Approximate Entropy as a Measure of Regularity in Hormone Series

Nichole E Carlson and Morton B Brown
University of Michigan, Department of Biostatistics
1420 Washington Heights
Ann Arbor, MI 48109-2029, USA
ntcarlso@umich.edu and mbb@umich.edu

1. Introduction

Hormones, such as luteinizing hormone (LH), are secreted into the blood in short bursts, called pulses, and removed by a decay process. Using hormone assays from blood samples obtained at equally spaced intervals, investigators study secretion characteristics such as pulse frequency and amplitude to learn whether various diseases affect hormone secretion.

Approximate entropy (ApEn) has been used to assess regularity in a series. Although it is not specifically tailored to the identification and characterization of pulses (Pincus and Keefe 1992), the authors state that ApEn evaluates both dominant and subordinate patterns in data and therefore detects differences in pulsatile series even when pulse frequency and height do not differentiate between the secretion patterns. ApEn does distinguish between complex mathematical systems that can generate hormone data (Pincus, 1994), but little has been studied about the sensitivity of ApEn to randomness in the main pulse characteristics (timing, mass, duration, and noise). We were interested in assessing whether ApEn was sensitive to timing, mass, and duration. We generated simulated data with fixed pulse characteristics and then added random effects to the timing, mass, and duration parameters. We studied combinations of fixed and random characteristics under various model assumptions (longer vs. shorter half-life, slow vs. rapid pulsing, and noise vs. no-noise) to assess which parameter characteristics cause changes in the ApEn statistic.

Since ApEn is being used to characterize the regularity of secretion we also wanted to assess whether conclusions made in peripheral blood, the samples collected in human studies, could be used to draw conclusions about the regularity of secretion closer to the source. We calculated the ApEn statistics for LH data collected in the portal and peripheral blood in six ewes. This enables us to evaluate the influence of hormone dispersion in the peripheral blood on the ApEn statistic.

2. The Approximate Entropy Statistic

ApEn depends on two parameters, m and r, where m specifies the length of runs for comparing closeness, and r defines closeness; both are specified by the user. For instance, if $m=1$, ApEn is a measure of whether two pairs of observations (x_t and x_{t+1}, and x_w and x_{w+1}) are close given that the first observation in each pair are close to each other. A lower value of ApEn corresponds to more regularity and, conversely, a higher ApEn corresponds to less regularity. For instance, the ApEn for a complete phase generated by a sine curve using 100 points of data is 0.16, while the ApEn for 100 points of a random sample from a uniform distribution is 1.68. ApEn increases with sample size (Pincus, 1994), thus comparing series of equivalent lengths series is necessary. For our analysis we used the software provided by The Center for Biological Timing at the University of Virginia. As suggested by the literature (Pincus et al., 1999) we used $m=1$ and $r=20\%$ of the standard deviation of the series.

3. Results

Midgley et al. (1996) assayed LH from blood samples taken simultaneously from the portal (near the pituitary gland) and the peripheral blood systems every 5 minutes from 6 ewes. The ApEn statistic was calculated for these pairs of series. For four of the six sheep, blood was collected for 6 hours (72 samples) and for two, blood was collected for 12 hours (144 samples).

Table 1 contains ApEn values for each LH series from the six ewes. Since pulses in the portal series are more regular than in the peripheral series, portal LH has lower ApEn statistics when compared to peripheral LH. In In 3 of the 6 series, the ApEn values for the peripheral series are closer to the value under random ordering than to the value for the portal system.

In addition, we simulated data series using parameters similar to those estimated

Table 1: ApEn Statistics for Six Ewes

m=1 and r=20%		Portal	Peripheral	Random
ID	N	Blood LH	Blood LH	Order
EWE1	71	0.75	1.29	1.57
EWE10	72	0.97	1.13	1.56
EWE2	71	0.92	1.26	1.56
EWE3	143	0.74	1.41	1.85
EWE5	136	0.86	1.28	1.79
EWE6	71	1.23	1.34	1.54

for the 6 ewes above. Parameters were either fixed throughout the simulated series or treated as random. Then the observations in each series were also randomly permuted 10 times to create series with noise instead of pulses.

