

A min–max combination of biomarkers to improve diagnostic accuracy

Chunling Liu,^{a,b} Aiyi Liu^{a*†} and Susan Halabi^c

Diagnostic accuracy can be improved considerably by combining multiple biomarkers. Although the likelihood ratio provides optimal solution to combination of biomarkers, the method is sensitive to distributional assumptions which are often difficult to justify. Alternatively simple linear combinations can be considered whose empirical solution may encounter intensive computation when the number of biomarkers is relatively large. Moreover, the optimal linear combinations derived under multivariate normality may suffer substantial loss of efficiency if the distributions are apart from normality. In this paper, we propose a new approach that linearly combines the minimum and maximum values of the biomarkers. Such combination only involves searching for a single combination coefficient that maximizes the area under the receiver operating characteristic (ROC) curves and is thus computation-effective. Simulation results show that the min–max combination may yield larger partial or full area under the ROC curves and is more robust against distributional assumptions. The methods are illustrated using the growth-related hormones data from the Growth and Maturation in Children with Autism or Autistic Spectrum Disorder Study (Autism/ASD Study). Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: area under curves; linear combinations; receiver operating characteristic (ROC) curve; robustness; sensitivity; specificity

1. Introduction

Powerful biomarkers have become and will continue to be important tools in diagnosis, clinical delivery, and research. In the area of diagnostic medicine, a biomarker is often used as a diagnostic tool to identify subjects with a disease, or at high risk of developing the disease. For example, prostate-specific antigen is used for diagnosis of prostate cancer, and high cholesterol levels are used as an indicator of coronary and vascular disease. However, even the most well known and commonly used biomarkers are not 100 per cent accurate and may incur considerable classification errors which will have serious implications on patients. It is therefore of great interest to utilize multiple biomarkers to increase accuracy so that these diagnostic tests will become part of standard of care.

For a biomarker whose values are measured on a continuous scale, its diagnostic performance in identifying diseased subjects is usually assessed via receiver operating characteristic (ROC) curves. Assuming that subjects with higher levels of the biomarker are more likely to develop the disease, a subject is classified as diseased (non-diseased) if the biomarker level is above (below) a certain threshold. Corresponding to this threshold, the biomarker's sensitivity is the probability that a diseased subject is correctly identified, and its specificity is the probability that a non-diseased subject is correctly classified. The ROC curve of the biomarker is then a plot of its sensitivity versus one minus its specificity for all possible values of the threshold. The area and partial area under the ROC curve are popular measures of the performance of a biomarker, with larger value of the area being the better [1]. There

^aBiostatistics and Bioinformatics Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Rockville, MD 20852, U.S.A.

^bDepartment of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

^cDepartment of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC 27708, U.S.A.

*Correspondence to: Aiyi Liu, Biostatistics and Bioinformatics Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Rockville, MD 20852, U.S.A.

†E-mail: liua@mail.nih.gov

are many reasons for the popularity of the ROC curve. It presents the benefits versus costs of screening with the biomarker. It also provides a scale where biomarkers can be meaningfully compared while original scales of measurements are often not comparable.

A number of authors have considered the problem of combining biomarkers to improve diagnostic accuracy by maximizing the area under the ROC curves. When the levels of the biomarkers in the diseased and non-diseased population both follow multivariate normal distributions, Su and Liu [2] derived the optimal linear combinations that yield largest area under the ROC curves. To date, optimal linear combinations have remained yet to be found if biomarkers follow non-normal distributions. Although Su and Liu's [2] optimal linear combinations can be used when data are not normally distributed, they may suffer from loss of efficiency. Pepe and Thompson [3] proposed to obtain an empirical solution of the optimal linear combinations that maximize the Mann–Whitney statistic, an empirical estimate of the ROC area. This procedure is distribution-free and thus robust against distributional assumptions. However, it may become less efficient than Su and Liu's [2] linear combination if data are normally distributed. Moreover, when the number of biomarkers is relatively larger, this empirical optimization process can be quite computationally intensive.

