EFFECT OF CORRELATION BETWEEN CLINICAL TESTS ON THE PERFORMANCE OF A MULTIPLE TEST-BASED DIAGNOSTIC SYSTEM

Study with a Logistic Model and Neural Nets

Noriaki Ikeda[1], Kai Ishida[2], Harukazu Tsuruta[1] and Akihiro Takeuchi[1]
[1]*Medical Informatics, School of AHS, Kitasato University, Sagamihara, Kanagawa Japan*
[2]*Graduate School of Medical Sciences, Kitasato University, Sagamihara, Kanagawa Japan*

Keywords: Multiple tests, Diagnostic performance, Correlation between tests, Logistic model, Neural nets.

Abstract: To examine the improvement of diagnostic performance by combining multiple tests, an algorithm was developed for generation of simulated data with arbitrary sensitivity, specificity and inter-test correlations. The effects of the number of tests and inter-test correlations on the diagnostic performance were studied using a logistic model and neural network (NN) models. The diagnostic performance measured by the concordance index, c, increased as the number of tests increased. For the same number of tests, the diagnostic performance was lowered by positive correlation and was elevated by negative correlation. Improvement of the performance was not obtained by increasing the number of NN layers.

1 INTRODUCTION

It is a common practice in clinical medicine to develop a better (more reliable) diagnostic system using multiple tests that individually are less reliable (Ikeda et al., 2006; Ikeda et al., 2007). For example, Hara et al. reported that a better diagnostic index for prediction of improvement of left ventricular ejection fraction (LVEF) after cardiac resynchronization therapy (CRT) in patients with heart failure could be obtained using a combination of three indices of cardiac function, such as Radial, OWD and IVMD (Hara, 2008).

A logistic model is often used for combining multiple tests, each of which has a sensitivity and specificity. The factor with a greater sensitivity and specificity has a larger regression coefficient. A neural network (NN) model may be effectively used for a case with strong nonlinearity.

If the tests are mutually independent the diagnostic performance is expected to increase as the number of combined tests becomes large. The first problem is to determine the relationship between the diagnostic performance and the number of tests. However, there are often correlations among tests. Improvement in diagnosis is clearly not possible if these correlations are strongly positive, whereas the effect of a negative correlation is less clear. Therefore, the second problem is to determine the effect of inter-test correlations on the diagnostic performance.

The purpose of the present study was to develop an algorithm that calculates the probability of the outcome of combined tests when the sensitivity and specificity of each test and the inter-test correlations are given, and to study the two problems described above based on simulated data generated by the algorithm.

In this study, we only deal with binary tests with outcomes that are positive (1) or negative (0).

2 METHODS

2.1 Joint Probability of Two Tests

The relationship between disease D and a clinical test T_i can be presented as a contingency table (Table 1), in which D reflects the status of the patient ($D = 1$ indicates having the disease and $D = 0$ indicates not having the disease) and T_i indicates the result of the i-th test (positive T_i=1, negative $T_i = 0$).

The sensitivity and specificity of the test are represented by α_i and β_i, respectively. For $D = 1$, the correlation coefficient between test T_i and test T_j is

Ikeda N., Ishida K., Tsuruta H. and Takeuchi A..
EFFECT OF CORRELATION BETWEEN CLINICAL TESTS ON THE PERFORMANCE OF A MULTIPLE TEST-BASED DIAGNOSTIC SYSTEM - Study with a Logistic Model and Neural Nets.
DOI: 10.5220/0003655703260329
In *Proceedings of the International Conference on Neural Computation Theory and Applications* (NCTA-2011), pages 326-329
ISBN: 978-989-8425-84-3

Table 1: Contingency table of test Ti and diagnosis D.

	$D=1$	$D=0$
$T_i=1$	α_i	$1-\beta_i$
$T_i=0$	$1-\alpha_i$	β_i

r_{ij}^+; and for $D=0$, the correlation coefficient is r_{ij}^-.

For a case with $D=1$, the joint probability $p_{km}^+,(k,m=0,1)$ of T_i and T_j are given by:

$$p_{11}^+ = \Pr(T_i=1,T_j=1) = \alpha_i\alpha_j + \phi_{ij}^+ \quad (1)$$

$$p_{10}^+ = \Pr(T_i=1,T_j=0) = \alpha_i(1-\alpha_j) - \phi_{ij}^+ \quad (2)$$

$$p_{01}^+ = \Pr(T_i=0,T_j=1) = (1-\alpha_i)\alpha_j - \phi_{ij}^+ \quad (3)$$

$$p_{00}^+ = \Pr(T_i=0,T_j=0) = (1-\alpha_i)(1-\alpha_j) + \phi_{ij}^+, \quad (4)$$

where

$$\phi_{ij}^+ = r_{ij}^+\sqrt{\alpha_i\alpha_j(1-\alpha_i)(1-\alpha_j)}. \quad (5)$$

