.3571</td>
<td>0.4376</td>
<td>1.5903</td>
<td>4.8000</td>
<td>0.2311</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gamma</td>
<td>2</td>
<td>MADET</td>
<td>0.4862</td>
<td>0.3993</td>
<td>0.3553</td>
<td>0.3779</td>
<td>1.6188</td>
<td>7.9300</td>
<td>0.3684</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6334</td>
<td>0.2872</td>
<td>0.2942</td>
<td>0.5434</td>
<td>1.7582</td>
<td>0.0000</td>
<td>1.2054</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.5124</td>
<td>0.3690</td>
<td>0.3606</td>
<td>0.4600</td>
<td>1.7019</td>
<td>3.2000</td>
<td>0.4210</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.4961</td>
<td>0.3726</td>
<td>0.3665</td>
<td>0.4455</td>
<td>1.6806</td>
<td>4.4100</td>
<td>0.3536</td>
<td></td>
</tr>
<tr>
<td>$n = m = k = l = 50$</td>
<td>Norm</td>
<td>1</td>
<td>MADET</td>
<td>0.3728</td>
<td>0.3850</td>
<td>0.3847</td>
<td>0.3696</td>
<td>1.5121</td>
<td>7.0400</td>
<td>0.0417</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.5614</td>
<td>0.2573</td>
<td>0.2570</td>
<td>0.5509</td>
<td>1.6267</td>
<td>0.0000</td>
<td>1.1844</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.4484</td>
<td>0.3428</td>
<td>0.3431</td>
<td>0.4488</td>
<td>1.5832</td>
<td>2.6700</td>
<td>0.3092</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.4320</td>
<td>0.3515</td>
<td>0.3520</td>
<td>0.4321</td>
<td>1.5676</td>
<td>3.6300</td>
<td>0.2293</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gamma</td>
<td>2</td>
<td>MADET</td>
<td>0.4791</td>
<td>0.4027</td>
<td>0.3525</td>
<td>0.3843</td>
<td>1.6186</td>
<td>5.5800</td>
<td>0.3591</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6270</td>
<td>0.2605</td>
<td>0.2674</td>
<td>0.5593</td>
<td>1.7143</td>
<td>0.0000</td>
<td>1.4069</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.4986</td>
<td>0.3594</td>
<td>0.3560</td>
<td>0.4563</td>
<td>1.6703</td>
<td>2.5700</td>
<td>0.4066</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.4828</td>
<td>0.3673</td>
<td>0.3644</td>
<td>0.4432</td>
<td>1.6577</td>
<td>3.3000</td>
<td>0.3249</td>
<td></td>
</tr>
<tr>
<td>$n = m = k = l = 100$</td>
<td>Norm</td>
<td>1</td>
<td>MADET</td>
<td>0.3801</td>
<td>0.3815</td>
<td>0.3786</td>
<td>0.3745</td>
<td>1.5147</td>
<td>5.9200</td>
<td>0.0187</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.5770</td>
<td>0.2310</td>
<td>0.2321</td>
<td>0.5699</td>
<td>1.6099</td>
<td>0.0000</td>
<td>1.4978</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.4467</td>
<td>0.3387</td>
<td>0.3387</td>
<td>0.4470</td>
<td>1.5710</td>
<td>2.4200</td>
<td>0.0613</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.5308</td>
<td>0.5056</td>
<td>0.5040</td>
<td>0.5124</td>
<td>2.0892</td>
<td>2.0000</td>
<td>0.0688</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gamma</td>
<td>2</td>
<td>MADET</td>
<td>0.4683</td>
<td>0.4043</td>
<td>0.3506</td>
<td>0.3944</td>
<td>1.6177</td>
<td>4.7800</td>
<td>0.3357</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6217</td>
<td>0.2516</td>
<td>0.2448</td>
<td>0.5806</td>
<td>1.6988</td>
<td>0.0000</td>
<td>1.5396</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.4921</td>
<td>0.3566</td>
<td>0.3531</td>
<td>0.4586</td>
<td>1.6604</td>
<td>2.2600</td>
<td>0.3937</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.4769</td>
<td>0.3658</td>
<td>0.3635</td>
<td>0.4447</td>
<td>1.6508</td>
<td>2.8300</td>
<td>0.3120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exp</td>
<td>3</td>
<td>MADET</td>
<td>0.6273</td>
<td>0.3772</td>
<td>0.5847</td>
<td>0.4763</td>
<td>2.0655</td>
<td>2.1700</td>
<td>0.0663</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GYI</td>
<td>0.6598</td>
<td>0.3923</td>
<td>0.5635</td>
<td>0.4958</td>
<td>2.1114</td>
<td>0.0000</td>
<td>0.6819</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MD</td>
<td>0.5237</td>
<td>0.5032</td>
<td>0.5387</td>
<td>0.5140</td>
<td>2.0797</td>
<td>1.5000</td>
<td>0.0705</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MV</td>
<td>0.5251</td>
<td>0.5020</td>
<td>0.5400</td>
<td>0.5131</td>
<td>2.0802</td>
<td>1.4800</td>
<td>0.0757</td>
<td></td>
</tr>
</tbody>
</table>
7 Discussion

For diseases with multiple stages, there are a few measures of diagnostic accuracy and cut-point selection criteria available. In this paper, we propose a new measure of diagnostic accuracy for the \( k \) stage setting. The new MADET method directly utilizes all the classification information and has an appealing geometric justification of its use as a diagnostic measure for disease with \( k \) stages \((k > 2)\).
Additionally, the new measure can be utilized as a criterion to select optimal cut-points. Simulation results show that MADET can detect the differences in diagnostic accuracy that GYI and HUM fail to observe. Furthermore, simulation results about cut-point selection show that the proposed MADET method has performance comparable to that of GYI, and is capable of achieving better balanced rates while maintaining minimal loss of TCCR.

The MADET will be zero when any two (or more) classes are indistinguishable regardless of whether or not there is separation in other classes. When two classes or more overlap, there is an indication that class-merging should occur before MADET is calculated. We suggest the users of MADET to check the SPM in equation (4) when MADET = 0. When two or more rows of SPM overlap, a good practice is to merge these classes so that only one row in SPM is kept. Subsequently, MADET can be calculated based on the reduced SPM.

In this paper, we focus on the scenarios for diseases with ordinal stages. In many applications, e.g. genomic studies, we need to deal with \( k \) nominal classes. For such cases, there actually exist \( k! \) possible classification decision rules; i.e. one for each possible permutation. The proposed measure MADET can be easily adapted to such cases due to the fact that switching columns (or rows) of SPM does not change the absolute value of the determinant of SPM (MADET), via two steps: (1) assume a hypothetical order for the \( k \) classes and calculate the corresponding MADET and the resulting SPM; (2) Permute the columns of SPM to maximize the CCRs or to minimize the false classification ones or both, depending on the interest. Consequently, the final optimal SPM as well as the corresponding optimal classification decision rule can be obtained.

It is worth noting that Skaltsa et al.\(^4^4\) and Batterton and Schubert\(^4^5\) have proposed two interesting metrics based on different cost functions in the context of incorporating the classification cost to evaluate a biomarker for the multiple-class diagnostic problems. For future work, we are also interested in incorporating classification costs to MADET.

An R-program is available on request from the corresponding author (ltian@buffalo.edu).

**Acknowledgments**

Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimers Association; Alzheimers Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics,
N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Mesoscale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. The ADNI research was also supported by NIH grants P30 AG010129 and K01 AG030514.

References