Using the Bernoulli trial approaches for detecting ordered alternatives

Chia-Hao Chang1*, Chih-Chien Chin2, Weichieh Wayne Yu1 and Ying-Yu Huang1

Abstract

**Background:** Diagnostic problems in clinical trials are sometimes ordinal. For example, colon tumor staging was performed according to the TNM classification. However, clinical data are limited by markedly small sample sizes in some stage.

**Methods:** We propose a distribution-free test for detecting ordered alternatives in a completely randomized design. The new statistic is based on summing all correctly (ascending) ordered samples.

**Results:** The exact mean and variance of the null distribution are derived and it is shown that this distribution is asymptotically normal. Furthermore, we show using Monte Carlo simulation that the proposed test is a significant improvement over the Terpstra-Magel test. That is, power is decreased where the investigator falsely assumes an a priori ordering relationship.

**Conclusions:** We conclude that these tests frequently detect an ordered trend when, in fact, one does not exist. However, the new test can reduce the error rate, at least not to the extent in which the Jonckheere-Terpstra test stated.

**Background**

This paper focuses on considering nonparametric tests for the non-decreasing ordered alternative of \( k(\geq 3) \) groups. The hypothesis to be tested is \( H_0^i : F_1(x) = F_2(x) = \cdots = F_k(x) \) for all \( x \) and \( H_1 : F_1(x) \geq F_2(x) \geq \cdots \geq F_k(x) \), for all \( x \) with \( F_1(x) > F_k(x) \) for some \( x \), where \( F_1(x), F_2(x), \cdots, F_k(x) \) are continuous distribution functions.

In this article, we assume the location model with \( F_i(x) = F(x - \mu - \theta_i) \), where \( \mu \) is a location parameter and \( \theta_i \) represents the effect of group \( i, i = 1, 2, \ldots, k \). This implies that the underlying populations may differ only in location. Throughout the article, let \( x_{i1}, x_{i2}, \ldots, x_{im_i}, i = 1, 2, \ldots, k \) represent independent random samples from the \( k \) populations with distribution functions \( F_i(x), i = 1, 2, \ldots, k \), respectively.

Nonparametric order restricted inference has been extensively investigated in past literature and new studies are continuing to emerge. For instance, Puri [1], Puri and Sen [2], and Padmanabhan et al. [3] applied the concept of Chernoff–Savage-type statistics to nonparametric ordered alternative tests. Studies that used power results to compare the validity of linear rank tests included Büning and Köessler [4], Beier and Büning [5], Büning and Köessler [6], Büning [7], Büning and Köessler [8], Büning and Köessler [9], Köessler [10] and Köessler [11].

The earliest and most classic treatment of \( k(\geq 3) \)-sample distribution-free statistic for ordered alternatives was proposed by Jonckheere [12] and Terpstra [13]. The test is known as the Jonckheere-Terpstra test (hereafter referred to as the JT test) and is based on a sum of \( C_k^1 \) Mann–Whitney statistics (Mann and Whitney, [14]; Hollander and Wolfe, [15]). In order to define the JT statistic, we express the Mann–Whitney statistics as

\[
U_{lm} = \sum_{j=1}^{n_l} \sum_{i=1}^{n_m} I(x_{lj}, x_{mi}), \quad 1 \leq l < m \leq k,
\]

where \( I(x_{lj}, x_{mi}) = \begin{cases} 1, & \text{if } x_{lj} < x_{mi} \\ 0, & \text{otherwise} \end{cases} \), and the JT statistic is given by

\[
JT = \sum_{l=1}^{k-1} \sum_{m=l+1}^{k} U_{lm}.
\]

Other tests for ordered alternatives were developed by Cuzick [16] and Le [17]. Among the JT, Cuzick, and Le
tests, the results from Mahrer and Magel [18] did not establish any of the tests as having overwhelmingly higher power over the others across different location parameters. Neuhauser et al. [19] presented a modified version of the JT test (hereafter referred to as the MJT test). The form of the MJT statistic is identical to the JT test except that the Mann–Whitney statistic \( U_{bm} \) multiplies the weight \( m \cdot l \) as the new kernel. Study results showed that the MJT test often produced a higher power than the JT test for the ordered alternative. We also noted that Tryon and Hettmansperger [20] presented the JT and MJT tests as members of a more general class of nonparametric tests. In the test statistics described above the kernels of the tests are almost all derived by comparing two pairs of sample observations at a time. However, Terpstra and Magel [21] proposed a test (hereafter referred to as the TM test) where the kernels of the TM test are based on information obtained simultaneously across all samples. The statistic is determined by adding the \( \prod_{i=1}^{k} n_i \) indicator functions, that is,

\[
TM = \sum_{l_1=1}^{n_1} \cdots \sum_{l_k=1}^{n_k} I(x_{l_1} \leq x_{l_2} \leq \cdots \leq x_{l_k})
\]

where \( I(x_{l_1} \leq x_{l_2} \leq \cdots \leq x_{l_k}) \) is equal to one, provided at least one strict inequality; otherwise, \( I(x_{l_1} \leq x_{l_2} \leq \cdots \leq x_{l_k}) \) is equal to zero.