Green and Swets [4] and Pepe *et al.* [5] noticed that if the disease status and the biomarkers' levels satisfy a generalized linear model, then the regression coefficients are the optimal linear combinations that maximize the area under the ROC curves. Recognizing the optimality of the fundamental Neyman–Pearson lemma [6] in hypothesis testing, Baker [7, 8], Eguchi and Copas [9], and Pepe *et al.* [5] pointed out that the likelihood ratio provides the globally optimal combination, that is, at any given specificity, the likelihood ratio possesses the largest sensitivity among all combinations. Thus, the likelihood ratio also yields the largest ROC area. Like Su and Liu's [2] combinations, these methods are optimal only if the (conditional) joint distribution function of the biomarkers given the disease status is correctly specified. Their performance can be substantially adversely affected if the true joint distribution functions depart considerably from the specified ones under which the combination is derived.

In this paper, we consider linear combinations of the minimum and maximum biomarkers' levels, hereafter referred to as 'min–max' combinations. We show in the following section that the maximum (minimum) of the biomarkers' levels yields larger sensitivity (specificity) than any individual biomarker. Therefore, a compromise between the maximum and the minimum, such as a proper linear combination of the two, is expected to perform well in certain situations. We derive the empirical optimal linear combination that maximizes the Mann–Whitney statistic, that is, the empirical area under the ROC curve. The proposed min–max procedure is nonparametric in nature and thus is robust against misspecified distributions. Furthermore, compared to the existing empirical linear combination approach, the min–max procedure involves the optimization of a single parameter, and thus numerically is more tractable. Diagnostic accuracy of the optimal min–max combination is investigated via simulation and real-data examples, in comparison with other combination methods. For the settings considered, numerical results demonstrate that the min–max combination outperforms the others, for at least a plausible range of specificity.

2. Combination of biomarkers

Consider k biomarkers whose levels are denoted by a vector $\mathbf{M}=(M_1, \dots, M_k)$. Conditioning on the disease status D , binary with 1 being diseased and 0 otherwise, the biomarkers's levels are denoted by $\mathbf{X}=(X_1, \dots, X_k)$ for a non-diseased subject ($D=0$) and by $\mathbf{Y}=(Y_1, \dots, Y_k)$ for a diseased subject ($D=1$), with density function $f=f(x_1, \dots, x_k)$ and $g=g(y_1, \dots, y_k)$, respectively. The corresponding marginal density function is denoted by f_i and g_i for the i th biomarker. The corresponding distribution functions are denoted by F , G , F_i , and G_i , respectively. According to Bamber [10], the area under the ROC curve of the i th biomarker is given by

$$A_i = \Pr\{Y_i - X_i > 0\}. \quad (1)$$

By combination of the k biomarkers we mean a real-valued function of the marker levels, $h=h(\mathbf{M})$, mapping from the k -dimensional Euclidean space \mathcal{R}^k onto the real line \mathcal{R}^1 . The area under the ROC curve of this combination is

$$A = \Pr\{h(\mathbf{X}) > h(\mathbf{Y})\}.$$

When the distributions F and G can be correctly specified, the likelihood ratio function

$$\text{LR}(\mathbf{M}) = \frac{g(\mathbf{M})}{f(\mathbf{M})} \quad (2)$$

yields an ROC curve that is above the ROC curves of any other combinations. Therefore, $\text{LR}(\mathbf{M})$ also has the largest area among all combinations.

In addition to its dependence on correct specification of distributions, the likelihood ratio combination can be quite complicated and hard to interpret to the medical investigators. A simple alternative is to use a linear combination

$$l(\mathbf{M}) = \lambda_1 M_1 + \dots + \lambda_k M_k \quad (3)$$

that may have a more meaningful practical interpretation. However, except for normal distributions, it is not clear whether an optimal solution for $\lambda = (\lambda_1, \dots, \lambda_k)$ exists, and even if it does, whether the solution can be found, although one can always search for optimal empirical solution using the observed biomarkers' levels.