For a case with $D=0$, similar relationships are obtained by replacing α_i by $1-\beta_i$ Cα_j by $1-\beta_j$Cr_{ij}^+ by r_{ij}^-Cand ϕ_{ij}^+ by ϕ_{ij}^-,

$$p_{11}^- = \Pr(T_i=1,T_j=1) = (1-\beta_i)(1-\beta_j) + \phi_{ij}^- \quad (6)$$

$$p_{10}^- = \Pr(T_i=1,T_j=0) = (1-\beta_i)\beta_j - \phi_{ij}^- \quad (7)$$

$$p_{01}^- = \Pr(T_i=0,T_j=1) = \beta_i(1-\beta_j) - \phi_{ij}^- \quad (8)$$

$$p_{00}^- = \Pr(T_i=0,T_j=0) = \beta_i\beta_j + \phi_{ij}^- \quad (9)$$

where

$$\phi_{ij}^- = r_{ij}^-\sqrt{\beta_i\beta_j(1-\beta_i)(1-\beta_j)} \quad (10)$$

2.2 Data Generation Algorithm

A general theory of the distribution of n binary items has been established (Bahadur, 1961).

Let X denote the set of all points $x=(x_1,x_2,...,x_n)$ with each $x_i=0$ or 1. Let $p(x)$ be a given probability distribution on X, i.e.,

$$p(x)\geq 0, \qquad \sum_{x\in X} p(x) = 1. \quad (11)$$

For each $i=1,...,n$, let

$$\alpha_i = E_p(x_i), \qquad 0<\alpha_i<1, \quad i=1,...,n \quad (12)$$

where E_p denotes the expected value about p. If the variables $x_1,x_2,...,x_n$ are mutually independent, we have

$$p(x) = \prod_{i=1}^{n}\alpha_i^{x_i}(1-\alpha_i)^{1-x_i} \quad (13)$$

When there are correlation among the variables, Bahadur gave the following theorem (Bahadur, 1961). mTheoremn@For all $x=(x_1,x_2,...,x_n)$ on X

$$p(x) = \prod_{i=1}^{n}\alpha_i^{x_i}(1-\alpha_i)^{1-x_i} f(x) \quad (14)$$

with

$$f(x) = 1 + \sum_{i<j} r_{ij}y_iy_j + \sum_{i<j<k} r_{ijk}y_iy_jy_k + ... + r_{12\cdots n}y_1y_2\cdots y_n \quad (15)$$

$$y_i = (x_i-\alpha_i)/\sqrt{\alpha_i(1-\alpha_i)} \quad (16)$$

$$r_{ij} = E_p(y_iy_j) \quad (17)$$

$$r_{ijk} = E_p(y_iy_jy_k) \quad (18)$$

$$... \quad (19)$$

$$r_{12\cdots n} = E_p(y_1y_2\cdots y_n), \quad (20)$$

where r_{ij} is the second-order correlation, r_{ijk} is the third-order correlation, etc.

Similarly, if we set

$$1-\beta_i = \Pr(x_i=1|D=0) \quad (21)$$

$$z_i = (x_i-1+\beta_i)/\sqrt{\beta_i(1-\beta_i)}, \quad (22)$$

then the probability distribution $q(x)$ for $D=0$ is given by

$$q(x) = \prod_{i=1}^{n}\beta_i^{1-x_i}(1-\beta_i)^{x_i}g(x) \quad (23)$$

with

$$g(x) = 1 + \sum_{i<j} s_{ij}z_iz_j + \sum_{i<j<k} s_{ijk}z_iz_jz_k + ... + s_{12\cdots n}z_1z_2\cdots z_n \quad (24)$$

$$s_{ij} = E_p(z_iz_j) \quad (25)$$

$$s_{ijk} = E_p(z_iz_jz_k) \quad (26)$$

$$... \quad (27)$$

$$s_{12\cdots n} = E_p(z_1z_2\cdots z_n). \quad (28)$$

With this theory, all probabilities of combination of outcomes of tests with arbitrary sensitivity, specificity and correlations among the tests can be computed.

2.3 Example of Test Data

A data set of N tests with the following conditions was generated by the method described in section 2.2. The sensitivity and specificity of each test were both set to 0.6:

$$\alpha_i = \beta_i = 0.6, \qquad i=1,2,...,N. \quad (29)$$

We define R^+ as the correlation matrix among tests for the population with disease (D^+), and R^- as that for the population with no disease (D^-). Higher order correlations (>2) were set to zero in this study,

although they can easily be considered. An example data set with $N = 4$ and the correlation matrix

$$R^+ = R^- = \begin{pmatrix} 1.0 & -0.3 & 0.0 & 0.0 \\ -0.3 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{pmatrix} \quad (30)$$

is shown in Table 2. The frequency of each outcome of the tests was calculated according to $p(x)$ and with the number of cases of D^+ and D^- set at 1000.

2.4 Diagnostic Systems

The following three models were examined as the diagnostic system.

(1) LG1: Logistic model.
(2) NN1: Neural net with a single layer.
(3) NN2: Neural net with two layers with 5 cells.

Table 2: Test data generated by the simulation.

