J Chron Dis Vol. 31 pp. 539–546

0021-9681/78/0801-539 $02.00/0

SEQUENCE LENGTH FOR REPEATED SCREENING TESTS

RICHARD G. CORNELL

Department of Biostatistics, School of Public Health,
University of Michigan, Ann Arbor, MI 48104, U.S.A.

(*Received in revised form* 20 *October* 1977)

Abstract—The cost of detecting asymptomatic colonic cancer with a sequence of occult blood tests by the guaiac method is investigated to determine the effect of sequence length. The results are sensitive to the definition of a screen positive and a variety of definitions of a positive screen are investigated. It is found that additional tests, when practical, can be used effectively to reduce the average cost of detection per case and to decrease markedly the probability of a false positive screen without sacrificing high sensitivity even when the prevalence of disease is low. The results hold for a variety of combinations of test and prevalence parameters. This strengthens the conclusion and makes the results more generally applicable.

THE COST of detecting asymptomatic colonic cancer with a six-sequential-stool-guaiac protocol was studied by Neuhauser and Lewicki [1]. They found that the marginal cost of the sixth test may be 20,000 times the average cost. In their calculations they estimated the sensitivity and specificity of an individual guaiac test and the prevalence rate for colonic cancer from a fairly small study by Greegor [2], but their findings would have been the same for any parameter specifications of the same order of magnitude. However, their conclusions are sensitive to the definition of a screen positive. They defined a person as screen positive if at least one of a sequence of guaiac tests were positive. Since the prevalence rate was assumed to be low (0.0072) and the sensitivity of the guaiac test to colonic cancer was assumed to be high, a sequence of, say, six guaiac tests with a single positive result was more apt to happen because of a test error than because of the presence of colonic cancer. Thus the probability that a positive screen was false was high, as was the average cost of detection of cancer cases.

In this study a variety of definitions of a positive screen are investigated for a sequence of tests. Protocols for both fixed length and variable length sequences are considered. It is found that additional tests can be used effectively to reduce the average cost of detection per case and to markedly decrease the probability of a false positive screen even when the prevalence of disease is low.

ANALYSIS

Sequences of fixed length

Consider a sequence of independent guaiac tests with results $X_1, X_2, \ldots, X_n$ where $X_i = 1$ if the ith test is positive and $X_i = 0$ if the ith test is negative, $i = 1, \ldots, n$. Let θ

denote the true state of disease with $\theta = 1$ if the colonic cancer is present and $\theta = 0$ otherwise. Assume that $P(X_i = 1|\theta = 1) = \alpha$. This is the sensitivity of an individual guaiac test, and has been assumed to be the same for each test and for all diseased individuals. Assume that $P(X_i = 1|\theta = 0) = \beta$. The specificity of an individual guaiac test is then $1 - \beta$. Assume that $P(\theta = 1) = \lambda$ where λ is the prevalence of colonic cancer. Thus far this is the same model as that considered by Neuhauser and Lewicki.

Let $Y = 1$ denote a positive screen based upon the entire sequence of n tests and let $Y = 0$ denote a negative screen. Let x denote the number of X_i's which equal 1. Neuhauser and Lewicki took Y to be 1 whenever $x \geqslant 1$. We shall first consider sequences of fixed length n where $Y = 1$ when $x \geqslant k$ and $Y = 0$ for $x < k$, for $k = 2, 3, \ldots, n$ as well as 1 and for $n = 1, \ldots, 6$. Sensitivity, specificity and predictive positive and negative probabilities have been calculated. So have marginal and average costs. The costs and number of cases detected are based on probabilities and are therefore expected values. Formulas for the calculation of these probabilities and costs are given in the Appendix. The definitions of marginal and expected costs are the same as those used by Neuhauser and Lewicki. Their cost figures have been used in calculations, namely, overhead expense of \$3 for a sequence of guaiac tests, \$1 additional for each guaiac test, and \$100 for follow-up testing of each screen positive individual with the barium enema procedure.