Note that the linear combination λ yields an ROC curve with area

$$A(\lambda) = \Pr\{\lambda_1(Y_1 - X_1) + \dots + \lambda_k(Y_k - X_k) > 0\}. \quad (4)$$

Assuming that the biomarkers' values follow multivariate normal distributions, with $\mathbf{X} \sim N(\mu_X, \Sigma_X)$ for non-diseased subjects and with $\mathbf{Y} \sim N(\mu_Y, \Sigma_Y)$ for diseased subjects, Su and Liu [2] derived the best linear combination (that maximizes the area under the ROC curves) with λ_i equal to, apart from a constant scalar, the i th component of

$$(\Sigma_X + \Sigma_Y)^{-1}(\mu_Y - \mu_X).$$

If the two normal distributions have a common variance-covariance matrix, that is, $\Sigma_X = \Sigma_Y$, then Su and Liu's [2] linear combination is identical to the likelihood ratio combination.

We now consider the min-max combination of the biomarkers. Let

$$M_{\max} = \max_{1 \leq i \leq k} M_i, \quad M_{\min} = \min_{1 \leq i \leq k} M_i \quad (5)$$

be, respectively, the maximum and minimum value of the k biomarkers. Then the 'new' marker M_{\max} has larger sensitivity and smaller specificity for any given threshold than any individual biomarker, and the 'new' marker M_{\min} has smaller sensitivity and larger specificity for any given threshold than any individual biomarker. To see this, let X_{\max} , X_{\min} , Y_{\max} , and Y_{\min} be similarly defined for each disease status. Then for sensitivity at a cut-point c we have

$$\Pr\{Y_{\max} > c\} = 1 - \Pr\{Y_{\max} \leq c\} = 1 - \Pr\{Y_j \leq c, 1 \leq j \leq k\} \geq 1 - \Pr\{Y_i \leq c\} = \Pr\{Y_i > c\}$$

and

$$\Pr\{Y_{\min} > c\} = \Pr\{Y_j > c, 1 \leq j \leq k\} \leq \Pr\{Y_i > c\}$$

for all $1 \leq i \leq k$.

For specificity we have

$$\Pr\{X_{\max} \leq c\} = \Pr\{X_j \leq c, 1 \leq j \leq k\} \leq \Pr\{X_i \leq c\}$$

and

$$\Pr\{X_{\min} \leq c\} = 1 - \Pr\{X_{\min} > c\} = 1 - \Pr\{X_j > c, 1 \leq j \leq k\} \geq 1 - \Pr\{X_i > c\} = \Pr\{X_i \leq c\}$$

for all $1 \leq i \leq k$.

These results indicate that M_{\max} (M_{\min}) is a plausible combination if sensitivity (specificity) is the primary concern, which, of course, is rarely the case. If our interest is to maximize the area or partial area under the ROC curves or to improve sensitivity over a range of specificity, then a compromise between M_{\max} and M_{\min} is expected to perform well, especially when distributional assumptions are difficult to justify. In this paper we confine our attention to linear combinations of M_{\max} and M_{\min} , that is, combinations of the form $M_\lambda = M_{\max} + \lambda M_{\min}$. The area under the ROC curve is given by

$$A(\lambda) = \Pr\{(Y_{\max} - X_{\max}) + \lambda(Y_{\min} - X_{\min}) > 0\}. \quad (6)$$

The optimal min–max combination is given by λ_{opt} that maximizes the area, that is

$$A(\lambda_{\text{opt}}) = \max_{\lambda} \Pr\{(Y_{\text{max}} - X_{\text{max}}) + \lambda(Y_{\text{min}} - X_{\text{min}}) > 0\}. \quad (7)$$

3. Estimation of optimal combinations

Optimal combinations of biomarkers usually depend on the distribution functions which are fully unspecified or involve several unknown parameters. A case–control study design can be employed to estimate these combinations. Let

$$\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{ki}), \quad i = 1, 2, \dots, n$$

be the observed values of the biomarkers from a random sample of size n in the non-diseased population, and

$$\mathbf{Y}_j = (Y_{1j}, Y_{2j}, \dots, Y_{kj}), \quad j = 1, 2, \dots, m$$

be the observed values of the biomarkers from a random sample of size m in the diseased population. If F and G are assumed to have known function forms apart from a few unknown parameters, then these parameters can be estimated by the corresponding maximum likelihood estimates obtained from the data. Replacing these parameters by their maximum likelihood estimates in the likelihood ratio function (2) then yields an estimate of the likelihood ratio combination.

