How the Threshold "R" Influences Approximate Entropy Analysis of Heart-Rate Variability

P Castiglioni, M Di Rienzo

Biomedical Technology Dept, Don Gnocchi Foundation, Milan, Italy

Abstract

Calculation of approximate entropy (ApEn) requires to select the correct threshold "r". Previous studies recommended r to be between 0.1 and 0.25 times the signal standard deviation, and now r=0.2 is used in almost all HRV studies. Recently it has been claimed that for fast signal dynamics, r=0.2 may lead to erroneous conclusions, while r maximizing ApEn, r_{MAX}, correctly assesses entropy. We verified 1) if r_{MAX} differs from r=0.2 also for HR; and 2) if all r values in the 0.1-0.25 range provide similar ApEn measures. For this aim, we recorded R-R intervals in 10 young subjects for 10', in supine and sitting positions, and calculated ApEn(r) for r between 0.02 and 1.20, identifying r_{MAX} and ApEn(r_{MAX}). r_{MAX} felt into the recommended range, but it significantly differed from 0.2. At the extremes of the range, the effects of posture change on ApEn were even opposite: ApEn(0.25) decreased while ApEn(0.1) increased. Therefore the choice of r is critical even in HRV studies.

1. Introduction

Lack of regularity in physiological time series is often quantified by computing the approximate entropy, ApEn. This index can be efficiently evaluated even over relatively short time series, making it particularly suitable for the analysis of physiological signals. ApEn is related to the probability that segments of "*m*" data samples which are similar (i.e., closer each other than a given distance "*r*") remain similar when the segment length increases to "*m+1*" [1-3]. Lower is this probability (and thus the predictability of the time series), greater is ApEn.

In order to calculate ApEn, one has to preliminary fix the values of the parameters "*m*" (the embedding dimension) and "*r*" (the threshold tolerance). Previous studies, based on the analysis of deterministic and stochastic processes, suggested to select *m*=2 and *r* in the range between 0.1 and 0.25 times the standard deviation of the time series [1,4], and actually ApEn is now calculated with *m*=2 and *r*=0.2 in almost all HRV studies.

Very recently, however, it has been claimed that when the signal dynamics is faster than the heart rate dynamics (as for neural signals), the recommended *r* values may lead to erroneous conclusions [5,6]. These authors alternatively suggest to use the *r* value which maximizes ApEn, r_{MAX}. The value of approximate entropy obtained by setting $r=r_{MAX}$, ApEn(r_{MAX}), quantifies the highest information difference between segments of length *m* and *m*+1, and therefore the choice of $r=r_{MAX}$ should allow to take into account more of the signal complexity than other values of *r*. Moreover, according to these authors, this is a less arbitrary choice than selecting one of the *r* values between 0.1 and 0.25, the recommended range.

Aim of this work is to verify: 1) whether r_{MAX} differs significantly from *r*=0.2 (value used in most HRV studies) also for the slower heart rate dynamics; and 2) whether all the *r* values within the recommended range (from 0.1 to 0.25) provide similar and consistent measures of regularity.

2. Methods

Data Collection. We enrolled 10 young healthy volunteers (age 21-25 yrs). We recorded the ECG in a quite room for 10 minutes twice: in the supine and in the sitting position. We selected these two conditions because we expect them to be characterized by different levels of entropy. In fact, these two conditions are known to be associated with different autonomic balances: supine at rest is a baseline condition where the vagal tone is high and the sympathetic tone is almost completely disengaged. By contrast, sitting at rest is characterized by a mild sympathetic activation, compared with supine rest, due to the blood redistribution associated with the posture change. Recordings in supine and sitting positions were preceded by an adaptation period to allow for stabilization of heart rate after the posture change.

Data Analysis. ECG was digitized (200 Hz sampling rate, 12 bits resolution). To derive the R-R interval (RRI) series from the ECG we identified each R peak by means of a detection algorithm based on the ECG band-pass filtering (to extract the QRS complex), differentiation, squaring and comparison with an adaptive threshold (7).