ApEn increases as pulse frequency increases regardless of the random vs. fixed state of the parameters. Increasing half-life does not affect ApEn when there is no noise, but ApEn increases with half-life when there is noise. Increasing the average pulse duration increases ApEn regardless of noise and fixed-random parameter combinations. The randomly ordered series have similar values for all models except for the slow frequency model. This indicates that there is always some pattern in the slow pulse frequency series. This is most likely due to the long decay patterns in the slow frequency series and suggests even under random ordering that this autoregressive relationship is still prevalent in runs of length two (the pattern we are detecting in our calculation of ApEn, m=1).

By comparing models where the only difference is in the fixed-random state of one parameter (interpulse interval, pulse mass, or pulse input duration), one can assess the ApEn statistic's sensitivity to the randomness of a particular characteristic. ApEn is most sensitive to the randomness of the interpulse interval.

4. Discussion

ApEn is a measure of long term trends in a series and increases when those long term trends are disrupted. The long term trend in pulsatile hormone data is the decay phase and pulses clearly disrupt these trends in the data. The more pulsing that there is in a series, the greater the disruption of the decay pattern and, therefore, in a sense, ApEn is a measure of pulse frequency.

It is important to understand what areas of the series contribute to the calculation of the ApEn statistic. In hormone data the decay and baseline areas contribute the most to the counts in the ApEn calculation. This supports the explanation of why ApEn is particularly sensitive to the pulse frequency and to the randomization of the interpulse interval. Both of these conditions disrupt the areas in the series that contribute the most matches to the ApEn calculation.

Randomness in the pulsatile input duration does not influence the ApEn statistic greatly. Therefore, one must exercise caution when interpreting an increase in ApEn as in increase in the randomness of pulsatile secretion. ApEn is related to frequency and randomness of the pulse locations, but does not always give information about the randomness of the input.

The ApEn statistic in the peripheral blood may be closer to the value generated by a random series than to the ApEn value from the portal blood. Therefore, use of the ApEn statistic to characterize pulsatility is premature.

REFERENCES

Mauger, D., et. al. A comparison of methods that characterize pulses in a time series. *Statistics in Medicine*. 14: 311-325, 1995.

Midgley, A.R. et al. Nonclassical secretory dynamics of LH revealed by hypothalamo-hypophyseal portal sampling of sheep. *Endocrine*. 6:133-143, 1997.

Pincus, S.M. and Keefe, D.L. Quantification of hormone pulsatility via an approximate entropy algorithm. *The American Journal of Physiology*. 262:E741-E754, 1992.

Pincus, S.M. Quantification of evolution from order to randomness in practical time series analysis. *Methods in Enzymology*. 240:68-89, 1994.

Pincus, S.M et al. Hormone pulsatility discrimination via coarse and short time sampling. *The American Journal of Physiology*. 277:E948-E957, 1999.

RESUME

Approximate entropy (ApEn) is a statistic that is used to characterize sequential regularity in a data series. For example, ApEn is used to quantify the regularity of the underlying pulsatile secretion patterns in time series of hormone concentrations. A higher ApEn value implies greater randomness in the underlying series, but this assumption has not been studied for pulsatile data. We use simulated data to study how randomness in models for the common pulsatile parameters (e.g. pulse amplitude, interpulse interval, and pulse duration) influence the ApEn statistics. The variability induced by random effects increases the ApEn statistics and masks the underlying regularity of the pulse generator. Luteinizing hormone from six ewes was assayed in peripheral and portal bloods to compare the ApEn statistics that would be obtained from circulatory blood sampling (as done in humans) to that obtained near the generator's source. Although lower than from randomly ordered series, the ApEn values from the jugular series are closer to the values obtained under random reordering than to the values obtained in the portal series.