RBE. VOL.6 N2. 1989

A MAXIMUM ENTROPY APPROACH TO CORONARY DISEASE DIAGNOSTIC USING STRESS TEST AND ST SEGMENT DEPRESSION

H. Palacios Fagerström, D. Rodríguez Schulz[1]

ABSTRACT -- The proper management of entropy, or average information content, can be used for increasing the power of a test. In this paper the ST segment depression of the electrocardiogram, associated to stress test, is used as an information source in an approach to determine the probability of coronary disease. After determining a noisy channel entropy equation, expressed in medical terminology, it can be concluded that the simple ON/OFF, 1 mm standard medical criteria can get only the 60% of the easily available information contents of the test. In order to get the remaining information, a q ary criteria should be used, with q=4, q=7 or q=13, instead of the q=2 standard binary criteria. The importance of this approach resides in the fact that in several cases a more involved, invasive, risky and high cost test, like coronary angiography, can be avoided and, in general terms, a better information about coronary disease can be obtained. The above criteria has been validated in more than 110 cases.

INTRODUCTION

Non invasive diagnostics have been increasing their popularity during last years and they are unanimously considered as the most desirable way to get a diagnostic [Patterson, 1984]; unfortunately the information contents of several non invasive tests are rather limited and often it is unavoidable to continue with invasive but more powerful test. Nevertheless, if a medical test is considered as an experiment, dedicated to increase the information about some subject, then it is possible to obtain a mathematical maximum for the entropy and, surprisingly, in several cases the standard procedure for the test is not the best for getting all the available information. In the ST segment depression, associated to stress test, the above mentioned concepts are fully applicable.

CONCEPTUAL BASIS

The maximum entropy approach is inserted in the following sequence of events,

- first, there must be a subject whose actual state is unknown and an a-priori probability of any of the possible states of the subject is known.

1 Dpto. Electrónica, Universidad Santa María, Valparaíso - CHILE

- second, to increase the information about the state of the subject, a sequence of experiments must be accomplished, where each experiment must increase the initial knowledge.

- third, when all the information, the initial and the contributions of all the experiments, is equal to or greater than the initial uncertainty, this latter is elucidated and the state of the subject is known.

- The maximum entropy approach is simply to make that each experiment, or test, gives the maximum possible information about the state of the subject, in which case the number of experiments is a minimum.

The above mentioned concepts can be applied to ST segment depression, STD, in the following way, coronary disease can be considered a q-ary subject, with q possible states or degrees of illness; the standard medical criteria consider coronary disease as a binary subject, the disease exists when at least one of the four major coronary arteries is obstructed in at least 50%; the other state is the complement of the just mentioned case [Diamond, 1979].

The ST segment depression can also be considered as a q-ary subject, a response of coronary disease induced by the stress test experiment; in this case, as in the above mentioned, also the standard medical practice is binary, equal or greater than 1 mm, 0.1 mV, means positive result and less than 1 mm is negative. As will be seen soon a much better criteria is obtained when $q = 4$, $q = 7$ or $q = 13$, where near 70% of more information is obtained over the same test.

It is possible to increase q to a larger number, say $q = 25$, but there are three reasons for not to do so, the first is the lack of reliable statistical data relating ST segment depression and coronary angiography, being the latter a perfect reference for the former; the second reason is the saturation of the information contents v/s q, so perhaps $q = 13$ is the best cost/performance trade off; the third is that a resolution of 0.025 mV is a practical limit for a standard electrocardiograph even if double sensibility is used.

MATHEMATICAL AND STATISTICAL BASIS

A noisy channel entropy equation, where i, the source, is binary and j, the destination, is q ary, can be determined as follows [Shannon, 1948]

$$H = \Sigma \Sigma p(i, j) \lg \frac{p(i,j)}{p(i)p(j)} \quad (1)$$

H : Entropy; i = states of source;

j : states of destination; $lg \equiv log_2$.