For each subject and for each of the two experimental conditions, we considered a time series of *N*=600 consecutive RRI values; we calculated the standard

deviation, SD, of RRI and evaluated ApEn setting m=2 and with r increasing from 0.02 to 1.20, with step 0.02.

For a given r value, ApEn(r) was calculated as follows. First we set m=2 and from the RRI series of N=600 beats, $\{RR(i)\}_{i=1,\ldots,N}$, we created the series of N-m+1 vectors of m components $R_m(i)=[RR(i), RR(i+1),\ldots, RR(i+m)]^T$. The vector $R_m(i)$ represents the sequence of m consecutive RRI values starting at the beat i. Two vectors $R_m(i)$ and $R_m(j)$ are similar if the absolute difference between each couple of corresponding scalar components is less than the distance $r\times$SD.

Calling $n_i^m(r)$ the number of N-m+1 vectors $R_m(j)$ which are similar to $R_m(i)$, then

$C_i^m(r)= n_i^m(r)/(N-m+1)$

is the probability to find a sequence of m beats similar to the sequence represented by $R_m(i)$, and $C^m(r)$, defined as the mean of all $C_i^m(r)$, quantifies the prevalence of similar strings of m beats. ApEn(r) is calculated as:

$ApEn(r)=-ln[C^{m+1}(r)/ C^m(r)]$

A high degree of regularity means that sequences which are similar for m points are likely to be similar also for the next $m+1$ point, while this is unlikely to occur for irregular time series. Thus low values of ApEn reflect high regularity.

To find r_{MAX} and ApEn(r_{MAX}), we considered the highest value of approximate entropy in the $0.02\le r\le 1.20$ range, ApEn(r_K), and interpolated it and the preceding and following values, ApEn(r_{K-1}) and ApEn(r_{K+1}), with a parabola. Position and value of the vertex of the parabola give r_{MAX} and ApEn(r_{MAX}).

The effects of changing posture from supine to sitting were quantified by computing the corresponding percent changes of ApEn, ΔApEn(r), separately at r=r_{MAX}, at r=0.2, and at the extremes of the recommended range: r=0.1 and r =0.25.

Statistics. We used non-parametric tests (8), taking p=0.05 as the level of statistical significance. We evaluated the significance of the difference between r_{MAX} and 0.2 by the sign test; and between ApEn(r_{MAX}) and ApEn(0.2), by the Wilcoxon matched pairs test. These tests were assessed separately in supine and sitting conditions. The Wilcoxon matched pairs test was also used to compare ΔApEn(r_{MAX}) with ΔApEn(0.2), and ΔApEn(0.1) with ΔApEn(0.25).

3. Results

Individual ApEn(r) functions in supine condition are shown in figure 1. In all subjects, ApEn(r) quickly increases up to a maximum, which falls within the recommended r range; then ApEn(r) monotonically decreases more slowly when r is greater than r_{MAX}. Median (1^{st} – 3^{th} quartiles) of r_{MAX} in supine condition is 0.17 (0.16 - 0.18), differing significantly (p=0.03) from r=0.20, the value employed in most HRV studies.

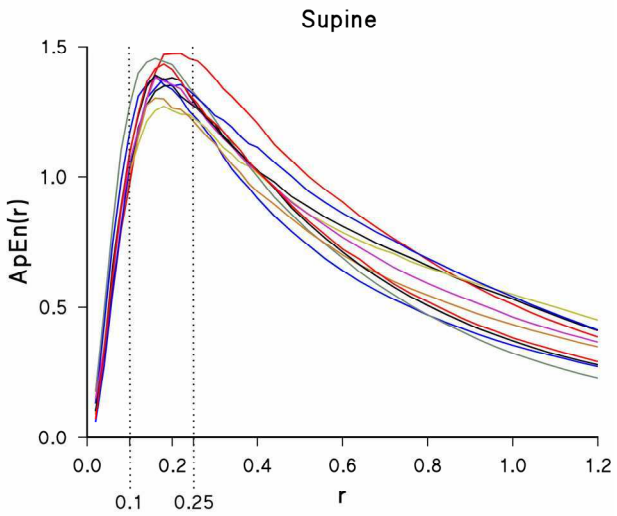

Figure 1. Individual ApEn(r) functions in 10 healthy volunteers during supine rest; the two dotted vertical lines delimitate the recommended r range.