If F and G are assumed to be of multivariate normal, then the mean vectors μ_X and μ_Y and the variance–covariance matrices Σ_X and Σ_Y can be estimated by the corresponding sample mean vectors and sample variance–covariance matrices. These estimates then yield an estimate of the likelihood ratio combination and Su and Liu's [2] optimal linear combination.

Without assuming any parametric forms for the distributions, the optimal linear combination can be estimated by maximizing the Mann–Whitney statistic

$$W(\lambda) = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m I(\lambda_1 Y_{1j} + \dots + \lambda_k Y_{kj} > \lambda_1 X_{1i} + \dots + \lambda_k X_{ki}) \quad (8)$$

the empirical estimate of the area under the ROC curve of the combination λ . Here, $I(\cdot)$ denotes the indicator function.

Similarly, the empirical optimal min–max combination can be derived by maximizing the corresponding Mann–Whitney statistic:

$$W(\lambda) = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m I(Y_{j,\text{max}} + \lambda Y_{j,\text{min}} > X_{i,\text{max}} + \lambda X_{i,\text{min}}), \quad (9)$$

where for the i th non-diseased subject

$$X_{i,\text{max}} = \max_{1 \leq l \leq k} X_{li}, \quad X_{i,\text{min}} = \min_{1 \leq l \leq k} X_{li}$$

and for the j th diseased subject

$$Y_{j,\text{max}} = \max_{1 \leq l \leq k} Y_{lj}, \quad Y_{j,\text{min}} = \min_{1 \leq l \leq k} Y_{lj}.$$

The values of λ s can be found numerically using the built-in function *optim* in the R software (free download at <http://www.r-project.org/>). The function provides general-purpose optimization based on Nelder–Mead, quasi-Newton and conjugate-gradient algorithms, and is widely used to solve various optimization problems such as finding the maximum likelihood estimates.

4. Simulations

It appears to be difficult to analytically compare the performance of the aforementioned combination methods because such a comparison involves the joint distribution of the maximum and minimum of

correlated random variables, which is quite complicated even when the variables are jointly normally distributed. Instead we rely on simulations to make comparison among the ROC curves resulted from the combinations under investigation.

We consider four settings of the joint distributions of five biomarkers ($k=5$) and estimated the ROC curves and their areas of the combinations for each setting. For each setting, 50 000 observations were generated from the underlying distributions that were used to estimate the ROC curves and their areas. We compare the min–max combination with Su and Liu’s [2] normal linear combination and the empirical linear combination. The combination parameter λ for min–max and empirical linear combinations were derived by maximizing the corresponding Mann–Whitney statistic. The normal linear combination was obtained from the sample mean vectors and sample variance–covariance matrices (In the figures below, TPF denotes Sensitivity, and FPF denotes 1-specificity).

4.1. Multivariate normal distributions with equal variance–covariance

We first considered two multivariate normal distributions with different mean vectors and equal variance–covariance matrices given by $\mu_X=0$ and

$$\mu_Y = \begin{pmatrix} 0.5 \\ 0.6 \\ 0.7 \\ 0.8 \\ 1.0 \end{pmatrix}, \quad \Sigma_X = \Sigma_Y = \begin{pmatrix} 1.0 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1.0 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1.0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1.0 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 1.0 \end{pmatrix}.$$

For this setting Su and Liu’s [2] linear combination is identical to the likelihood ratio combination and thus is optimal in that it yields an ROC curve which is higher than any other ROC curves. This is reflected in Figure 1, which presents the three ROC curves. In this case the min–max combination is inferior to both linear combinations which are identical. The area under the ROC curve is found to be 0.859 for min–max combination, and 0.938 for the two linear combinations.

