Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices

Douglas E. Lake and J. Randall Moorman

Am J Physiol Heart Circ Physiol 300:H319-H325, 2011. First published 29 October 2010;
doi:10.1152/ajpheart.00561.2010

Am J Physiol Heart Circ Physiol 300: H319–H325, 2011.
First published October 29, 2010; doi:10.1152/ajpheart.00561.2010.

Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices

Douglas E. Lake and J. Randall Moorman

Cardiovascular Division, Department of Internal Medicine, and Cardiovascular Research Center, University of Virginia Health System, Charlottesville, Virginia

Submitted 11 June 2010; accepted in final form 22 October 2010

Lake DE, Moorman JR. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am J Physiol Heart Circ Physiol* 300: H319–H325, 2011. First published October 29, 2010; doi:10.1152/ajpheart.00561.2010.—Entropy estimation is useful but difficult in short time series. For example, automated detection of atrial fibrillation (AF) in very short heart beat interval time series would be useful in patients with cardiac implantable electronic devices that record only from the ventricle. Such devices require efficient algorithms, and the clinical situation demands accuracy. Toward these ends, we optimized the sample entropy measure, which reports the probability that short templates will match with others within the series. We developed general methods for the rational selection of the template length m and the tolerance matching r. The major innovation was to allow r to vary so that sufficient matches are found for confident entropy estimation, with conversion of the final probability to a density by dividing by the matching region volume, $2r^m$. The optimized sample entropy estimate and the mean heart beat interval each contributed to accurate detection of AF in as few as 12 heartbeats. The final algorithm, called the coefficient of sample entropy (COSEn), was developed using the canonical MIT-BIH database and validated in a new and much larger set of consecutive Holter monitor recordings from the University of Virginia. In patients over the age of 40 yr old, COSEn has high degrees of accuracy in distinguishing AF from normal sinus rhythm in 12-beat calculations performed hourly. The most common errors are atrial or ventricular ectopy, which increase entropy despite sinus rhythm, and atrial flutter, which can have low or high entropy states depending on dynamics of atrioventricular conduction.

heart rate; heart rate variability; statistical analysis

PATIENTS WITH REDUCED CARDIAC FUNCTION receive cardiac implantable electronic devices (CIEDs), such as implantable cardioverter-defibrillators, to prevent sudden death due to ventricular tachycardia or ventricular fibrillation, but they also commonly develop atrial fibrillation (AF). This irregular rhythm, which often develops or is recognized only after device implantation, can lead to stroke and other clinical deteriorations and often mandates new therapies with anticoagulation or drugs to control the heart rate or to restore sinus rhythm. Diagnosing AF using a single-lead CIED (from which only ventricular electrographic recordings and RR intervals, the times between heartbeats, are available) is an important clinical goal but is not currently available, perhaps because the limited processing capacity of these devices puts a premium on efficient detection. For this reason, we studied the clinical problem of AF detection using very short RR interval time series; we felt that diagnosis using only 12 beats was a clinical imperative.

The hallmark of AF is its irregularity. The nonsensical descriptor, "irregularly irregular," that clinicians use underscores this fundamental difference from normal sinus rhythm, and we expect RR interval time series in AF to have higher entropy than ventricular tachycardia or sinus rhythm. In this context, the meaning of entropy follows the work of Shannon (21), Kolmogorov (8, 9), Sinai (22), Grassberger and Procaccia (4), Eckmann and Ruelle (2), and others, who conceived of entropy as a measure of the degree to which template patterns repeat themselves. Repeated patterns imply order and lead to reduced values of entropy. Estimates of entropy, such as sample entropy (SampEn), rely on counts of m-long templates matching within a tolerance r that also match at the next point (12, 19, 20) and have found utility in predicting infection and death in premature infants (5–7). More formally, entropy is the negative natural logarithm of the conditional probability that any two sequences of length m that match within tolerance r will also match at the $m + 1$st point. Counting the number of times that templates find matches is the central activity of entropy estimation, and the result is a ratio: the number of matches of length $m + 1$ divided by the number of matches of length m. More matches means more confident estimation of this ratio and, up to a point, better entropy estimation.

In long heart rate records, when matches abound, entropy measures distinguish AF well from sinus rhythm (1). There is a challenge, though, in assuring a sufficient number of matches when the data sets are short. Thus, for the rapid diagnosis of AF using entropy estimation, the selection of the parameters m and r is critically important. If m is too large or r is too small, then the number of template matches will be too small for confident estimation of the conditional probability. If, on the other hand, m is too small and r is too large, then all templates will match each other, and there will be no discrimination among rhythms. We and others (12, 17) have suggested strategies such as picking m based on the autocorrelation function and picking r based on minimizing the error of the entropy estimates.

Of these, the larger problem in implementing entropy estimation is picking the value of the tolerance r. The original recipe has been to select r as 20% of the SD of each time series segment, based on the "preliminary" conclusions of Pincus (14) in 1991 for implementing approximate entropy (ApEn) calculations. Systematic approaches to picking r have been presented and are usually based on the analysis of relative errors in large data sets (12, 17). An important new insight, though, was presented by Lake (11) in 2006, who approached the problem from the standpoint of stochastic processes and applied concepts of probability density estimation. The direct

Address for reprint requests and other correspondence: D. E. Lake, Univ. of Virginia Health Science Center, Box 800158, Charlottesville, VA 22908 (e-mail: dlake@virginia.edu).

result was to convert the measured conditional probability to a density by normalizing to the volume of the matching region, or $2r^m$, an operation that reduces to adding $\ln(2r)$ to the entropy estimate. The result, called the quadratic sample entropy (QSE) (11), allows any r to be used for any time series and the results compared with any other estimate. This approach frees the investigator to vary r as needed to achieve confident estimates of the conditional probability (CP) and is an important component of our new approach to AF detection.

