Older males secrete luteinizing hormone and testosterone more irregularly, and jointly more asynchronously, than younger males

(approximate entropy/aging/pulsatility/reproduction/variable lag process)

STEVEN M. PINCUS[†‡], THOMAS MULLIGAN[§], ALI IRANMANESH[¶], SYLVIA GHEORGHIU[§], MICHAEL GODSCHALK[§], AND JOHANNES D. VELDHUIS[∥]

[†]990 Moose Hill Road, Guilford, CT 06437; [§]Hunter Holmes McGuire Veterans Affairs Medical Center and Medical College of Virginia, Richmond, VA 23249; [¶]Endocrine Section, Veterans Affairs Medical Center, Salem, VA 24153; and [∥]Department of Internal Medicine, and National Science Foundation Center for Biological Timing, University of Virginia Health Sciences Center, Charlottesville, VA 22908

Communicated by Wylie Vale, The Salk Institute for Biological Studies, San Diego, CA, August 15, 1996 (received for review May 17, 1996)

ABSTRACT New statistical perspectives on the secretory patterns of both luteinizing hormone (LH) and testosterone (T) may prove useful in further understanding the aging process, and possibly ultimately in improving the diagnosis and treatment of spermatogenetic failure and loss of sexual interest. We examined serum concentration time-series for LH and T in 14 young (21–34 years of age) and 11 aged (62–74 years of age) healthy men. For each subject, blood samples were obtained at 2.5-min intervals during a sleep period, with an average sampling duration of 7 hr. For each of LH and T, we used the model-independent statistic approximate entropy (ApEn) to quantify the irregularity of the serum concentration time-series; to quantify joint LH–T secretory asynchrony, we employed the recently introduced cross-ApEn. Although mean (and SD) LH and T concentrations were indistinguishable in the two age groups ($P > 0.25$), for LH, aged subjects had greater ApEn values (1.525 ± 0.221) than younger individuals (1.207 ± 0.252), $P < 0.003$, indicating more irregular secretion in the older cohort. For T, aged subjects also had greater ApEn values (1.622 ± 0.120) than younger counterparts (1.384 ± 0.228), $P < 0.004$. In young, but not older men, ApEn(T) significantly exceeded ApEn(LH), $P < 0.02$. Aged subjects had greater cross-ApEn values (1.961 ± 0.121) than younger subjects (1.574 ± 0.249), $P < 10^{-4}$, with nearly 100% sensitivity and specificity, indicating greater LH–T asynchrony in the older group. In conjunction with previous findings of greater irregularity of growth hormone release with increasing age, we propose that increased secretory irregularity with advancing age may be a widespread hormonal phenomenon. Finally, theoretically, we clarify the need for quantifications such as ApEn and cross-ApEn via a study of a "variable lag" pulsatile process, and empirically note the potential wide applicability of cross-ApEn to quantify asynchrony in interconnected (hormonal) networks.

Greater understanding of the evolution of the hypothalamo–pituitary–testicular axis with aging is of vital importance both scientifically, in elucidating the physiology of reproductive capacity, and clinically, in assessing, e.g., a loss of libido or decreased reproductive performance. In recent years, there has been considerable study of luteinizing hormone (LH) and testosterone (T) serum concentration time-series in both younger and older males to develop such understanding, and to determine whether a hypothesized male climacteric (or so-called andropause) at least partially analogous to menopause in the female exists, and if so, in what precise sense. Such studies have evaluated changes in (*i*) mean concentrations of total and free T, and LH and the ratio of biological to immunological (B/I) LH activity (1–6); (*ii*) "near-term" (cir-

choral) pulsatility characteristics of LH and T, via changes in mean frequencies and amplitudes (7–10); (*iii*) "longer term," i.e., nyctohemeral characteristics of LH and T release (1). While considerable insight has already been gained, there remain nontrivial controversies, e.g., primary determinations of whether overall mean levels of LH and T decrease with increasing age, as discussed below. Furthermore, the precise neuroendocrine mechanisms that underlie such age-related changes remain largely unresolved.