4.2. Multivariate normal distributions with unequal variance–covariance

We next consider two multivariate normal distributions with different mean vectors and unequal variance–covariance matrices. The means and variance–covariances are the same as in Section 4.1

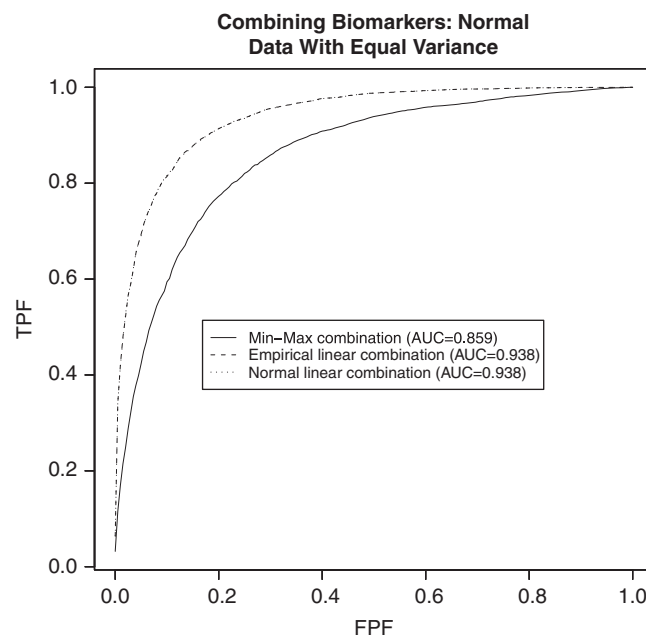


Figure 1. Comparison of min–max combination to linear combinations under multivariate normality with equal variances.

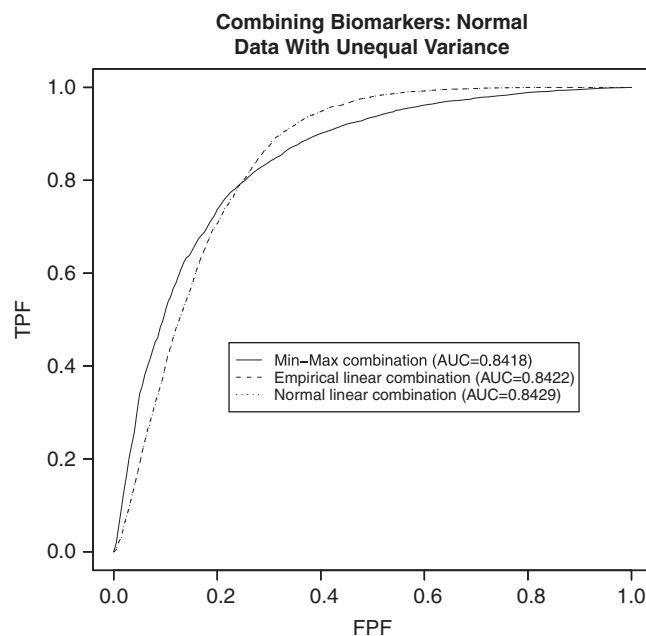


Figure 2. Comparison of min–max combination to linear combinations under multivariate normality with unequal variances.

except that Σ_X is set to be a diagonal matrix with

$$\Sigma_X = \text{diag}(1.5, 1, 2.5, 2, 3).$$

Under this setting, the likelihood ratio yields a combination that is a quadratic function of the biomarkers' values. Su and Liu's [2] linear combination is thus no longer the best combination, though it is among combinations with linear forms. The three ROC curves are presented in Figure 2. The areas under the three ROC curves are quite close, indicating almost identical performance when judged by the area under the ROC curve. However, a more detailed investigation reveals that the min–max yields larger sensitivity for specificity larger than 75 per cent. This indicates that if relatively high specificity is desirable, then the min–max combination may turn out to be more appealing than linear combinations, even when levels of the biomarkers follow multivariate normal distributions with unequal variance–covariance matrices. Note that under multivariate normality assumptions, the optimal empirical linear combination is, in theory, identical to Su and Liu's [2] optimal linear combination, as reflected in Figures 1 and 2.

4.3. Multivariate log-normal distributions with unequal variance–covariance

Linear combinations usually perform better when the distributions of the levels of the biomarkers are close to normal, and the variance–covariance matrices do not differ much. In this section, we investigate the diagnostic accuracy of the min–max combination as compared with the linear combinations, assuming that the biomarkers' levels follow multivariate log-normal distributions, that is, the log-transformed levels are jointly normally distributed. To this end, data from multivariate log-normal distributions are generated by exponentiating the simulated data in Section 4.2 from the multivariate normal distributions.