Terpstra et al. ([22]a, b) proposed a new nonparametric test statistic (hereafter referred to as the KTP test) that is a generalization of the TM test. The idea is to replace the indicator kernel from the TM test with Spearman’s rank correlation coefficient, that is,

\[
KTP = \sum_{l_1=1}^{n_1} \cdots \sum_{l_k=1}^{n_k} r(x_{l_1}, x_{l_2}, \ldots, x_{l_k}),
\]

where \( r(x_{l_1}, x_{l_2}, \ldots, x_{l_k}) \) denotes Spearman’s rank correlation coefficient between the observed data and the corresponding group number.

In this study, we propose a new test is based on the information present in the \( N^* = \prod_{i=1}^{k} n_i \) k-tuplets, where a k-tuplet includes one observation from each treatment group. All correctly (ascending) ordered samples are then summed to form a statistic that is distributed approximately as a normal distribution. Details of this new test and its asymptotic distribution are provided, and the computational algorithm is presented in the Additional file 1.

A colon cancer data example is given in data example section. Finally, we present a finite sample simulation study which compares the proposed test, the JT test, MJT test, TM test, and the KTP test in terms of power. A computer program written in R that implements the proposed methods will be available from the first author upon request. It is recommended that readers who are not interested in the details of the computational algorithm skip the Additional file 1.

**Methods**

**Test statistic**

The new nonparametric test for non-decreasing alternatives is based on the following statistic,

\[
T = \sum_{l_1=1}^{n_1} \cdots \sum_{l_k=1}^{n_k} k(x_{l_1}, x_{l_2}, \ldots, x_{l_k}).
\]

Where \( k(x_{1}, x_{2}, \ldots, x_{k}) = \sum_{i=1}^{k} I(R(x_i) = i) \), \( R(x_i) \) denotes the rank of \( x_i \) with respect to \( x_{1}, x_{2}, \ldots, x_{k} \), and \( I(.) \) denotes the indicator function.

The remainder of this section presents and derives results pertaining to the null distribution of the proposed test statistic. We assume throughout this section that the observed data, \( \{X_{ij}\} \) is essentially a random sample from some continuous probability distribution function \( F \). Hence, the possibility of ties has a probability of zero. In principle the test statistic uses the k-tuplet method of Terpstra and Magel. Additionally, in the null hypothesis each \( k(x_{1}, x_{2}, \ldots, x_{k}) \) follows the Binomial \((k, 1/k)\) distribution. For these reasons, we will refer to this test as the KTMB test.

**The exact null distribution**

Let \( N \) denote the sum of the sample sizes for each treatment. Namely, let \( N = n_1 + \cdots + n_k \). Here, we have \( N!/(n_1! \cdots n_k!) \) partitions of the numbers 1, ..., \( N \). The null distribution of \( T \) means each one of these partitions is equally likely so the mean and variance can be calculated directly by multiplying each possible value of \( T \) with its probability. When the number of partitions is small, we can easily calculate the exact distribution by hand or with the computer. Table 1 shows the probabilities, means, and variances of the test statistic \( T \) for sample size arrangements (2, 1, 1), (2, 1, 2), and (1, 1, 3) respectively.

Real world cases are not always as simple as the above illustration. For example, when \( k = 4 \), \( n_1 = n_2 = n_3 = n_4 = 5 \), we have \( 20!/(5!5!5!5!) = 1.1733 \times 10^{10} \) partitions. Such a distribution function cannot be calculated even with the most efficient personal computers. We will therefore introduce a Monte Carlo approximation to the null distribution. On the other hand, if the distribution of \( T \) can be approximated or can be shown to converge to a well-known distribution, we can avoid computational complexity altogether.
Table 1 Some exact null distributions for the proposed test statistic

| Sample sizes (n1, n2, n3) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 12 | Mean and variance |
|---------------------------|---|---|---|---|---|---|---|---|---|---|----|------------------|
| (2, 1, 1)                 | 2/12 | 3/12 | 4/12 | 1/12 | 1/12 | - | 1/12 | - | - | - | - | 2 | 2.667             |
| (2, 1, 2)                 | 2/30 | 2/30 | 6/30 | 4/30 | 6/30 | 2/30 | 3/30 | 2/30 | 2/30 | - | 1/30 | 4 | 6.867             |
| (1, 1, 3)                 | 2/20 | 3/20 | 4/20 | 6/20 | 1/20 | 1/20 | 1/20 | 1/20 | 1/20 | - | 1/20 | - | 3 | 5.000             |