These considerations led us to develop an entropy parameter that is optimized for the rapid detection of AF with rapid heart rates. We call it the coefficient of sample entropy (COSEn). The major improvements include flexibility in the choice of r and the use of the heart rate itself in the calculation. We evaluated this new measure with a clinician's eye, in particular with regard to records with atrial flutter (AFL). This rhythm can produce very regular RR intervals when in a stable n:1 conduction regime, but is clinically treated much like AF. Thus, we categorized AFL as AF, even though its regular and low-entropy RR interval time series are very different from irregular, high-entropy AF. Throughout, we analyzed AF and AFL together, although the diagnostic performance of any algorithm would improve if we excluded AFL. We note that the episode-by-episode comparison algorithms specified by the current American National Standard for ambulatory ECG analyzers (ANSI/AAMI EC38:1998) do not include AFL segments in the evaluation process, emphasizing the difficulty of diagnosing atrial rhythms from ventricular beats. We felt that our approach more closely mimicked real-world management problems.

MATERIALS AND METHODS

RR interval databases. We studied RR intervals in the MIT-BIH AF database available at Physionet (www.physionet.org). It consists of 10-h recordings from 25 patients with AF with 1,221,578 intervals, of which 519,815 (42.5%) are labeled as AF. It and the other Physionet databases have been widely used in the study of AF and other aspects of heart rate variability. SampEn and multiscale entropy (MSE) have been used to discriminate AF in long records (n = 20,000) of the MIT-BIH databases (2).

In addition, we analyzed 1,461 24-h RR interval time series from Holter monitor recordings from the University of Virginia (UVa) Health System Heart Station over the period of 2/2005 to 5/2008 that were ordered for clinical reasons by UVa physicians. These were manually overread for the presence of AF, with the correction of beat labels as necessary. For each recording, the entire RR interval time series was inspected with rhythm verification from the ECG every 5 min. Since the incidence of AF is very low before the age of 40 yr old (3), we report on 940 Holters from patients (480 men) over this age. We labeled a 12-beat segment as AF if it contained a single AF beat. This approach lowers performance estimates but is clinically justified. The Institutional Review Board of UVa approved the study.

Study design. We used the MIT-BIH dataset for algorithm development and threshold determination. We then validated the algorithm in the UVa data set.

Statistical analysis. We used logistic regression to make multivariable models with different subsets of predictor variables (e.g., entropy, variability, and mean RR interval) and evaluated their ability to detect AF by measuring the receiver operating characteristic (ROC) curve area, which evaluates all possible threshold cutoffs. The significance of models and coefficients was evaluated using the Wald χ^2-statistic adjusted for repeated measures (5). We compared the models and developed the COSEn algorithm as an optimal method for detecting AF.

SampEn estimation. SampEn can be conceived as the CP that two short templates that match within an arbitrary tolerance will continue to match at the next point. A central idea for this work is that SampEn, like any probability, is estimated more accurately when more events are counted.

A data record consists of a series of N consecutive interbeat (RR) intervals, $x_1, x_2, \ldots, x_n$, where the record may be as short as $N = 12$. For a length $m < n$ and starting point i, the template $x_m(i)$ is the vector containing the m consecutive intervals $x_i, x_{i+1}, \ldots, x_{i+m-1}$. For a matching tolerance $r > 0$, an instance where all the components of $x_m(i)$ are within a distance r of any other $x_m(j)$ in the record, is called a match (or template match). For example, the template $x_2(1)$ matches $x_2(3)$ if both $|x_1 - x_3| < r$ and $|x_2 - x_4| < r$. Let B_i denote the number of matches of length m with template $x_m(i)$ and A_i denote the number of matches of length $m + 1$ with template $x_{m+1}(i)$.

Let $A = \Sigma A_i$ and $B = \Sigma B_i$ denote, respectively, the total number of matches of length $m + 1$ and m. The ratio $p = A/B$ is then the CP that subsequent points of a set of closely matching m intervals also remain close and match. SampEn [or SampEn(m,r) to indicate its dependency on m and r] is the negative natural logarithm of this probability, as follows:

$$\text{SampEn} = -\ln(p) = -\ln(A/B) - \ln(B) - \ln(A)$$

We are challenged in interpreting SampEn because of its strong dependency on the size of the tolerance r and in comparing entropy estimates made using different values of r. Generally, smaller values of r lead to higher and less confident entropy estimates because of falling numbers of matches of length m and, to an even greater extent, matches of length $m + 1$. To address this issue, a measure called the quadratic entropy rate, based on densities rather than probability estimates, was introduced (11). To normalize for the value of r, SampEn was modified by dividing the probability p by the length of the overall tolerance window $2r$. The resulting quantity, called QSE, is as follows:

$$\text{QSE} = -\ln[p/(2r)] = -\ln(p) + \ln(2r) = \text{SampEn} + \ln(2r)$$

With this formulation, estimates made with different values of r measure the same inherent quantity and can be compared directly. Another advantage to this approach is that that the tolerance r can be optimally varied for each individual data record. This is analogous to varying the bin widths of histograms to optimally depict the distribution of a particular data set.