The estimated ROC curves are presented in Figure 3. The min–max combination appears to perform considerably better than the two linear combinations; it yields a relatively larger area of 0.847 under the ROC curve and the ROC curve is higher than the other two over a quite large range of specificity (from about 55 to 100 per cent). The empirical linear combination appears to be uniformly better than Su and Liu's [2] linear combination in that it produces a higher ROC curve. This should not be surprising because Su and Liu's [2] linear combination is optimal only under multivariate normality, and may suffer substantial loss in diagnostic accuracy if the multivariate normality assumption is violated.

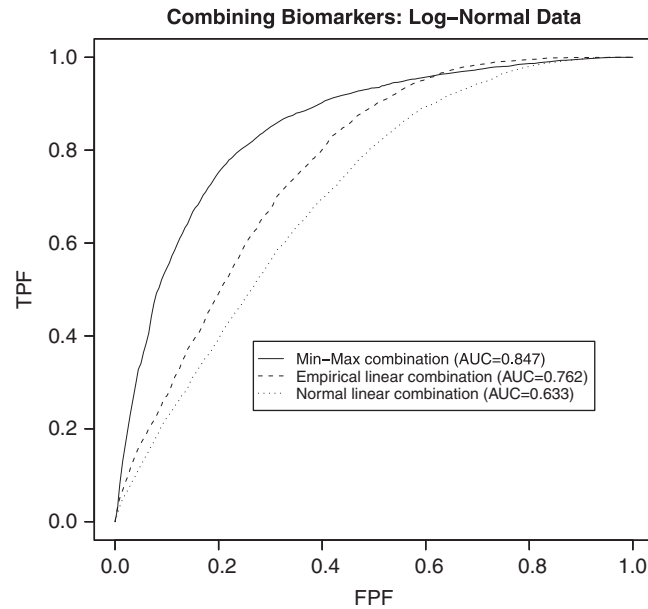


Figure 3. Comparison of min-max combination to linear combinations under multivariate log-normality.

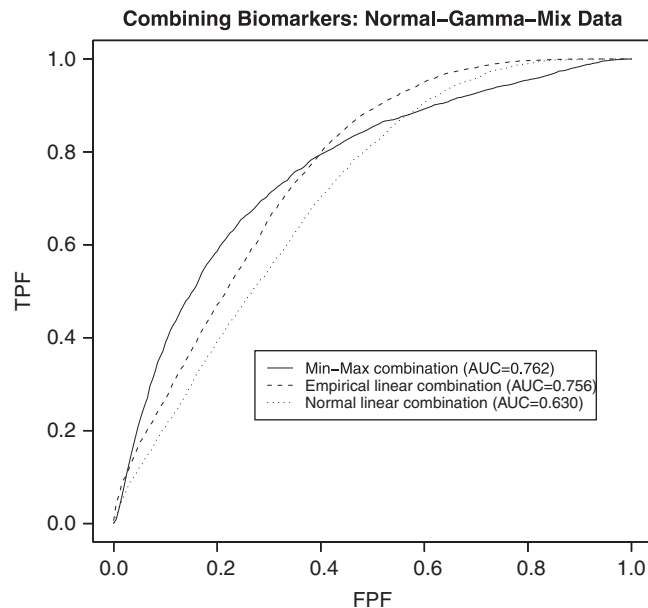


Figure 4. Comparison of min-max combination to linear combinations with multivariate normal and gamma mixture.

4.4. Normal-gamma mixture

We further investigate the three combinations assuming that levels of biomarkers follow a mixed distribution of normal and gamma with equal allocation. The simulated data were obtained by adding simulated gamma data to the multivariate normal data in Section 4.2. To be specific, the biomarkers' levels \mathbf{Y} from the diseased population were generated by adding a gamma variate with a common shape parameter of 0.1 to the multivariate normal data. For biomarkers' levels \mathbf{X} in the non-diseased population, a gamma variate with shape parameter $0.1 \times i$ was added to the i th biomarker ($i = 1, 2, 3, 4, 5$).