The mean and variance

If the asymptotic null distribution of a test statistic is normal and the exact mean and variance of T under H0 in standard form can be established we can then standardize T by using the exact mean and variance to obtain Z_{KTMB} where \( Z_{KTMB} = \frac{\{T - E_0(T)\}}{\sqrt{V_0(T)}} \). In this case, we can find critical values from the standard normal table.

We will start by finding the mean value of T, \( E_0(T) \), noting that T is nothing but a sum of the \( k \times \prod_{i=1}^{k} n_i \) Bernoulli (1/k) distribution. It is straightforward to get

\[
E_0(T) = \prod_{i=1}^{k} n_i
\]  

(1)

Here, and in the following, we let \( n^* = \prod_{i=1}^{k} n_i \).

\[
V_0(T) = v_0^2 + \sum_{i=1}^{k-1} v_i^2 + v_k^2
\]  

(2)

where \( v_0^2 = \{1 - (1/k)\} \prod_{i=1}^{k} n_i \) for no tie, \( v_k^2 = n^* k(k-1) \)

\( \{ (k-2)!/(k-1)!/k^2 \} \) for k ties for \( i \neq j \).

For the case of \( i \) ties, we present an algorithm for the computation of the \( \sum_{j=1}^{k-1} v_j^2 \) in Additional file 1. Readers who are not interested in the details of this algorithm may want to skip the Additional file 1 and go to data example section, in which examples based on real data are provided.

The asymptotic null distribution

In this section we will look to see if the asymptotic null distribution of test statistic T follows the standard normal distribution. In other words, we prove that

\[
T^* = \frac{T - E_0(T)}{\sqrt{V_0(T)}} \xrightarrow{D} N(0, 1).
\]  

(3)

H_0 will therefore be rejected for large values of \( T^* \). The normal approximation for the procedure is to reject \( H_0 \) if \( T^* \geq z_{1-\alpha} \), otherwise do not reject \( H_0 \). Note that the critical value \( z_{1-\alpha} \) is chosen to make the Type I error probability equal to \( \alpha \). That is, \( \alpha = P(T^* \geq z_{1-\alpha}|H_0 \text{ true}) \).

We note that (3) is a direct consequence of Theorem 1, which we now state.

**Theorem 1** Let \( N = \sum_{i=1}^{k} n_i \) and assume \( n_i = \lambda_1 + o(1) \) where \( \lambda_1 \in (0, 1) \). Then, under \( H_0 \), \( T_N \) defined as

\[
T_N = \frac{1}{k N^{-1/2}} \sum_{i=1}^{n} \sum_{j=1}^{m} [I(j = m) Z_{ik}]
\]

\[
x(\lambda_{ik}) + i(j = m) \frac{1}{k} \frac{1}{k^2} = \frac{L_a(l)}{N^{k-1/2}} \left[ Z_{ik}(x_{ik}) - \frac{1}{k} \right]
\]

\[
Z_{ik}(x_{ik}) = \frac{(k-1)!}{(l-1)!(k-l)!} \left[ 1 - F(x) \right]^{k-l} \left[ 1 - F(x) \right]^{l-1}, \quad l = 1, \ldots, k
\]

where

\[
L_a(l) = \prod_{j=1}^{l} j^{(l-j)}.
\]

The projection of \( T_N \) say, \( P_N \) can be defined as,

\[
P_N = \sum_{i=1}^{k} \sum_{m=1}^{n_i} E[T_N | X_{im}]
\]

and

\[
E[Z_{ik}(x_{ik})] = \frac{1}{k} \quad \text{and} \quad V[Z_{ik}(x_{ik})] = \sigma_{ik}^2
\]

can be proved by Beta distribution. The convergence criteria on the \( k \) sample sizes imply that,

\[
L_a(l) \sqrt{n_i} \left\lfloor \frac{n_i}{N} \right\rfloor = \sqrt{\lambda_{ik}} + o(1).
\]  