An important aspect of the SampEn algorithm is that self-matches are not counted (19, 20). This significantly reduces bias but contributes to the problem of falling counts of template matches to the point that A and even B could be zero, leading to infinite or indeterminate estimates. This becomes an increasing concern for short records. In addition, the accuracy of a probability estimate A/B is dependent on both the magnitude of the numerator A and denominator B. For example, an estimate of 0.1 with 100/1,000 is more accurate than an estimate with 1/10. Because QSE allows the flexibility to vary r, inaccurate probability estimates can often be avoided. As discussed in Ref. 11, one approach to accomplish this, called the minimum numerator count method, is to vary r until a specified number of matches A is attained. For example, as shown here, a minimum numerator count of 5 was found to give optimally accurate estimates detecting AF in short records of length $n = 12$.

RESULTS

Figure 1 shows ECGs from a patient in the MIT-BIH AF database. Figure 1 shows the most common rhythms we encountered: sinus rhythm (A) and AF (B). The obvious differences in regularity justify an approach to detecting AF based

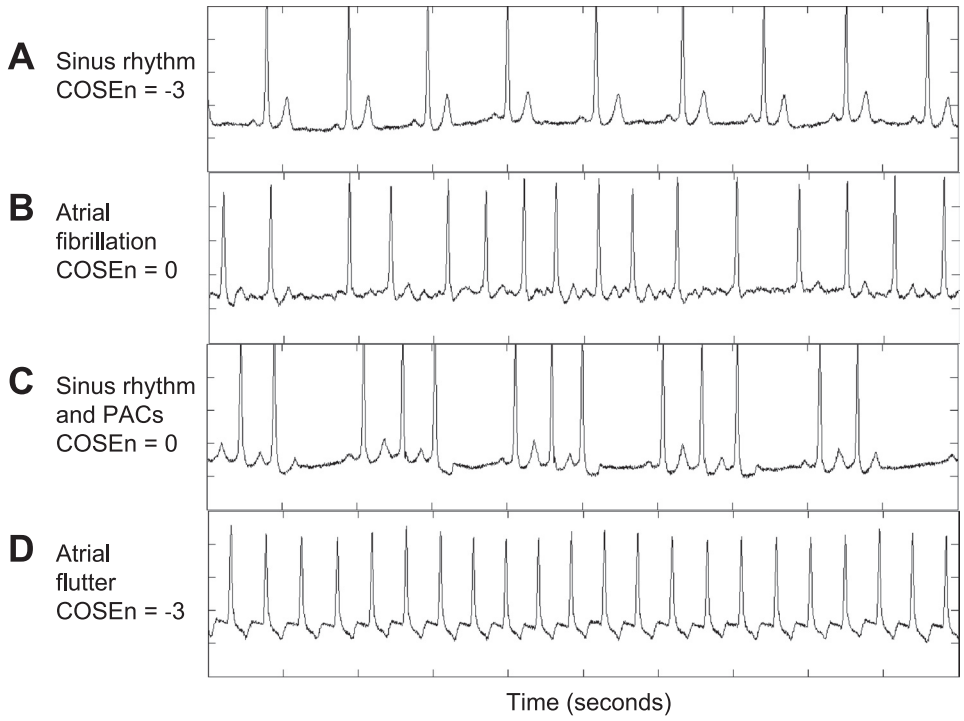

Fig. 1. ECGs from *patient 202* in the MIT-BIH atrial fibrillation (AF) database. While the main exercise was to distinguish sinus rhythm (*A*) from AF (*B*), two important problems are sinus rhythm with ectopy (*C*) and atrial flutter (AFL) with regular *n*:1 (here 2:1) atrioventricular conduction (*D*). COSEn, coefficient of sample entropy (SampEn); PAC, premature atrial contraction.

on entropy calculations. Two other ECGs from this patient are shown, each representing more difficult problems, and they are discussed further below (in *COSEn: an entropy estimate optimized for the detection of AF*).

The rest of this section examines the selection of parameters m and r and introduces the important notion of varying r until a confident entropy estimate is possible. In particular, we explored an approach using a minimum numerator count. Unless otherwise noted, all the Holter results reflect COSEn calculations performed once per hour using 12-beat segments, with a total sample of 288 beats per 24-h recording.

Optimizing AF detection using entropy estimation: is $m = 1$ sufficient? Figure 2 shows justifications for selecting m to be 1. Figure 2*A* shows an autocorrelation function of RR interval data from AF and AFL compared with non-AF segments from the MIT-BIH AF database. In AF, there was very reduced correlation at lag = 1 beat, suggesting that there was no additional information about order in longer templates and that $m = 1$ is a valid selection. Figure 2*B* shows plots of the ROC area for detecting AF as a function of m for RR interval time series of lengths 8, 12, 16, 25, and 50. In each case, the best distinction was made at $m = 1$. Such a short template length is well suited to the clinical problem of rapid entropy estimation using minimal calculations.

What should the minimum numerator count be? Figure 3*A* shows the ROC area for distinguishing AF from non-AF in the MIT-BIH and UVa databases as a function of the minimum numerator count. (In 12-beat segments, the maximum number of matches is 66.) The ROC area peaks near the count of 5, suggesting this to be a reasonable minimum numerator count for this clinical and numerical data set. Figure 3*B* shows the directly related result that COSEn estimates are stable across a wide range of minimum numerator counts.