Outcome of the tests				Frequency	
T_1	T_2	T_3	T_4	D^+	D^-
0	0	0	0	14	104
1	0	0	0	21	69
0	1	0	0	21	69
1	1	0	0	32	46
0	0	1	0	50	112
1	0	1	0	75	75
0	1	1	0	75	75
1	1	1	0	112	50
0	0	0	1	50	112
1	0	0	1	75	75
0	1	0	1	75	75
1	1	0	1	112	50
0	0	1	1	46	32
1	0	1	1	69	21
0	1	1	1	69	21
1	1	1	1	104	14

2.5 Evaluation of Diagnostic Performance

As the indices of performance of the system, we calculated the Somers'D (Gini coefficient), Goodman-Kruskal gamma, Kendall's Tau-a, and the concordance index, c, which are closely related to each other. We chose to use the value of the concordance index for each result, because this index is known to give the area under the receiver operating characteristic (ROC) curve of the diagnostic system.

2.6 Computation Methods

SAS 9.1.3 was used for logistic analysis and MATLAB (Neural Net Toolbox) was used for the NN1 and NN2 calculations.

3 RESULTS

For cases with $N = 3-7$, the sensitivity and specificity were set to 0.6. For each case, computation was performed under the following three conditions:

(a) Independent: $R^+ = R^- = I$
(b) Positive correlation: $R^+(1,2) = R^-(1,2) = 0.3$
(c) Negative correlation: $R^+(1,2) = R^-(1,2) = -0.3$

3.1 Comparison of the Diagnostic Systems

We did not find any significant differences among the three diagnostic systems, LG1, NN1 and NN2. The results from NN1 are shown in Table 3.

Table 3: Concordance index c.

N	(a)Independent	(b)Positive R	(c)Negative R
3	0.683	0.665	0.697
4	0.710	0.693	0.714
4	0.737	0.720	0.740
4	0.758	0.745	0.761
4	0.759	0.757	0.781

3.2 Effect of the Number of Tests

The concordance index, c, increased as the number of tests increased. The ROC curve for each case is shown in Figure 1.

3.3 Effect of Correlation between Tests

As shown in Table 3, the diagnostic performance of the combined tests was worse in a case of positive correlation between tests and better in a case of negative correlation, compared to the independent case.

4 CONCLUSIONS

Examination of the improvement of diagnostic performance by combining multiple tests requires an algorithm for generating simulated data with arbitrary sensitivity, specificity and inter-test correlations.

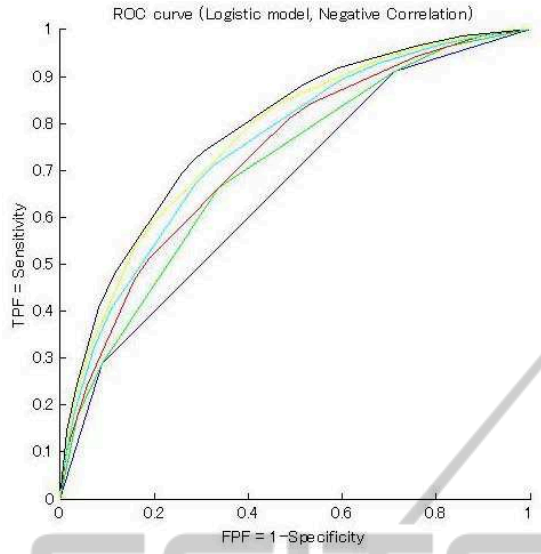

Figure 1: ROC curve of the diagnostic system for different number of tests, N. N=2 (blue), 3 (green), 4 (red), 5 (cyan), 6 (yellow) and 7 (black).

The effects of the number of tests and inter-test correlations on the diagnostic performance were studied using a logistic model and neural network models.

The diagnostic performance measured by the concordance index, c, increased as the number of tests increased. For the same number of tests, the diagnostic performance was reduced by positive correlation and elevated by negative correlation. Improvement of the performance was not obtained by increasing the number of NN layers.

ACKNOWLEDGEMENTS

This study was funded in part by a grant from the Kitasato University School of Allied Health Sciences (No. 2010-6604).

REFERENCES

Bahadur, R. R. (1961). A representation of the joint distribution of responses to n dichotomous items. In Solomon, H., editor, *Studies in Item Analysis and Prediction*, pages Chapter 9:158–160. Stanford University Press.

Hara, H. (2008). A logistic analysis of left ventricular ejection fraction (LVEF) after CRT. In *American Heart Association 2008*.

Ikeda, N., Bax, L., Henmi, O., Mamorita, N., Tsuruta, H., Shibata, S., and Takeuchi, A. (2007). Study of a logistic model with mutually correlated variables using a generation algorithm of dichotomous data with arbitrary sensitivity, specificity and correlation. In *MEDINFO 2007*. , Brisbane, Australia. (Proc 2486-2488).

Ikeda, N., Shibata, S., Bax, L., Henmi, O., Mamorita, N., Tsuruta, H., and Takeuchi, A. (2006). Diagnostic performance of combined tests using a generation algorithm of multiple tests with arbitrary sensitivity, specificity and correlation. In *MEDSIP 2006*. Glasgow, UK.