Tables 1, 2, and 3 give the sensitivity, specificity and average cost, respectively, of a sequence defined by the various pairs of n and k values studied. The results in the first column of each table are for $k = 1$, the setting studied by Neuhauser and Lewicki. The sensitivity probabilities listed in column 1 of Table 1 are listed as true-positive percentages in their Table 1; one minus the specificity probabilities in column 1 of Table 2 are listed as false-positive percentages in their Table 1. The average cost figures in column 1 of Table 3 differ from those listed in the last column of their Table 2 which are in error because of a mistake in the calculation of false-positive cases in the last column of their Table 1. However, the average cost figures in column 1 of Table 3 more than double as n increases from 1 to 6 as did those calculated by Neuhauser and Lewicki and thus these corrected cost figures for $k = 1$ support their conclusions.

The results in Table 3 show that k should be taken to be greater than the value of 1 used in Neuhauser and Lewicki's calculations and particularly that values of $n \geqslant 3$ should be considered. From Table 3 it can be seen that the average cost per case detected is lower for 11 pairs of n and k with $k > 1$ than for any entry with $k = 1$. For nine of those pairs, $n \geqslant 3$. Six pairs of n and k, all with $n \geqslant 3$, led to average cost figures of the order of magnitude of \$2,000, about 1/3 the lowest average cost for $k = 1$. Moreover, the marginal cost of an additional test for fixed k, although not displayed in these Tables, actually decreases as n goes from 5 to 6 for $k = 5$, and is never more than \$6,643 for $k \geqslant 3$, which is similar to the cost per case detected of a single test ($n = 1$, $k = 1$).

Tables 1 and 2 also support the choice of larger values of n and k. From Table 2 it can be seen that the specificity of a sequence of tests is over 0.85 for eight pairs of n and k values, including all entries with $n \geqslant k \geqslant 4$. The highest specificity probability for $k = 1$ is 0.635 when $n = 1$. All of the sensitivity probabilities are $\geqslant 0.99$ for $n \geqslant 2$ when $k = 1$, but this is also true for $k = 2$ when $n \geqslant 4$, $k = 3$ when $n \geqslant 5$ and $k = 4$ when $n = 6$. Sensitivities over 0.94 were also obtained for three other pairs of n and k values, including $k = 4$ when $n = 5$.

TABLE 1. SENSITIVITY OF A SEQUENCE OF n TESTS BY THE NUMBER OF POSITIVE TESTS, k, REQUIRED FOR A POSITIVE SEQUENCE, $\alpha = 0.916667$, $\beta = 0.365079$, $\lambda = 0.0072$

n	1	2	3	4	5	6
1	0.917					
2	0.993	0.840				
3	0.999	0.980	0.770			
4	1.000	0.998	0.963	0.706		
5	1.000	1.000	0.995	0.941	0.647	
6	1.000	1.000	0.999	0.990	0.917	0.593

TABLE 2. SPECIFICITY OF A SEQUENCE OF n TESTS BY THE NUMBER OF POSITIVE TESTS, k, REQUIRED FOR A POSITIVE SEQUENCE, $\alpha = 0.916667$, $\beta = 0.365079$, $\lambda = 0.0072$

n	1	2	3	4	5	6
1	0.635					
2	0.403	0.867				
3	0.256	0.697	0.951			
4	0.163	0.536	0.859	0.982		
5	0.103	0.400	0.741	0.937	0.994	
6	0.066	0.292	0.616	0.866	0.973	0.998

TABLE 3. AVERAGE COST IN DOLLARS PER CASE DETECTED OF A SEQUENCE OF n TESTS BY THE NUMBER OF POSITIVE TESTS, k, REQUIRED FOR A POSITIVE SEQUENCE, $\alpha = 0.916667$, $\beta = 0.365079$, $\lambda = 0.0072$

n	1	2	3	4	5	6
1	6.198					
2	9.087	3.114				
3	11.199	5.205	2.053			
4	12.621	7.842	3.134	1.824		
5	13.577	9.489	4.807	2.201	1.955	
6	14.236	11.119	6.643	3.234	1.870	2.262