What should be the initial value of r? While the minimum numerator count approach should be robustly applicable, some implementations call for rapid detection or parsimonious computation. Thus, we determined an optimal value of r to suffice in most cases of AF detection in 12 beats using the MIT-BIH AF database. Figure 4*A* shows the effect of the value of r on the important quantities in entropy estimation to detect AF for the MIT-BIH and UVa data sets. The horizontal axis is r, and the measured quantities are ROC area for AF detection,

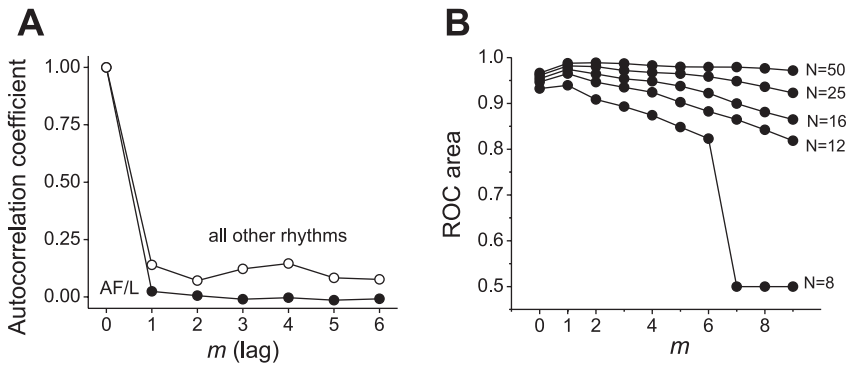

Fig. 2. Justification of selection of length $m = 1$ for entropy estimation. *A*: plot showing the average autocorrelation function for AF and AFL (filled symbols) and all other rhythms (open symbols). There was less correlation between successive beats in AF, suggesting that $m = 1$ was sufficient for entropy estimation. *B*: receiver operating characteristic (ROC) area for detecting AF or AFL for different segment lengths and values of m. The value of 1 had a high discrimination even at segment lengths of 12 beats.

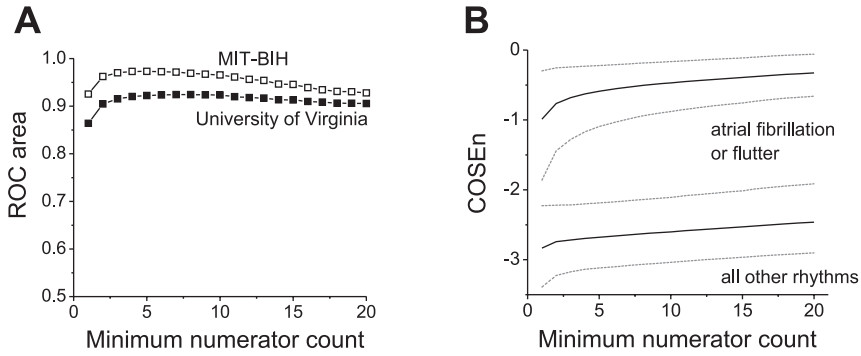

Fig. 3. Minimum numerator count approach for AF detection using entropy estimation. *A*: ROC curve areas for detecting AF in the MIT-BIH AF database and in 900 consecutive Holter monitor recordings for the University of Virginia Heart Station as a function of the minimum numerator in the conditional probability (CP) fraction. Detection was optimal near a minimum numerator count of 5. *B*: stability of the entropy estimates over a range of minimum numerator counts. The solid lines are the median and the dotted lines are the 25th and 75th percentiles.

average CP of a match, and proportions of 12-beat segments that have no template matches (CP = 0) or all matches (CP = 1). The optimal selection for r should have the ROC near maximal, average CP near 0.5, and near-minimal $p(\text{CP} = 0)$ and $p(\text{CP} = 1)$. This is satisfied for $r = 30$ ms, as noted by vertical shaded bar, and we used this as our initial value of r.

Heart rate adds information to entropy in detecting AF. To determine the impact of the heart rate itself on the detection algorithms, we used logistic regression analysis. Heart rate was introduced as a predictor variable along with corrected entropy to distinguish AF from non-AF in the MIT-BIH AF database. The regression coefficients for both heart rate and entropy were significant, suggesting that both variables add independent information. Since the coefficients were approximately equal in magnitude but opposite in sign, we adopted the convenient method of subtracting the logarithm of the mean RR interval from the corrected entropy estimate.

COSEn: an entropy estimate optimized for the detection of AF. Thus, the final algorithm was to calculate SampEn, allowing r to vary from 30 ms if necessary so that the CP was reasonably estimated, and calculate COSEn = SampEn − ln(2r) − ln(mean RR interval). We sought a COSEn threshold that would lead to an observed AF burden of 43% in the MIT-BIH data set, and we found this to be −1.4. We found this cutoff to have a sensitivity of 91% and a specificity of 94% in the MIT-BIH data set. Thus, the final step was to assign the diagnosis of AF for values greater than −1.4.

Comparison with other methods. Figure 4*B* compares methods of AF detection using entropy and variability. The ordinate is the ROC area for AF detection in the MIT-BIH database, and the abscissa is the number of beats. We compared SampEn with $m = 1$ and $r = 0.2$ (SD), a popular formulation, and the coefficient of variation (CV), the SD normalized by the mean RR interval. COSEn reached a high ROC area in as few as 10 beats, whereas SampEnt, under the usual conditions, required as many as 50 or more beats, and CV never reached as high a value of ROC area.