(5)
It now follows from (4), (5), and limiting moment generating function theory that,
\[ P_N \overset{D}{\rightarrow} N\left(0, \sum_{i=1}^{k} \lambda_i^2 \sigma_i^2 \right). \tag{6} \]

Let us now consider \( V[T_N] \), which we write as,
\[ V[T_N] = \frac{1}{k!N^{2k-1}} \sum_{i=1}^{n_1} \cdots \sum_{i=1}^{n_k} \left( \sum_{i=1}^{n_1} \cdots \sum_{i=1}^{n_k} \cdots \sum_{i=1}^{n_k} \right) \times \text{COV}\left[k(x_{1i}, \ldots, x_{ki}), k(x_{1i}, \ldots, x_{ki})\right]. \tag{7} \]

Consider first the case of \( k \) ties for \( i \neq j \), it is straightforward to show \( \text{COV}\left[k(x_{1i}, \ldots, x_{ki}), k(x_{1i}, \ldots, x_{ki})\right] = k \frac{\binom{k-2}{k-1} \frac{N!}{k!}}{k!} \). Next, consider the case in which there are exactly three ties among the different subscripts. For example, if we let \( R_u \) denotes the rank of \( x_u \) with respect to \( x_1, x_2, \ldots, x_k \), \( R_v \) denotes the rank of \( x_v \) with respect to \( x_1, x_2, \ldots, x_k \), \( R_u < R_v \) and \( x_1, x_2, x_3 \) denote the tied observations then the covariance term has the form \( \text{COV}[I_u, I_v] \) where, \( R_u \) denotes the rank of \( x_u \) with respect to \( x_1, x_2, \ldots, x_k \), \( x_{t_1}, x_{t_2}, \ldots, x_{t_k} \), \( x_{t_1+1}, \ldots, x_{t_2}, \ldots, x_{t_k} \) and \( I_u = I(R_u = 1) \), and \( I_v = I(R_v = 1) \). Under \( H_0 \), \( E[I_u] = E[I_v] = \frac{1}{k!} \). Next, consider \( E[I_u I_v] \). This expectation contains \( 2k-3 \) observations, so that under \( H_0 \), each of the \( (2k-3)! \) permutations of the observations are equally likely. However, there are only the numbers of \( (1: R_u < R_v) \cap (1: R_v < 1) \) possible ways, say \( I_{u,v}(R_u, R_v) \), plus the numbers of \( (R_u < 1: 2k-3) \cap (1: R_v < 1) \) possible ways, say \( I_{u,v}(R_u, R_v) \), plus the numbers of \( (R_u < 1: 2k-3) \cap (R_v < 1: 2k-3) \) possible ways, say \( I_{u,v}(R_u, R_v) \), to preserve \( x_1, x_2, \) and \( x_3 \). Furthermore, there are \( (k-3)! \) possible ways to preserve \( x_{t_1}, \ldots, x_{t_k} \), \( x_{t_1+1}, \ldots, x_{t_k} \) to the left of \( x_u \), \( C_{u-1}(R_u, R_v) \) possible ways to preserve \( x_{t_1}, \ldots, x_{t_k} \), \( x_{t_1+1}, \ldots, x_{t_k} \) to the left of \( x_v \), \( C_{v-1}(R_v, R_u) \) possible ways to preserve \( x_{t_1}, \ldots, x_{t_k} \), \( x_{t_1+1}, \ldots, x_{t_k} \) to the right of \( x_u \) and \( x_v \), \( C_{u-1}(R_u, R_v) \) possible ways to preserve \( x_{t_1+1}, \ldots, x_{t_k}, x_{t_1+1}, \ldots, x_{t_k} \) to the right of \( x_u \) and \( \frac{1}{k!} ! \) possible way to preserve \( x_{t_1+1}, \ldots, x_{t_k+3}, x_{t_1+2}, \ldots, x_{t_k}, \ldots, x_{t_k-3}, x_{t_1+2}, \ldots, x_{t_k}, \ldots, x_{t_k-3} \) to the right of \( x_v \). Hence, these arguments imply that,
\[ \text{COV}(I_u, I_v) = \frac{3!(k-4)! C_{t_1+2+t_2+1}(R_u, R_v) + C_{t_1+2+t_2+1}(R_v, R_u) + C_{t_2+1+t_2+1}(R_u, R_v)}{(2k-3)!} \times \frac{1}{k!}! \]

where \( t_1 + t_2 + t_3 = 3, t_1, t_2, \) and \( t_3 = 0, 1, 2, 3 \).