The actual choice of a pair of n and k values involves a trade-off between the criterion variables tabulated in Tables 1, 2, and 3. The lowest average cost per case detected in Table 3 occurs for $n = 4$, $k = 4$, but from Table 1 the sensitivity for this n, k pair is only 0.706, which means that the probability of missing a case with this type of sequence would be 0.294. The other five pairs of n and k values with average costs per case detected of the order of magnitude of $2,000 also have sensitivities which are too low. Two of the three n, k pairs with average costs per case detected of the order of magnitude of $3,000 have sensitivities over 0.96 and specificities of about 0.86. One of these pairs, $n = 6$, $k = 4$, describes a sequence of tests with an average cost of $3,234 with a sensitivity of 0.990 and a specificity of 0.866. This represents a loss of sensitivity of slightly less than one percent with at least a 2-fold gain in specificity and well over a 2-fold decrease in average cost per case detected as compared to sequences with $k = 1$ and $n \geq 2$. Thus, for the model and parameter values studies by Neuhauser and Lewicki, the best choice of a sequence length ≤ 6 is 6 where four or more of the sequence of tests should be required to be positive for a patient to be declared screen positive.

One concern still remains about this choice of a sequence of tests as well as about the sequences with $k = 1$ studied by Neuhauser and Lewicki. For all of these tests, the predictive probability that a person who is found to be screen negative is actually

negative, denoted by $((\theta = 0|Y = 0)$, is $\geqslant 0.999$ and is 0.9999 when $n = 6$ and $k = 4$. However, the predictive probability that a person who is found to be screen positive is actually positive, denoted by $P(\theta = 1|Y = 1)$, is never greater than 0.018 for $k = 1$ and is only 0.051 for $n = 6$ and $k = 4$. The next section describes a modification of the $n = 6$, $k = 4$ procedure which yields a much higher value of $P(\theta = 1|Y = 1)$ but is still satisfactory with respect to the criteria described earlier.

A sequential procedure

One way to increase the predictive probability of a positive is to repeat the entire sequence of n guaiac tests whenever $Y_1 = 1$ for the first sequence. The subscript on Y indicates the sequence number and equals 1 or 2; Y without a subscript will still indicate an overall assessment of the screening protocol for an individual. We will take $Y = 1$ (positive) whenever both Y_1 and $Y_2 = 1$. When $Y_1 = 1$, but $Y_2 = 0$, Y will be taken to be 0 (negative). When $Y_1 = 0$, no further testing will be done and Y will be taken to be 0 (negative).

The sensitivity of this sequential procedure is π_1^2, where π_1 is the sensitivity of the corresponding single-sequence procedure described by Table 1. Thus when $n = 6$, $k = 4$ for each sequence, the sensitivity of this sequential procedure is $(0.990)^2 = 0.981$. The specificity can be shown to be $[1 - (1 - \pi_0)^2]$ which is 0.982 for $n = 6, k = 4$, where π_0 is the specificity of the corresponding single-sequence procedure given in Table 2. The average cost, again calculated as the total expected cost over the expected number of cases detected, is \$1748 for this sequential procedure provided no additional overhead expense is incurred. (See the Appendix for all the calculation formulas for this procedure.) Thus this sequential procedure has considerably higher specificity than the corresponding single-sequence procedure (0.982 versus 0.866) and about half as much cost per case detected (\$1748 versus \$3234) at the expense of only 1% decrease in sensitivity (0.980 versus 0.990). The predictive probability of a negative screen is the same (0.9999) to four decimal places for both procedures. However, the predictive probability of a positive screen is higher by more than a factor of 5 for the sequential procedures (0.282 versus 0.051). Thus the sequential procedure would be preferred.

Extensions to n > 6

In the calculations thus far attention has been focused on single sequences with $n \leqslant 6$, the maximum considered by Neuhauser and Lewicki. It would, of course, not be

TABLE 4. SENSITIVITY, SPECIFICITY AND AVERAGE COST PER CASE DETECTED AND PREDICTIVE PROBABILITIES OF NEGATIVE AND POSITIVE SCREENS FOR SINGLE-SEQUENCE AND SEQUENTIAL PROCEDURES FOR $n = 6, 7$ AND 8 AND SELECTED VALUES OF k. $\alpha = 0.916667$, $\beta = 0.365079$, $\lambda = 0.0072$