Examples. Figure 5 shows an example of COSEn analysis in a patient from the MIT-BIH AF data set that had paroxysmal AF and for whom ECGs are shown in Fig. 1. The blue dots show points of agreement between COSEn analysis and the electrocardiographer, whether in sinus rhythm (Fig. 5*A*) or AF (Fig. 5*B*). As noted above, the points in Fig. 5, *C* and *D*, confound entropy analysis. Figure 1*C* shows sinus rhythm with frequent ectopy, in this case, atrial in origin. The RR intervals were irregular, and the resulting entropy estimate had a high value, more like AF than sinus rhythm. Figure 5*D* shows AFL with 2:1 conduction. The RR intervals were very regular, and the entropy estimate was low, more like sinus rhythm than AF. These two scenarios, frequent ectopy and AFL, will not be correctly labeled by entropy estimates and are important in understanding and interpreting entropy estimates of heart rate.

Validation of COSEn in the UVa data set. As noted, Fig. 3 shows the ROC area for detecting AF/AFL in both the MIT-BIH and UVa data sets using COSEn calculations on 12 beats every hour. The distinction between groups was very good, with ROC areas of 0.9 or better. The reduced performance in the UVa data set may be due to our practices of labeling a 12-beat segment as AF if it contained only a single AF label, the inclusion of AFL RR interval time series as AF despite low entropy estimates, and the effect of ectopic beats during sinus rhythm.

Figure 6 shows histograms of entropy estimates for AF/AFL and all other rhythms in the MIT-BIH and UVa datasets using all the 12-beat segments. Generally, AF detection was better in the MIT-BIH data set, where it was more prevalent. We

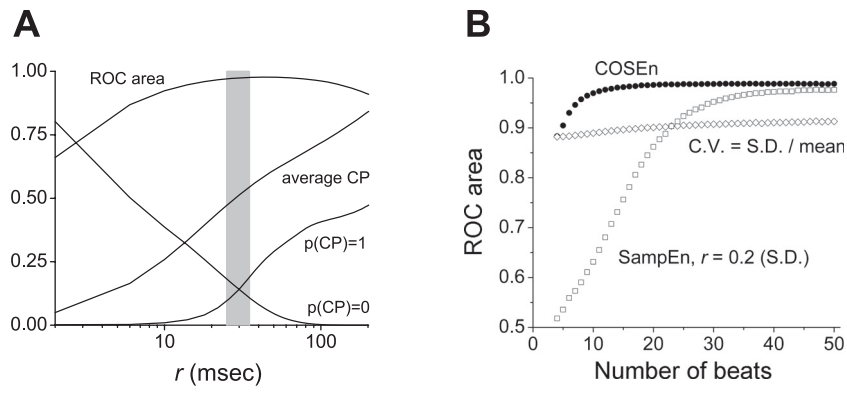

Fig. 4. *A*: quantities important for entropy estimation. The results shown in the plot justify the initial selection of tolerance $r = 30$ ms (vertical shaded bar) to detect AF in 12 beats. At that point, the probability that a template will have no matches [$p(\text{CP}) = 0$] was the same as the probability it will find all matches [$p(\text{CP}) = 1$]. For $r = 30$ ms, the ROC area was near 1 and the average CP was near 0.5. *B*: improved diagnostic performance of COSEn compared with coefficient of variation and SampEn using standard approach of taking r to be 20% of the sample SD.

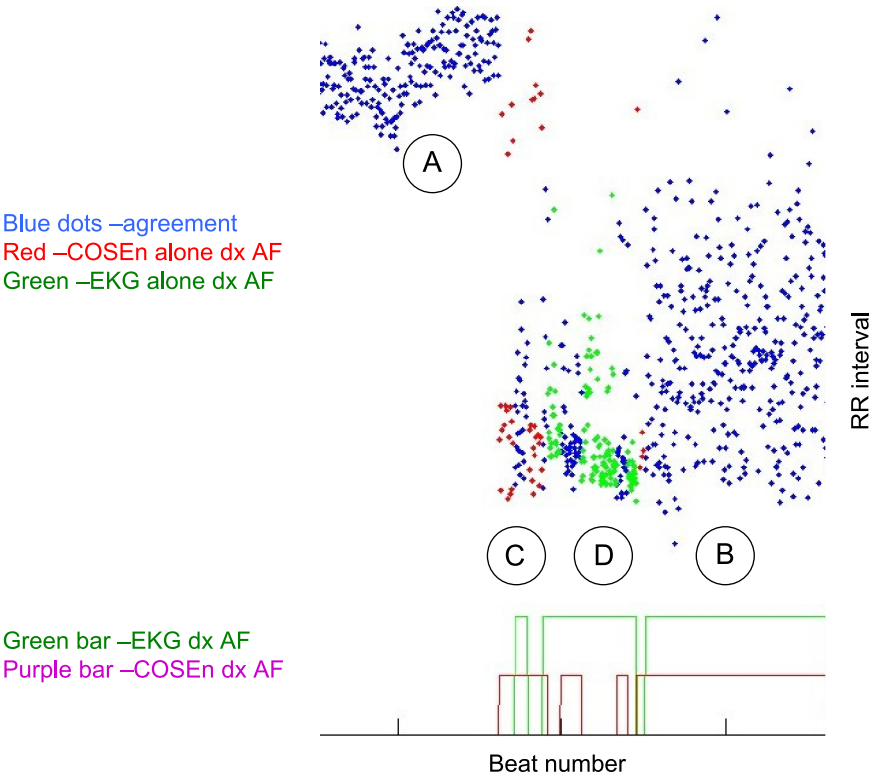

Fig. 5. Performance of COSEn analysis using 12-beat segments in a representative record from the MIT-BIH AF database. The blue dots show points of agreement between COSEn analysis and the electrocardiographer, whether in sinus rhythm (*A*) or AF (*B*). Entropy failed to detect sinus rhythm when ectopy was present (*C*) and failed to detect AF when there was organized atrial activity and regular atrioventricular conduction (*D*).