Introducing Bayes' rules

$$p(i,j) = p(j)\ p(i/j) = p(i)\ p(j/i) \qquad (2)$$

And reordering, the classical i-ary source j-ary destination noisy channel entropy can be obtained.

$$H = - \sum_i p(i) lg\ p(i) + \sum_j p(j) \{\sum_i p(i/j) lg\ p(i/j)\} \qquad (3)$$

If the following relations are introduced in (3)

$$p(j) = \Sigma\ p(i)\ p(j/i) \qquad (4)$$

$$p(i/j) = p(i)\ p(j/i) \qquad (5)$$

And renaming i, the source, as the binary variable E, coronary disease, where E+ is the assertion and E- the negation, and also renaming j as Dq, diagnostic, being Dq a q-ary variable, where q can be 2, 4, 7, 13 and some other discrete values, a final expression for entropy can be obtained.

$$\begin{aligned} H = &- \{ p(E+)\ lg\ p(E+) + p(E-)\ lg\ p(E-) \} \\ &+ \sum_q \{ p(E+)\ p(Dq/E+)\ lg\ p(E+)\ p(Dq/E+) \\ &\quad + p(E-)\ p(Dq/E-)\ lg\ p(E-)\ p(Dq/E-) \\ &\quad - [\ p(E+)\ p(Dq/E+) + p(E-)\ p(Dq/E-)] * \\ &\quad lg\ [\ p(E+)\ p(Dq/E+) + p(E-)\ p(Dq/E-)\ ] \} \end{aligned} \qquad (6)$$

E+ coronary disease, E- non coronary disease.

This equation represent the noisy channel binary source q-ary destination mentioned above. The first row is directly related to the a-priori information contents, prevalence in medical terms. The second and following rows are related to the result of the actual experiment or test.

The statistical data available follows; the coronary angiography has been considered as a reference i.e. maximal information test.

1. Prevalence

Pain Symptom	p(E+)	p(E-)
Typical angina	0.899	0.101
Atypical angina	0.534	0.466
Non anginal pain	0.190	0.810
No pain	0.065	0.935

Averages valid for men from 40 to 60 years old [Patterson, 1984].

2. Conditional probabilities q = 13

p(Dq/E+)	p(Dq/E-)	STD* ≥ mm
.930	.535	0.25
.875	.375	0.50
.784	.239	0.75
.649	.148	1.00
.515	.091	1.25
.416	.038	1.50
.376	.019	1.75
.328	.017	2.00
.245	.012	2.25
.125	.005	2.50
.131	.004	2.75
.050	.003	3.00

* STD: ST Segment depression, equal or greater than the value indicated, in millimeters.

From this table, any q less than 13 table can be obtained. As an example, q = 7 conditional probability table follows.

3. Conditional probabilities q = 7

p(Dq/E+)	p(Dq/E-)	STD ≥ mm
.875	.375	0.5
.649	.148	1.0
.416	.038	1.5
.328	.017	2.0
.195	.005	2.5
.050	.003	3.0

4. Conditional probabilities q = 4

p(Dq/E+)	p(Dq/E-)	STD ≥ mm
.649	.148	1.0
.328	.017	2.0
.050	.003	3.0

The selected values of q are related to electrocardiogram amplitude resolution, for q = 4 resolution is 1 mm, 0.1 mV, for q = 7 resolution is 0.5 mm and for q = 13 resolution must be 0.25 mm, the practical higher limit for a standard electrocardiograph with double sensibility. The next step could be q = 25, out of range for standard devices. The source of information for the tables is in [Weiner, 79]. [Diamond, 79] and [Patterson, 84].

The above tables consider total probability for each value of STD; it is now necessary to get probabilities by range, so the STD obtained during stress test is allocated to a compartment defined by a range of STD values; each compartment is associated to conditional probabilities of having or not coronary disease. It is obvious that while smaller each compartment more information provides. The following table is directly derived from the number 2. above.

5. Conditional probabilities by range, q = 13

p(Dq/E+)	p(Dq/E-)	< STD ≤	mm
.070	.465	0.00	0.25
.055	.160	0.25	0.50
.091	.136	0.50	0.75
.135	.091	0.75	1.00
.134	.057	1.00	1.25
.099	.053	1.25	1.50
.004	.019	1.50	1.75
.048	.002	1.75	2.00
.083	.005	2.00	2.25
.050	.007	2.25	2.50
.064	.001	2.50	2.75
.081	.001	2.75	3.00
.050	.003	3.00	∞

Similar tables can be obtained for q = 7 and q = 4.

After the stress test, the ST segment depression obtained determines the entry point to the table in use, say q = 13 conditional probabilities by range. From the table, p(Dq/E+) and p(Dq/E-) are obtained.