Now, for a given \( I_1, I_2, \) and \( I_3 \) there are \( n_1 n_2 n_3 \prod_{l=1}^{k} [n_l(n_l-1)]^{(l+1)l(l+1)} \) of these covariance terms. Next, consider all possible \( R_u, R_v \), and all possible treatment locations (\( m \) and \( n \)), in the case of one tie \( X_i \), (7) reduces to,
\[ \sum_{i=1}^{n} \sum_{k=0}^{N^2} \left[ \sum_{m=0}^{n_1} \sum_{n=0}^{n_2} \right] \frac{C_{t_1+2+t_2+1}(R_u, R_v) + C_{t_1+2+t_2+1}(R_v, R_u) + C_{t_2+1+t_2+1}(R_u, R_v)}{(2k-3)!} \times \frac{1}{k!}! \]

where \( t_1 + t_2 + t_3 = 1, t_1, t_2, \) and \( t_3 = 0, 1, 2 \).

From (6) and (8) it follows that \( V(T_N) - V[P_N] = o(1) \). Asymptotic normality results are attainable.

**Patient characteristics**

The institutional review board of Chang Gung Memorial Hospital approved the present study. Detailed information about patients with colon cancer, such as patient- and tumor-related factors and follow-up status, was retrieved from the Colorectal Section Tumor Registry at Chang Gung Memorial Hospital, Taiwan. All the data in this registry were prospectively collected.

**Results and discussion**

**Data examples**

Between January 2006 and December 2010, 154 consecutive patients with histologically confirmed colonic adenocarcinoma underwent curative surgeries at the Chang Gung Memorial Hospital in Chiayi. The stage IV colon cancer, non-curative surgeries, rectal cancer and mucinous adenocarcinomawere excluded in this study. Tumor staging was performed according to the TNM classification described in the 6th edition of the cancer staging manual of the American Joint Committee on Cancer (Stage I, II, IIIIA and IIIB). The different tumor staging require a different treatment to optimize patient and hospital outcomes. An ordinal logistic regression model was developed with predictors as follows: age, gender, tumor location, histologic differentiation, preoperative
albumin level, preoperative carcinoembryonic antigen level, and underlying medical illnesses.

To illustrate the KTMB test, assume an outcome with four stages and a set of cases consisting of one case from each stage. The case from Stage I has risks of 0.50, 0.25, 0.15 and 0.10 for Stage I, II, IIIA and IIIB, respectively. The case from Stage II has risks 0.26, 0.52, 0.17 and 0.05; the case from Stage IIIA has risks 0.06, 0.32, 0.42 and 0.20; the case from Stage IIIB has risks 0.12, 0.18, 0.30 and 0.40. The risk for Stage IIIB (say, event) is higher for the case that belongs to this stage (0.40) than for the other cases (0.10, 0.05 and 0.40). The risk for event is second-highest for the case from Stage IIIA (0.20 versus 0.10, 0.05 and 0.40). However, the risk for event is lowest for the case from Stage II (0.05 versus 0.10, 0.20 and 0.40). The risk for event is third-highest for the case from Stage I (0.10 versus 0.05, 0.20 and 0.40). Therefore, the risks correctly identify the cases from Stage IIIA and IIIB but not Stage I and II, resulting in a score of 2 for this set \( k(x_1, x_2, x_3, x_4) \).

Hence, the set of hypotheses was \( H_0 : F_I(x) = F_{II}(x) = F_{IIIA}(x) = F_{IIIB}(x) \) for all \( x \) and \( H_1 : F_I(x) \geq F_{II}(x) \geq F_{IIIA}(x) \geq F_{IIIB}(x) \), where \( F(x) = F_{IIIB}(x) \) for some \( x \).

Five test statistics and the corresponding p-values are given in Table 2. Since a plot of this data set in Figure 1 exhibits a non-increased trend, it appears that the JT, MJT and KTP tests have falsely conclusion (p < 0.05). The KTMB test has the largest p-value (See Table 2).

Moreover, Stage IIIB patients reported significantly more risk for Stage IIIB than Stage II subjects (ANOVA, post hoc: IIIB > II, p = 0.003) while Stage I, II and IIIA patients did not differ in risk for Stage IIIB (ANOVA, post-hoc: p>0.05) Hence, we conclude that the risk for Stage IIIB do not increase with the patient’s TNM in the model. That is, the discrimination performance of the ordinal logistic model is not very well between Stage I, II and IIIA.

### Comparison with respect to size and power

To determine if the underlying population came from different skew and kurtosis distributions that impact on the power of the test statistic, we used log-F distributions with combinations of 2, 4.5 and 10 degrees of freedom.
freedom to generate the random variable. We can therefore define random variable $X_{ij}$ as: $X_{ij} = \theta_i + \epsilon_{ij}$, where $\epsilon_{ij}$ is the iid log-F distribution, and $\theta_i$ are location parameters.