attribute the overlap in the tails of the distributions to the practices cited above.

We investigated the correlation of AF burdens detected using COSEn of 12-beat segments every hour with those derived from an ECG inspection of the entire record. For the 940 patients over the age of 40 yr old, the AF burden using COSEn of only 12-beat segments measured only once every hour was compared with that from ECG inspection of the entire recording. The correlation coefficient r between the two methods was 0.88, but a more clinically useful way to quantify the performance is as a binary diagnosis of AF versus non-AF. Table 1 shows the confusion matrix for this type of analysis, where, for purposes of comparison, a burden exceeding 10% for both methods was used to diagnose AF. The sensitivity and specificity of COSEn were 91% and 98%, respectively, and the positive predictive value (PPV) was 63%. The majority of the 79 false positives were recordings with frequent, complex ventricular ectopy or electronic pacemakers. Accounting for these cases can significantly increase the performance. For example, discounting segments with >2 ectopic beats (as labeled by the Holter) increases the PPV to 72% and has 27 fewer false positives.

Age-related changes in entropy of sinus rhythm and AF. While heart rate dynamics during sinus rhythm reflect complex and varying coupling of the sinus node to the autonomic nervous system, dynamics during AF should result only from atrioventricular nodal properties. Figure 7 shows mean COSEn measured from 12-beat segments every hour in UVa Holters that contained only sinus rhythm or AF as a function of age. Figure 7*B* shows the same data on a log scale to clarify the change in COSEn of sinus rhythm in the young. Entropy of sinus rhythm rose over the first 10–15 yr of life and then fell, consistent with a prior report (13). Note that the COSEn of sinus rhythm in the elderly was sometimes high. This was usually due to frequent ventricular ectopy. In AF, on the other hand, COSEn was higher ($P < 0.0001$) but changed less with

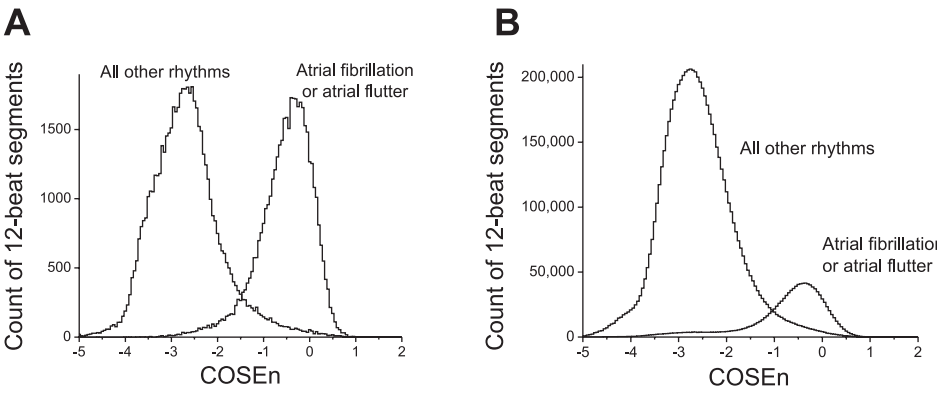

Fig. 6. Distributions of COSEn for AF and non-AF in the MIT-BIH (*A*) and University of Virginia (*B*) data sets. Note the much larger size of the University of Virginia data set as well as the reduced proportion of AF and AFL.

Table 1. *Diagnostic performance of COSEn using 12-beat segments every hour*

	Holter AF burden		
COSEn AF burden	≥10%	<10%	Total
≥10%	131	79	210
<10%	13	717	730
Total	144	796	940

COSEn, coefficient of sample entropy; AF, atrial fibrillation.

age. Cases of low COSEn labeled AF were usually due to AFL. The slope of the regression line fell from −0.0082 in sinus rhythm to −0.0049 in AF.

DISCUSSION

We studied the optimization of entropy estimation for the important clinical problem of detecting AF in short heart rate records. Our major innovation is the idea that there should be systematic choice of parameters to increase confidence in the CP estimate that is at the core of the entropy estimation. To achieve the necessary minimum numerator counts for this clinical problem, we propose the template length parameter $m = 1$ and that the tolerance parameter r be allowed to vary. We propose that the resulting entropy estimate, the negative natural logarithm of CP, be corrected for the choice of r (11), allowing a direct comparison of results despite differences in r. In addition, we propose correction of the entropy estimate by the mean heart rate because it contributed significant and independent diagnostic information in the detection of AF. This optimized entropy estimation technique we call COSEn.