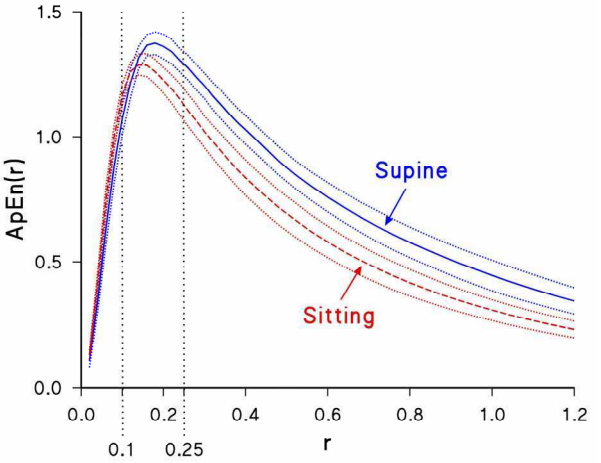

Figure 2. Mean and 95% confidence intervals of ApEn(r) evaluated over the group of N=10 healthy volunteers in the supine (continuous line) and sitting (dashed line) positions.

Mean values and 95% confidence intervals over the group are shown in figure 2 separately for supine and sitting conditions. In both conditions the ApEn maximum falls within the recommended range. However, it appears shifted at lower r values in the sitting with respect to the supine condition. In sitting condition, the median (1^{st}-3^{th} quartile) of r_{MAX} is 0.14 (0.13-0.15), and r_{MAX} is even more significantly lower (p=0.004) than the traditionally selected r=0.20 value. Differences in the shape of the two

ApEn(r) functions are not limited to a shift in the position of r_{MAX}: in fact, ApEn(r) is lower in the sitting with respect to the supine position when r>0.14, while the opposite is true for r<0.12.

Table 1. ApEn(r) values at r= r_{MAX} and at r=0.2: values are shown as median (1st-3th quartiles)

	ApEn(r_{MAX})	**ApEn(0.2)**	**p**
Supine	1.39 (1.36-1.42)	1.35 (1.34-1.40)	0.005
Sitting	1.30 (1.26-1.35)	1.22 (1.18-1.28)	0.005

p is the significance of the difference between ApEn(r_{MAX}) and ApEn(0.2).

Table 1 compares ApEn calculated at r_{MAX} and at 0.2. While it reports only a relatively small discrepancy in the supine position (≈3%), a more substantial difference appears in the sitting condition (≈7%).

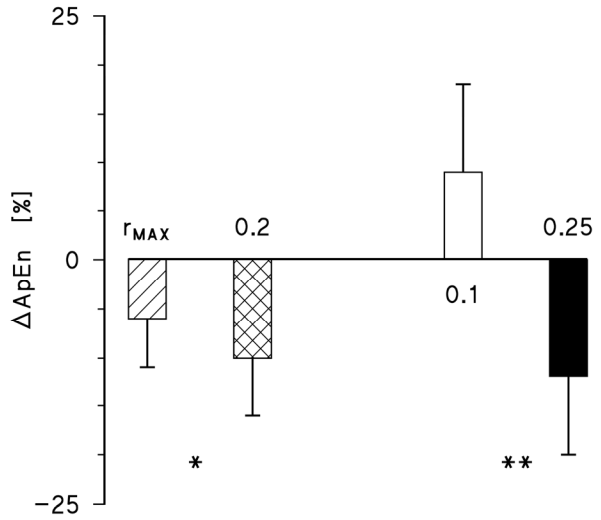

Figure 3. Percent change of ApEn(r) from supine to sitting, at r= r_{MAX} and 0.2, and at the two extremes of the recommended r values: r=0.1 and 0.25; mean +SD over the N=10 volunteers. The * and ** indicate significant differences between ΔApEn(r_{MAX}) and ΔApEn(0.2) and between ΔApEn(0.1) and ΔApEn(0.25), at the 0.05 and 0.01 levels.