A final coronary disease probability can be obtained using Bayes' rule

$$p(E+/D) = \frac{p(E+)\ p(D/E+)}{p(E+)\ p(D/E+) + p(E-)\ p(D/E-)} \qquad (4)$$

It is now possible to plot H v/s q. The peak, ideal value of H is one and it corresponds to p(E+) = 0.5, as any standard binary entropy. Coronary angiography performs very near of the optimum, so for E+ = 0.5 H is almost one. The stress test ST segment depression performs as follows

- for q = 2, H = 0.198 bit, almost 20% of the maximum
- for q = 13, H = 0.330 bit, this is 33% of the maximum and 67% more information than binary case
- for q = 7, H = 0.277 bit
- for q = 4, H = 0.228 bit

PRACTICAL RESULTS

A population of 112 men, with ages between 40 and 60 years old, were analyzed, all of them with stress test and coronary angiography reported.

The Bayes' probability of coronary disease was sectorized in 3 ranges

- High risk, $p(E+/D) \geq 0.8$

- Medium or uncertain risk

  $0,2 \leq p(E+/D) < 0.8$

- Low risk

  $p(E+/D) < 0.2$

The higher uncertainty is related to atypical angina, where $p(E+)$ is near 0.5; in this case 12 patients with negative binary STD, but shifted to uncertain risk under $q = 13$ criteria, were reported positive in coronary angiography.

In the case of asymptomatic prevalence, 5 patients with positive binary STD were shifted to low risk and the angiographic test reported negative, so the latter test could have been avoided.

In the case of typical angina, under $q = 13$ criteria, there is no way that $p(E+/D)$ can go below 0.2, so a confusing negative binary STD would be avoided.

In the case of non-anginal pain, a positive binary STD always give a $p(E+/D) \geq 0.2$, so $q = 13$ criteria is coincident with the standard medical criteria.

CONCLUSIONS

- The use of $q = 7$, or $q = 13$ criteria, associated with an a-posteriori probability of disease, presents solid advantages over ON/OFF 1 mm criteria, due to the fact that it is also simple to use and it has a better correlation with coronary angiography.

- In atypical angina, a significative number of patients with negative 1 mm diagnostic can be considered as in the uncertain risk zone, and the angiography reported positive.

- In asyntomatic cases, again a significative number of patients with positive 1 mm diagnostic, can be shifted to low risk, which was confirmed by negative angiographic test, so, for these patients, the invasive angiographic test could have been avoided.

- One practical disadvantage resides in the fact that good quality electrocardiograms are needed in order to permit the application of q-ary criteria, specially when q = 13, where double sensibility must be used.

ACKNOWLEDGMENT

To cardiology division personnel of "Hospital Dr. G. Fricke", Viña del Mar, Chile, specially to Jorge Bartolucci J., MD. and Pedro Chadid S. M.D., by their encouragement and help.

This work has been finantially supported by "Fondo Nacional de Desarrollo Científico y Tecnológico", Chile.

REFERENCES

A-HAMID HAKKI, DE PACE N., COLBY J. (1983), "Implications of normal exercise electrocardiographic results in patients with angiographically documented coronary artery disease". The American Journal of Medicine, Volume 75, pages 439-443.

DIAMOND B. A., FORRESTER J. S. (1979), "Analysis of probability as an aid to the clinical diagnosis of coronary artery disease", New England Journal of Medicine, volume 300, page 1350.

OKADA M. (1978), "A method for clinical data reduction based on weighted entropy". IEEE Transactions on Biomedical Engineering, volume BME-25, page 462.

PALACIOS H. (1985), "Aplicación de la teoría de información en el diagnóstico clínico". IX Congreso Brasileño de Ingeniería Biomédica, Campinas, Brasil, Sept. 1985. Curso Tutorial, profesor invitado.

PATTERSON R., C. ENG., S. HOROWITZ (1984), "Practical, diagnosis of coronary artery disease: A Bayes's theorem nomogram to correlate clinical data with noninvasive exercise test". The American Journal of Cardiology, volume 53, pages 252-256.

SANTASH K., ROITMAN D., BRADLEY L. (1983), "Enhanced evaluation of tread mill test by mean of scoring based on multivariate analysis". The American Journal of Cardiology, volume 52, pages 1155-1160.

SHANNON C. E. (1948), "A mathematical theory of communications", Bell System Technical Journal. July and October.

VATTEROTT P., HAMMILL S. (1988), "Signal averaged electrocardiography: A new non invasive test". Mayo Clin. Proc. September, Vol 63.

WEINER D. A., RYAN T. J. (1979), "Exercise stress testing. Correlation among history of angine, ST segment response and prevalence of coronary artery disease in CASS". New England Journal of Medicine, volume 301, page 230.

WEINTRAUB W., BARR-ALDERFER V., SEELAUS P. (1985), "A sequential approach to the diagnosis of coronary artery disease using multivariate analysis". American Heart Journal, volume 109, page 999.