We selected parameters for the detection of AF in short RR interval time series using these ideas. First, we selected $m = 1$ based on autocorrelation analysis. Second, we selected an initial value of r that optimized the parameters of the CP calculation. Third, for each segment, we assured a sufficient number of matches of length $m + 1$, the numerator of the CP fraction; we call this the minimum numerator count. We allowed r to vary until a minimum numerator count was achieved and normalized the entropy estimate by $\ln(2r)$ (11). Finally, we normalized the entropy estimate by the mean heart rate.

The fourth step, normalization by $\ln(2r)$, is a new consideration in entropy estimation in biological and clinical time series analysis. In particular, this is a significant departure from the usual practice of taking r as a fixed proportion of the SD of the time series values. Previously, ApEn estimators follow the prescription that r be a fraction of the sample SD, usually 0.2. To assure non-zero values in the CP calculation, ApEn allows templates to match themselves. This results in a bias of ApEn toward lower values (15). Porta and coworkers (16) developed conditional corrected entropy, which penalizes entropy estimates when few matches are found. We (12, 19, 20) later developed SampEn as a more robust entropy estimate that does not allow self-matches, and we and Richman (10, 11, 18) have reported on its statistical properties. SampEn is the basis for MSE analysis as developed by Costa and coworkers (1) and implemented in AF detection in long records.

We (12) previously proposed a rational selection of parameters from plots of an estimate of efficiency (a combination of low error and CP near 0.5) as a function of m and r for neonatal heart rate data. We also propose that the minimum numerator count be selected so as to assure confidence in the count and to avoid CP too near 0 or too near 1; here, there will may little distinction between groups.

Ramdani and coworkers (17) used SampEn to characterize postural swaying and systematically assessed choices of m and r. They picked m based on the convergence of SampEn results and then picked r based on the measurement of relative errors. Richman (18) proposed the selection of m based on the amount of additional information gained, stopping with the smallest value that added new information to the estimate at template length $m - 1$.

Importantly, we note that these parameter values were derived in ECG databases for the specific purpose of detecting AF. They might differ for other data sets from other kinds of sources or from these data sets for other clinical questions. Future studies of entropy estimation should systematically evaluate choices for these parameters.

Limitations: the problem of AFL. The strategy of diagnosing atrial arrhythmia using RR intervals is particularly poorly suited to detecting AFL because the ventricular rhythm can either be fixed at integer fractions of the atrial rate, which is usually near 300 beats/min, or can be as variable as in AF. Despite the obvious impossibility of diagnosing AFL with a regular ventricular rhythm using our approach, we included AFL with AF because the clinical imperative for AFL is the same as AF, especially with regard to the important issue of anticoagulation. This is one of the reasons why the diagnostic performance of our (or anyone's) numerical algorithm is better

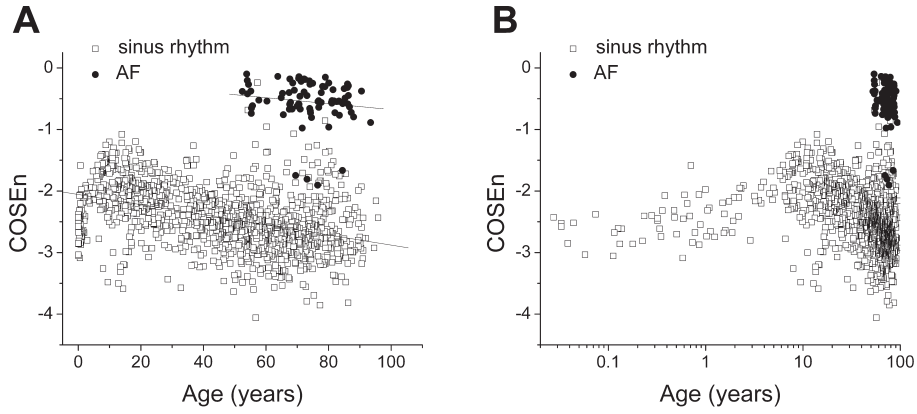

Fig. 7. Entropy of sinus rhythm falls with age more so than that of AF. Both plots show COSEn as a function of age. Open symbols are data from Holters with all AF ($N = 80$), and filled symbols are data from Holters with all sinus rhythm ($n = 1{,}250$). The linear plot (*A*) also shows regression lines (for AF, $R = -0.14$ and $P = 0.21$; and for sinus rhythm, $R = -0.41$ and $P < 0.0001$). The logarithmic plot (*B*) shows a rise in COSEn in the first years of life.

in the MIT-BIH AF database and worse in the real-world UVa database. For example, removing the 18 patients with AFL from our analysis increased the ROC area from 0.928 to 0.955. We know, though, of no better approach that does not use the atrial electrogram.

Clinical implications. An important implementation for this method to detect AF in short records is the measurement of AF burden from single-lead pacemakers or defibrillators. New onset or new recognition of AF in patients with reduced left ventricular systolic function is common, and therapeutic decisions are made easier by an accurate estimate of the AF burden. At the same time, these devices have limited storage and calculating capacity, and a simple algorithm requiring only a few RR intervals might find widespread clinical use.

We considered the clinical impact of the imperfect diagnostic performance of this efficient algorithm. Misdiagnosis of AFL as non-AF was frequent, leading to lower estimates of AF burden using COSEn in our real-world Holter data set. In the clinical context, mistaking an AF burden of 100% for 50% is not a clinical catastrophe, as therapy with anticoagulation would be undertaken for either. Mistaking an AF burden of 0% for 1% is probably not important either, as therapy is not likely to result for either. Reporting an AF burden of 0% when it is higher than, say, 5%, is where trouble might lie, as this represents a failure to diagnose clinically important AF. We found this to be the case in 10 of 642 Holters in patients over the age of 40 yr old who had either sinus rhythm or AF throughout the recording, and 7 of them had AFL, our nemesis. While this result will vary depending on the prevalence of AF, it is reassuring that missing an AF/AFL burden of $>5\%$ does not seem to be common.