For the numbers of treatment ($k$), sample sizes ($n_j$) and location parameters ($\theta$) we examine the different combinations of $k = 3$ and 4, $n_j = 4, 5, 8$ and 10, $\theta = 0, 0.25, 0.5, 0.75, 1$ and 1.25. We investigated designs under assumed alternatives which are of the forms of concave and convex. Programs to compare powers were written in R 2.9.2 (R Development Core Team, Vienna, Austria).

The estimations were conducted by simulating 10,000 different sets of samples. Furthermore, we estimated the power by counting the number of times $H_0$ was rejected and using the value to divide by 10,000. Ideally, we believe that the test should have higher power than a general alternative test when $H_1$ is true, and should have low power for any alternative that does not fit the profile given in $H_1$.

In general, the JT and KTP tests have the highest powers for the ordered alternative cases. Comparing with TM test, the gain percentage in power, $DP = (KTMB - \cdot 4.5) - 0.492 \cdot 0.496 - 0.494 \cdot 0.496 - 0.494 \cdot 0.494$

| Location parameter | KTMB | KTP | JT | MJT | TM | DP    |
|-------------------|------|-----|----|-----|----|------|
| n1=4, n2=4, n3=4 | (0, 0, 0) | 0.0492 | 0.0496 | 0.0489 | 0.0496 | 0.0494 | (0, 0, 0) | 0.0494 | 0.0481 | 0.0488 | 0.0494 | 0.0509 |
| (0, 0.25, 0.5)   | 0.1176 | 0.1243 | 0.1096 | 0.1243 | 0.1255 | -6.29% |
| (0.5, 0, 0.25)   | 0.0295 | 0.0344 | 0.0295 | 0.0344 | 0.0346 | 0.00%  |
| (0.25, 0.75, 0)  | 0.021 | 0.029 | 0.0212 | 0.029 | 0.0236 | 0.95%  |
| Log-F (4.5, 4.5) | (0, 0, 0) | 0.0511 | 0.0521 | 0.0518 | 0.0521 | 0.0531 | (0, 0, 0) | 0.0492 | 0.0478 | 0.0501 | 0.0488 | 0.0512 |
| (0, 0.25, 0.5)   | 0.1533 | 0.1612 | 0.1447 | 0.1612 | 0.1604 | -4.43% |
| (0.5, 0, 0.25)   | 0.0189 | 0.0259 | 0.0212 | 0.0259 | 0.0237 | 12.17% |
| (0.25, 0.75, 0)  | 0.0207 | 0.0246 | 0.0213 | 0.0246 | 0.0239 | 2.90%  |
| Log-F (10, 4.5)  | (0, 0, 0) | 0.0509 | 0.0523 | 0.0519 | 0.0523 | 0.0527 | (0, 0, 0) | 0.0488 | 0.0494 | 0.0517 | 0.0505 | 0.051 |
| (0, 0.25, 0.5)   | 0.1802 | 0.1961 | 0.1742 | 0.1961 | 0.1843 | -2.22% |
| (0.5, 0, 0.25)   | 0.016 | 0.021 | 0.0166 | 0.021 | 0.0192 | 3.75%  |
| (0.25, 0.75, 0)  | 0.0161 | 0.0196 | 0.0161 | 0.0196 | 0.0198 | 0.00%  |
| Log-F (24.5)     | (0, 0, 0) | 0.0512 | 0.0519 | 0.0515 | 0.0519 | 0.0514 | (0, 0, 0) | 0.0505 | 0.0484 | 0.0518 | 0.0507 | 0.0479 |
| (0, 0.25, 0.5, 0.75) | 0.165 | 0.1806 | 0.188 | 0.1806 | 0.1615 | 2.17%  |
| (0.75, 0, 0.25, 0.5) | 0.0388 | 0.0388 | 0.0417 | 0.0388 | 0.0471 | 0.00%  |
| (0.25, 0.75, 1.25, 0) | 0.0225 | 0.0315 | 0.0313 | 0.0315 | 0.0276 | 22.67% |
| Log-F (4.5, 4.5) | (0, 0, 0) | 0.0507 | 0.0503 | 0.0509 | 0.0503 | 0.0508 | (0, 0, 0) | 0.0505 | 0.0487 | 0.0508 | 0.0477 | 0.048 |
| (0, 0.25, 0.5, 0.75) | 0.2242 | 0.2513 | 0.2606 | 0.2513 | 0.2143 | 4.62%  |
| (0.75, 0, 0.25, 0.5) | 0.0297 | 0.0282 | 0.0304 | 0.0282 | 0.0385 | -5.05% |
| (0.25, 0.75, 1.25, 0) | 0.0237 | 0.0291 | 0.0298 | 0.0291 | 0.0277 | 16.88% |
| Log-F (10, 4.5) | (0, 0, 0) | 0.0527 | 0.0513 | 0.0509 | 0.0513 | 0.0527 | (0, 0, 0) | 0.0504 | 0.0492 | 0.0493 | 0.0514 | 0.0508 |
| (0, 0.25, 0.5, 0.75) | 0.2882 | 0.3179 | 0.3316 | 0.3179 | 0.259 | 11.27% |
| (0.75, 0, 0.25, 0.5) | 0.0168 | 0.0211 | 0.0204 | 0.0211 | 0.0216 | 21.43% |
| (0.25, 0.75, 1.25, 0) | 0.025 | 0.0259 | 0.0229 | 0.0259 | 0.0317 | 3.60%  |