	Sensitivity	Specificity	Average cost per case detected	Predictive probability of screen	
				Negative	Positive
Single-sequence					
$n = 6, k = 4$	0.990	0.866	3234	0.9999	0.051
$n = 7, k = 4$	0.999	0.775	4603	1.0000	0.031
$n = 8, k = 5$	0.997	0.876	3351	1.0000	0.055
Sequential					
$n = 6, k = 4$	0.981	0.982	1748	0.9999	0.282
$n = 7, k = 4$	0.997	0.949	2421	1.0000	0.125
$n = 8, k = 5$	0.995	0.985	1996	1.0000	0.318

TABLE 5. SENSITIVITY, SPECIFICITY AND AVERAGE COST PER CASE DETECTED AND PREDICTIVE PROBABILITIES OF NEGATIVE AND POSITIVE SCREENS FOR SELECTED PAIRS OF n AND k BY α, β and λ

α, β, λ	n, k	Sensitivity	Specificity	Average cost per case detected	Predictive probability of screen	
					Negative	Positive
0.90,0.10,0.01	7,4	1.00	1.00	1130	1.00	0.79
0.05	5,3	0.99	0.99	278	1.00	0.86
	7,4	1.00	1.00	306	1.00	0.95
0.10	5,3	0.99	0.99	188	1.00	0.93
	7,4	1.00	1.00	203	1.00	0.98
0.50	5,3	0.99	0.99	117	0.99	0.99
	7,4	1.00	1.00	120	1.00	1.00
0.20,0.01	7,4	1.00	0.97	1436	1.00	0.23
0.05	7,4	1.00	0.97	364	1.00	0.61
0.10	7,4	1.00	0.97	230	1.00	0.77
0.50	5,3	0.99	0.94	122	0.99	0.94
	7,4	1.00	0.97	123	1.00	0.97
0.30,0.01	7,4	1.00	0.87	2354	1.00	0.07
0.05	7,4	1.00	0.87	541	1.00	0.29
0.10	7,4	1.00	0.87	314	1.00	0.47
0.50	5,3	0.99	0.84	133	0.99	0.86
	7,4	1.00	0.87	133	1.00	0.89
0.95,0.10,0.01	6,4	1.00	1.00	1015	1.00	0.89
0.05	6,4	1.00	1.00	283	1.00	0.98
	7,4	1.00	1.00	305	1.00	0.95
0.10	6,4	1.00	1.00	191	1.00	0.99
	7,4	1.00	1.00	202	1.00	0.98
0.50	5,3	1.00	0.99	117	1.00	0.99
	6,4	1.00	1.00	118	1.00	1.00
	7,4	1.00	1.00	120	1.00	1.00
0.20,0.01	7,5	1.00	1.00	1150	1.00	0.68
0.05	7,5	1.00	1.00	310	1.00	0.92
0.10	7,5	1.00	1.00	205	1.00	0.96
0.50	6,4	1.00	0.98	120	1.00	0.98
	7,5	1.00	1.00	121	1.00	1.00
0.30,0.01	7,5	1.00	0.97	1390	1.00	0.26
0.05	7,5	1.00	0.97	356	1.00	0.65
0.10	7,5	1.00	0.97	226	1.00	0.79
0.50	6,4	1.00	0.93	125	1.00	0.93
	7,5	1.00	0.97	123	1.00	0.97

practical to require many more stool specimens than this in the absence of symptoms or without a positive indication from the first sequence of tests. However, the fact that $n = 6$ for the best single sequence studied thus far makes slightly longer sequences of interest. Results for $n = 7$ and 8 for both single-sequence and sequential procedures are given in Table 4 for k values which ensure sensitivity over 0.99, specificity as high as possible and low average cost per case detected. Results for $n = 6, k = 4$ are repeated in Table 4 for comparison. Note that on each of the criteria except average cost slight improvement is achieved in going from $n = 6$ to $n = 8$, but the overall impression is that both for single-sequence and sequential plans, the $n = 6, k = 4$ and $n = 8, k = 5$ screens have similar characteristics. However, the preferred screen among those listed

in Table 4 would be the sequential plan with $n = 8$ and $k = 5$ because its sensitivity is slightly higher than the sequential screen with $n = 6$ and $k = 4$.