Figure 4 with the simulated ROC curves reveals trends similar to Figure 3. The min-max combination appears to be the best among the three which yields the largest area under the ROC curve. The min-max ROC curve again is higher than the other two over a quite large range of specificity (from about

75 to 100 per cent). Su and Liu's [2] linear combination again is the worst among the three, and appears to be uniformly inferior to the empirical linear combination, reflecting the substantial loss in accuracy with non-normal data.

5. An example: diagnosis of autism by combining growth-related hormones

In this section we illustrate the application of the proposed min–max combination to the diagnosis of autism, a disorder of neural development characterized by impaired social interaction and communication, and by restricted and repetitive behavior.

To investigate whether children aged 4–8 years old with autism/autism spectrum disorders (ASD) are larger than matched control children and the role that growth-related hormones may play, a case–control study was conducted by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development in 2002–2005. The study participants consisted of 81 subjects (75 boys and 6 girls) diagnosed with autism/ASD and 80 age-matched controls (59 boys and 21 girls). Blood samples from the study participants were assayed for insulin-like growth factors (IGF1, IGF2), insulin-like growth factor-binding protein (IGFBP3), and growth hormone-binding protein (GHBP), as well as for dehydroepiandrosterone (DHEA) and DHEA sulfate (DHEAS). See Mills *et al.* [11] for details of subject enrollment and data collection.

To serve the purpose we selected three growth hormones, IGFBP3, IGF1, and GHBP, to illustrate the performance of the min–max combination in identifying autism/ASD disease. Owing to limited number of enrollments, data from the girls are excluded from the analysis. Levels of the hormones are standardized to avoid different measurement units. To ensure the optimal linear combinations to have a unique solution, the combination coefficient corresponding to IGFBP3 is set to 1. The optimal min–max combination is found to be

$$\max\{\text{IGFBP3}, \text{IGF1}, \text{GHBP}\} + 0.563 \times \min\{\text{IGFBP3}, \text{IGF1}, \text{GHBP}\}.$$

The empirical linear combination is found to be

$$\text{IGFBP3} - 0.05 \times \text{IGF1} + 0.3 \times \text{GHBP}$$

close to the normal linear combination given by

$$\text{IGFBP3} - 0.039 \times \text{IGF1} + 0.433 \times \text{GHBP}.$$

The corresponding area under the ROC curve is found to be 0.759, 0.753, and 0.754, respectively. Thus, judging by the area under the ROC curve, the min–max combination has higher diagnostic accuracy than linear combinations in identifying autism disease. Furthermore, the min–max combination yields higher specificity for sensitivity larger than 70 per cent, a more appealing optimality property (Figure 5).

6. Discussion

Diagnostic accuracy, as measured by the area or partial area under the ROC curves, can be substantially improved by combining multiple biomarkers. The optimal combinations such as the likelihood ratio combination or Su and Liu's [2] combination, however, are vulnerable to the underlying distributional assumptions, which, when violated, can yield undesirable results and will have serious implications on patients. The min–max combination proposed in this paper, on the other hand, is nonparametric in nature. Thus, this approach is more robust against parametric models. This is demonstrated from the simulation results and the autism example. Our numerical results also showed that the min–max combination may achieve higher sensitivity over a desirable range of specificity. Similar results (data not shown) were observed from simulation on multivariate distributions that can be transformed to a multivariate normal distribution via a Box–Cox transformation, and on Ramabhadran's [12] multivariate gamma distribution.

The min–max procedure aims at combining repeated measurements of a single biomarker over time or multiple biomarkers that are measured with the same unit. If the biomarkers are measured with different units, then the measurements need to be converted to have the same unit or to be standardized to be unit-less, as what was done in the autism/ASD example, before combining.