Table 3 Estimated powers and type I error rates of ordered tests under significance level 0.05
TM)/TM, ranges from −6.29% to 11.27% with the average gain percentage in power being 2.63% (difference of percentage).

Consider the corresponding alternatives of the form of concave and convex shapes. The powers of the KTMB test outperforms (lower power) the KTP, JT, MJT, and TM tests when balanced design. The loss percentage in power, $DP = (\min(\text{KTP, JT, MJT, and TM}) - \text{KTMB})/\text{KTMB}$, ranges from −7.28% to 20.00% with the average loss percentage in power being 4.25% for $k = 3$. The DP ranges from −13.2% to 27.68% with the average loss percentage in power being 8.83% for $k = 4$.

When the sample sizes corresponding to the non-decreasing trend location parameters are comparatively large, the KTMB test is better than KTP, MJT, JT, and TM tests. The DP ranges from −7.29% to 9.30% with the average loss percentage in power being 2.47% for $k = 3$. The DP ranges from −6.52% to 30.93% with the average loss percentage in power being 12.07% for $k = 4$. However, the KTP test slightly better than KTMB test when the underlying population is skewed to the right (see Table 3).

Based on the simulation results above, we conclude that the KTMB test is better than the TM test in regards to the power against ordered alternatives. Moreover, the KTMB test offers built in protection for the situation when an investigator falsely assumes an a priori ordered relationship.

Table 3 just represent a small subset of the many different scenarios that we simulated. For example, we also conducted simulations for numerous other alternative patterns. Interested persons may contact the corresponding author for these simulated results.

Conclusions
This research proposes a new nonparametric test for the ordered alternative problem. The new test statistic is based on the calculating all $k(x_{i1}, x_{i2}, \ldots, x_{ik})$ in proper (ascending) order. In other words, the new test statistic collects the information of each observation for each treatment to provide the message of “increasing” to the test statistics. A higher test statistics means a stronger “increasing” message. This is also why we expect the new test statistics to offer better power under certain situations.

Due to the small number of groups and sample sizes, we tabulated and listed their distribution as well as the exact mean and variance of the null distribution. From the equation for the exact mean and variance of the null distribution was derived and the asymptotic null distribution is normal were given.

We also use the example of ordinal risk prediction of colon cancer to compare the test statistics mentioned in the papers. A finite sample simulation study was also used to explore in-depth how the powers of JT, MJT, TM, KTP and KTMB tests under different underlying populations, treatment numbers and sample sizes. Based on the example and simulation results, we conclude that these tests frequently detect an ordered trend when, in fact, one does not exist. However, the KTMB test can reduce the error rate, at least not to the extent in which the JT and MJT tests do.

Ben Van Calster et. al. extend the main measure of binary discrimination, the c-statistic or area under the ROC curve, to nominal polytomous settings by polytomous discrimination index (PDI) [24]. They mention it is desirable that the risk of each group is highest for the case that belongs to this group in a set of cases. Therefore, the PDI score awarded to a set equals the number of groups for which this holds. Based on this point of view, in our opinion, the KTMB test can not only be used for detecting the non-decreasing alternatives but can also be measured to summarize polytomous discrimination.

Additional file

Additional file 1: Algorithm for computing $\sum_{i=1}^{k-1} v_i^2$.

Competing interests
The authors declare that they have no competing interest.

Authors’ contributions
CHC designed the study, prepared the manuscript and performed statistical analyses. WY drafted and assisted with the revision of the article. CCC and YYH participated in and carried out the field work. All authors read and approved the final manuscript.