Extensions to other parameter combinations

Thus far the numerical results in this paper have been for individual tests characterized by $\alpha = 0.916667$ and $\beta = 0.365079$ and for prevalence rate $\lambda = 0.0072$, the parameter values used in Neuhauser and Lewicki's study. Other values of these parameters are of interest both with respect to the robustness of the results to changes in the parameter specifications for guaiac tests for colonic cancer and with respect to possible extensions of the results to different screening tests or higher risk populations. Table 5 presents values of the same variables as those contained in Table 4 for selected n and k pairs for alternative combinations of α, β, and λ. All results in Table 5 are for single-sequence procedures with $n \leqslant 7$. In every instance, n and k were picked to give a sensitivity $\geqslant 0.99$ with relatively low average cost per case detected and specificities and predictive probabilities as high as possible. More than one pair of n and k values are given for some combinations of α, β, and λ when these pairs describe sequences with similar characteristics. Only one pair is presented when it is clearly the best choice. Note that a pair of n and k values with $n = 7$ has been chosen for every combination of α, β, and λ. Also note that the choice of k varies with the parameter combination. For $\alpha = 0.90$ with $n = 7, k = 4$ in every instance in Table 5; for $\alpha = 0.95$ with $n = 7, k = 4$ for $\beta = 0.10$ but $k = 5$ for $\beta = 0.20$ and 0.30.

DISCUSSION

It has been shown that for the screening model for repeated guaiac tests for colonic cancer studied by Neuhauser and Lewicki that large values of n such as 6, 7, and 8 give the most informative screens when the values of k, the number of positives required for a sequence to be positive, are chosen appropriately. In each instance, the best choice is $k \geqslant n/2 > 1$. Moreover, if the entire sequence is repeated when the first sequence is positive with retention of this initial positive designation only when the second sequence is also positive, then an ever better screen protocol is obtained, particularly with respect to the predictive probability of a positive screen.

The results have been found to hold for a variety of combinations of test and prevalence parameters, thus strengthening the conclusion that longer sequences of inexpensive tests are better screens than shorter sequences when the information obtained is used effectively.

Acknowledgement—Support for this research was provided by contract N01-CB-53907 to the Michigan Cancer Foundation from the National Cancer Institute. The calculations were programmed by Paula J. Beitler.

REFERENCES

1. Neuhauser D, Lewicki AM: What do we gain from the sixth stool guaiac? **New Engl J Med** 293:226–228, 1975
2. Greegor DH: Detection of silent colon cancer in routine examination. **CA** 19:330–337, 1969

APPENDIX

Formulas used in the calculations reported in this paper for single sequences of length n are as follows:

Sensitivity $$\pi_1 = P(Y = 1|\theta = 1) = \sum_{x=k}^{n} (n_x)\alpha^x(1-\alpha)^{n-x},$$

Specificity: $\pi_0 = P(Y = 0|\theta = 0) = 1 - \sum_{x=k}^{n} (n_x)\beta^x(1-\beta)^{n-x}$,

Average cost per case detected $= (3+n)/(\lambda\pi_1) + 100\tau$

where $\tau = 1 + [(1-\pi_0)(1-\lambda)]/(\pi_1\lambda)$,

Predictive probability of a screen positive $= P(\theta = 1|Y = 1) = 1/\tau$,

Predictive probability of a screen negative $= P(\theta = 0|Y = 0) = 1/\mu$

where $\mu = 1 + [(1-\pi_1)\lambda]/[\pi_0(1-\lambda)]$.

Formulas for the sequential test presented in this paper are as follows:

Sensitivity: $\rho_1 = \pi_1^2$,

Specificity: $\rho_0 = 1 - (1-\pi_0)^2$,

Average of cost per case detected $= (3+n)/(\lambda\rho_1) + 100v + v\tau\pi_1/\rho_1$

where $v = 1 + [(1-\rho_0)(1-\lambda)]/(\rho_1\lambda)$,

Predictive probability of a screen positive $= 1/v$,

Predictive probability of a screen negative $= 1/\omega$

where $\omega = 1 + [(1-\rho_1)\lambda]/[\rho_0(1-\lambda)]$.