Entropy of sinus rhythm rises until an age of 10 yr old or so and then falls with aging, consistent with the idea of reduced complex physiological variability (13) or altered coupling of the sinus node to the autonomic nervous system. We find, though, that entropy of AF does not depend on age. While the mechanism is not known, one explanation might be that the impulse conduction properties of the atrioventricular node are more resistant to aging than the impulse-making properties of the sinus node. Whatever the cause, an appealing clinical result is enhanced diagnostic accuracy of COSEn in the elderly, where AF is increasingly prevalent.

Summary and conclusions. We have developed new methods for systematic approaches to entropy estimation in short time series and implemented them in the important clinical problem of detecting AF. We call the result COSEn, which detects AF in as few as 12 beats. Here, we use new ideas about how to estimate entropy in very short time series of all descriptions and uniquely allow for variation in the tolerance measure r. Entropy estimation may be a useful adjunct for patients with single-lead CIEDs, a population at particular risk for AF and all the clinical ills that follow.

ACKNOWLEDGMENTS

The authors thank Dr. D. DeMazumder, Y. Zhou, P. Iazzetti, B. Dickinson, and T. Moss for help in overreading the Holter recordings.

GRANTS

This work was supported by an American Heart Association Mid-Atlantic grant-in-aid.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

1. **Costa M, Goldberger AL, Peng CK.** Multiscale entropy analysis of biological signals. *Phys Rev E Stat Nonlin Soft Matter Phys* 71: 021906, 2005.
2. **Eckmann JP, Ruelle D.** Ergodic theory of chaos and strange attractors. *Rev Mod Phys* 57: 617–654, 1985.
3. **Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE.** Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnticoAgulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 285: 2370–2375, 2001.
4. **Grassberger P, Procaccia I.** Estimation of the Kolmogorov entropy from a chaotic signal. *Phys Rev A* 28: 2591–2593, 1983.
5. **Griffin MP, Lake DE, Bissonette EA, Harrell FE Jr, O'Shea TM, Moorman JR.** Heart rate characteristics: novel physiomarkers to predict neonatal infection and death. *Pediatrics* 116: 1070–1074, 2005.
6. **Griffin MP, O'Shea TM, Bissonette EA, Harrell FE Jr, Lake DE, Moorman JR.** Abnormal heart rate characteristics are associated with neonatal mortality. *Pediatr Res* 55: 782–788, 2004.
7. **Griffin MP, O'Shea TM, Bissonette EA, Harrell FE Jr, Lake DE, Moorman JR.** Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness. *Pediatr Res* 53: 920–926, 2003.
8. **Kolmogorov AN.** Entropy per unit time as a metric invariant of automorphism. *Doklady Rus Acad Sci* 124: 754–755, 1959.
9. **Kolmogorov AN.** New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces. *Doklady Rus Acad Sci* 119: 861–864, 1958.
10. **Lake DE.** Nonparametric entropy estimation using kernel densities. *Methods Enzymol* 467: 531–546, 2009.
11. **Lake DE.** Renyi entropy measures of heart rate Gaussianity. *IEEE Trans Biomed Eng* 53: 21–27, 2006.
12. **Lake DE, Richman JS, Griffin MP, Moorman JR.** Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol* 283: R789–R797, 2002.
13. **Pikkujamsa SM, Makikallio TH, Sourander LB, Raiha IJ, Puukka P, Skytta J, Peng CK, Goldberger AL, Huikuri HV.** Cardiac interbeat interval dynamics from childhood to senescence: comparison of conventional and new measures based on fractals and chaos theory. *Circulation* 100: 393–399, 1999.
14. **Pincus SM.** Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA* 88: 2297–2301, 1991.
15. **Pincus SM, Goldberger AL.** Physiological time-series analysis: what does regularity quantify? *Am J Physiol Heart Circ Physiol* 266: H1643–H1656, 1994.
16. **Porta A, Baselli G, Liberati D, Montano N, Cogliati C, Gnecchi-Ruscone T, Malliani A, Cerutti S.** Measuring regularity by means of a corrected conditional entropy in sympathetic outflow. *Biol Cybern* 78: 71–78, 1998.
17. **Ramdani S, Seigle B, Lagarde J, Bouchara F, Bernard PL.** On the use of sample entropy to analyze human postural sway data. *Med Eng Phys* 31: 1023–1031, 2009.
18. **Richman JS.** Sample entropy statistics and testing for order in complex physiological signals. *Commun Stat Theory Methods* 36: 1005–1019, 2006.
19. **Richman JS, Lake DE, Moorman JR.** Sample entropy. *Methods Enzymol* 384: 172–184, 2004.
20. **Richman JS, Moorman JR.** Physiological time series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278: H2039–H2049, 2000.
21. **Shannon CE.** A mathematical theory of communication. *Bell Syst Tech J* 27: 379–423, 1948.
22. **Sinai YG.** On the notion of entropy of a dynamical system. *Doklady Rus Acad Sci* 124: 768–771, 1959.