Acknowledgements
This research was supported in part by the National Science Council, Taiwan, ROC, under NSC99-2118-M-255-001.

Author details
1Department of Nursing, Chang Gung University of Science and Technology, Chiayi Campus, Chiayi, Taiwan 61363. 2Department of Surgery, Section of Colon and Rectal Surgery, Chang Gung Memorial Hospital, Chiayi, Taiwan.

Received: 4 September 2013 Accepted: 29 November 2013
Published: 5 December 2013

References
1. Puri ML: Some distribution-free k-sample rank tests of homogeneity against ordered alternatives. Commun Pure Appl Math 1965, 18:51–63.
2. Puri ML, Sen PK: On chernoff-savage tests for ordered alternatives in randomized blocks. The Annals of Mathematical Statistics 1968, 39(3):967–972.
3. Padmanabhan AR, Puri ML, Saleh AKME: A non-parametric test for equality against ordered alternatives in the case of skewed data with a biomedical application, Statistics and related topics (Ottawa, Ont. 1980). North-Holland, Amsterdam: North-Holland Publishing Company; 1981:279–283.
4. Buning H, Kössler W: Robustness and efficiency of some tests for ordered alternatives in the C-Sample location problem. J Stat Comput Simul 1996, 55:337–352.
5. Beier F, Buning H: An adaptive test against ordered alternatives. Computational Statistics & Data Analysis 1997, 25:441–452.
6. Büning H, Kössler W: The asymptotic power of Jonckheere-type tests for ordered alternatives. Australian & New Zealand Journal of Statistics 1999, 41(1):67–77.
7. Büning H: Adaptive Jonckheere-type tests for ordered alternatives. J Appl Stat 1999, 26:541–551.
8. Kössler W, Büning H: The asymptotic power and relative efficiency of some c-sample rank tests of homogeneity against umbrella alternatives. Statistics 2000, 34(1):1–26.
9. Kössler W, Büning H: The efficacy of some c-sample rank tests of homogeneity against ordered alternatives. Journal of Nonparametric Statistics 2000, 13(1):95–106.
10. Kössler W: Some c-sample rank tests of homogeneity against ordered alternatives based on U-statistics. Journal of Nonparametric Statistics 2005, 17(7):777–795.
11. Kössler W: Some c-sample rank tests of homogeneity against umbrella alternatives with unknown peak. J Stat Comput Simul 2006, 76(1):57–74.
12. Jonckheere AR: A distribution-free k-sample test against ordered alternatives. Biometrika 1954, 41:133–145.
13. Terpstra T: The asymptotic normality and consistency of Kendall’s test against trend when ties are present in one ranking. Indagationes Mathematica 1952, 14:327–333.
14. Mann HB, Whitney DR: On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 1947, 18(1):50–60.
15. Hollander M, Wolfe DA: Nonparametric statistical methods. New York: John Wiley, 1999.
16. Cuzick J: A wilcoxon-type test for trend. Stat Med 1985, 4:87–90.
17. Le CT: A new rank test against ordered alternatives in k-sample problems. Biom J 1988, 30(1):87–92.
18. Mahier J, Magel R: A comparison of tests for the k-sample, non-decreasing alternative. Stat Med 1995, 14(8):863–871.
19. Neuhausser M, Liu PY, Hothorn L: Nonparametric tests for trend: Jonckheere’s test, a modification and a maximum Test. Biom J 1998, 40(8):899–909.
20. Tryon VP, Hettmansperger TP: A Class of non-parametric tests for homogeneity against ordered alternatives. Ann Stat 1973, 1:1061–1070.
21. Terpstra JT, Magel RC: A new nonparametric test for the ordered alternative problem. Journal of Nonparametric Statistics 2003, 15(3):289–301.
22. Terpstra JT, Chang CH, Magel RC: On the use of spearman’s correlation coefficient for testing ordered alternatives. J Stat Comput Simul 2011, 81(11):1381–1392.
23. Hettmansperger TP, McKean JW: Robust nonparametric statistical methods. Great Britain: Arnold, 1998.
24. Calster Br, Belle Wv, Vergouwe Y, Timmerman D, Huffel VS, Steyerberg EW: Extending the C-statistic to nominal polytomous outcomes: the polytomous discrimination index. Stat Med 2012, 31:2610–2626.

doi:10.1186/1471-2288-13-148
Cite this article as: Chang et al: Using the Bernoulli trial approaches for detecting ordered alternatives. BMC Medical Research Methodology 2013 13:148.