Percent changes of ApEn from supine to sitting conditions are shown in figure 3. Approximate Entropy decreases with the change of posture both when it is estimated at r_{MAX} and at 0.2; however, the decrease at 0.2 is almost twice (-10%) the value obtained at r_{MAX} (-6%). Even more striking differences appear comparing the changes of approximate entropy assessed at the two extremes of the recommended r value: in fact, ApEn decreases by about 12% at r=0.25, while it increases by 9% at r=0.1.

4. Discussion and conclusions

The first works which introduced the calculation of approximated entropy for assessing the regularity of physiological signals recommended to select r within the 0.1-0.25 range [1,4]; in the following years, this suggestion was followed in all papers which applied ApEn on the heart rate dynamics, most of them selecting r=0.2. Recently, however, the analysis of signals with faster dynamics provided evidence that the recommended range may not be appropriate, and it has alternatively suggested to select the r which maximize ApEn [5,6].

First result of our study is that, as far as the heart rate dynamics is considered, the selection of r_{MAX} is not incompatible with the traditionally recommended range, because in all of our subjects, both in supine and sitting conditions, r_{MAX} felt between 0.1 and 0.25.

However, this does not mean that any r value within the recommended range could be indifferently selected for evaluating ApEn. In fact we showed that even a simple manoeuvre which is expected to slightly change the heart rate entropy, like a change of posture from supine to sitting, does not have simple effects on ApEn. Actually we found that this change of posture decreases ApEn, if large r values are used, while it increases ApEn, for small r. More importantly, the transition from an increase to a decrease of ApEn exactly occurs within the recommended range of r, so that arbitrary selections of the r parameter may lead to opposite results.

In conclusion we showed that the criteria for selecting r should be critically revised even for heart rate studies. Selecting the r value which maximizes ApEn seems a reasonable approach, because this choice would allow to quantify more of the time series irregularity than any other choice of r. In any case, whatever will be the criterion for selecting r, we would recommend to preliminary quantify the whole ApEn(r) profile to get a more complete picture of the phenomenon and verify how critical could be the choice of r in the quantification of ApEn.

References

[1] Pincus SM, Gladstone IM, Ehrenkranz RA. A regularity statistic for medical data analysis. J Clin Monit 1991 Oct;7(4):335-45.

[2] Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA 1991 Mar 15;88(6):2297-301.

[3] Pincus SM, Goldberger AL. Physiological time-series

analysis: what does regularity quantify? Am J Physiol 1994 Apr;266(4 Pt 2):H1643-H1656.
[4] Pincus SM, Keefe DL. Quantification of hormone pulsatility via an approximate entropy algorithm. Am J Physiol 1992 May;262(5 Pt 1):E741-E754.
[5] Lu S, Chen X, Kanters JK, Solomon IC, Chon KH. Automatic selection of the threshold value R for approximate entropy. IEEE Trans Biomed Eng 2008 Aug;55(8):1966-72.
[6] Chen X, Solomon IC, Chon KH. Parameter selection criteria in approximate entropy and sample entropy with application to neural respiratory signals. Am. J. Physiol. Regul. Integr. Comp. Physiol., to be published.
[7] Kohler BU, Hennig C, Orglmeister R, The principles of software QRS detection. IEEE Eng Med Biol Mag 2002 Jan;, 21(1):42-57.
[8] Krauth J. Distribution-free statistics. An application-oriented approach. Amsterdam: Elsevier; 1988.

Address for correspondence

Paolo Castiglioni
Polo Tecnologico, Fondazione Don C.Gnocchi
Via Capecelatro 66,
I 20148 Milano, Italy
pcastiglioni@dongnocchi.it