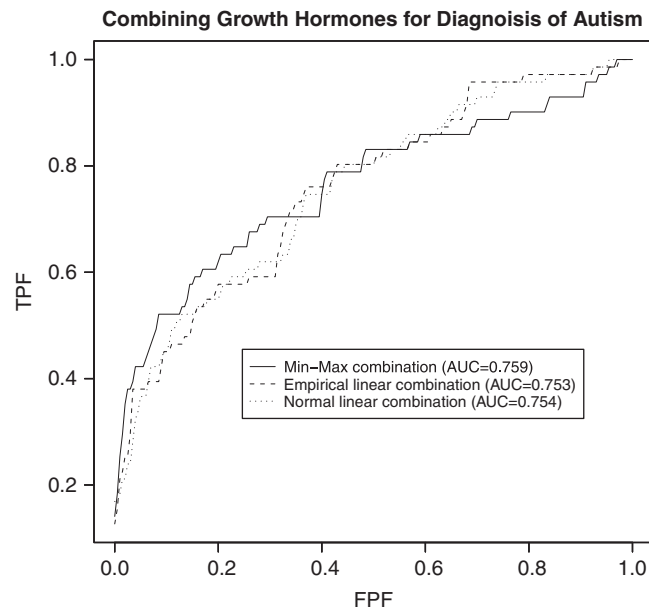


Figure 5. Min–max combination of growth hormones IGF2, IGFBP3, IGF1, and GHBP for diagnosis of autism.

If, instead of sensitivity, the specificity of a biomarker is of primary concern, then the min–max combination adopts a form of specificity + $\lambda \times$ sensitivity. This combination is expected to be able to achieve higher specificity over a desirable range of sensitivity.

The estimated ROC curves of the combinations can be quite jagged when the sample sizes are relatively small. Smoothed ROC curves can be obtained using methods developed by Zou *et al.* [13] and Zou and Hall [14], among others.

Future research is necessary for inference on the optimal min–max combination coefficient λ_{opt} , such as consistent estimation and confidence intervals. The performance of the min–max combination under other multivariate distributions is also of particular interest.

Acknowledgements

Research of the first two authors was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the National Institutes of Health (NIH). The authors thank Dr James L. Mills and Dr Kai F. Yu for helpful discussions on the autism data example. The authors thank the Editor, Associate Editor, and two referees for their constructive comments, and Dr Paul Albert for helpful discussions.

References

- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**:29–36.
- Su JQ, Liu JS. Linear combinations of multiple diagnostic markers. *Journal of the American Statistical Association* 1993; **88**:1350–1355.
- Pepe MS, Thompson ML. Combining diagnostic test results to increase accuracy. *Biostatistics* 2000; **1**:123–140.
- Green DM, Swets JA. *Signal Detection Theory and Psychophysics*. Wiley: New York, 1966.
- Pepe MS, Cai T, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* 2006; **62**:221–229.
- Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypothesis. *Philosophical Transactions of the Royal Society of London, Series A* 1933; **231**:289–337.
- Baker S. Identifying combinations of cancer markers for further study as triggers of early intervention. *Biometrics* 2000; **56**:1082–1087.
- Baker S. The central role of receiver operating characteristic (ROC) curves in evaluating tests for the early detection of cancer. *Journal of the National Cancer Institute* 2003; **95**:511–515.
- Eguchi S, Copas J. A class of logistic-type discriminant functions. *Biometrika* 2002; **89**:1–22.
- Bamber D. The area above the ordinal dominance graph and the area below the receiver operating characteristics graph. *Journal of Mathematical Psychology* 1975; **12**:387–415.

11. Mills JL, Hediger ML, Molloy CA, Chrousos GP, Manning-Courtney P, Yu KF, Brasington M, England LJ. Elevated levels of growth-related hormones in autism and autism spectrum disorder. *Clinical Endocrinology* 2007; **67**: 230–237.
12. Ramabhadran VR. A multivariate gamma-type distribution. *Sankhya* 1951; **11**:45–46.
13. Zou KH, Hall WJ, Shapiro DE. Smooth nonparametric receiver operating characteristic curves for continuous diagnostic tests. *Statistics in Medicine* 1997; **16**:2143–2156.
14. Zou KH, Hall WJ. Two transformation models for estimating an ROC curve derived from continuous data. *Journal of Applied Statistics* 2000; **26**:621–631.