Here, we consider possible reproductive aging changes from two perspectives entirely different from those mentioned above, namely by directly evaluating the degree of irregularity of each of the LH and T time-series, via approximate entropy (ApEn) (11, 12) and by quantifying the degree of asynchrony in the joint LH–T series, via cross-ApEn (12). In this study, we reanalyzed data collected by frequent venous sampling (every 2.5 min) overnight (8) to delineate the nature of changes in the secretion of these two hormones in healthy older men. Methodologically, first, it is imperative to note the importance of matching corresponding parts of the circadian epoch in the older and younger age cohorts, because aged men lose circadian fluctuations in serum T concentrations (1).

In addition to the biological relevance of assessing LH and T release from a distinct statistical perspective, we particularly note the potential broad statistical utility of the recently introduced cross-ApEn to quantify asynchrony or conditional irregularity in interconnected (hormonal) networks. In the *Appendix* we consider: (*i*) the manner in which ApEn and cross-ApEn have a quite different and complementary primary orientation from both linear correlation and the power spectrum and (*ii*) why this separate perspective affords the biologist distinct tools from which changes in the extent of synchrony in interconnected hormonal systems can be clearly determined. This quantification strategy is relevant to many feedback and/or control systems and models for which cross-correlation and cross-spectral methods fail to fully highlight markedly changing features of the data sets under consideration.

MATERIALS AND METHODS

The study group comprised 14 young (21–34 years of age) and 11 aged (62–74 years of age) healthy nonsmoking men within 20% of ideal body weight. For each subject, blood samples were obtained during a sleep period on a second night of study in the General Clinical Research Center at the University of Virginia, at 2.5-min intervals commencing at 2300 hr, with sampling terminated when the subject spontaneously awakened, for an average sampling duration of 7 hr. Serum LH concentrations were measured in duplicate by using a two-site monoclonal

Abbreviations: LH, luteinizing hormone; T, testosterone; ApEn, approximate entropy; GnRH, gonadotropin-releasing hormone; GH, growth hormone.
‡To whom reprint requests should be addressed.

immunoradiometric assay (Nichols Institute, San Juan Capistrano, CA). Assay sensitivity was 0.2 unit/liter according to the First International Reference Preparation. Serum total T concentrations were quantified in duplicate for each sample by using a solid-phase RIA (Diagnostic Products, Los Angeles). Assay sensitivity was 20 ng/dl. For both LH and T assays, intra- and interassay imprecision was less than 10%. Further subject and assay descriptions are given by Mulligan *et al.* (8).

ApEn. To quantify irregularity, we use ApEn, defined in ref. 11, further mathematical properties, and biological applications as given in refs. 12–18. ApEn is complementary to pulse-detection algorithms widely employed to evaluate hormone secretion time-series (19). ApEn evaluates both dominant and subordinant patterns in data; notably, it will detect changes in underlying episodic behavior not reflected in peak occurrences or amplitudes (17). Additionally, ApEn provides an explicit barometer of feedback system change in many coupled systems (17, 20). Within various endocrine contexts, ApEn has unveiled vivid distinctions ($P < 10^{-10}$) between normal and tumor-bearing subjects for GH (21) and aldosterone release (22), a pronounced and consistent gender difference in GH time-series irregularity in both human and rat (23), and a positive correlation between greater irregularity of GH secretion and advancing age (24).

ApEn assigns a nonnegative number to a time-series, with larger values corresponding to greater apparent process randomness (serial irregularity), and smaller values corresponding to more instances of recognizable patterns in the data. Two input parameters, m and r, must be specified to compute ApEn. Briefly, ApEn measures the logarithmic likelihood that runs of patterns that are close (within r) for m contiguous observations remain close (within the same tolerance width r) on next incremental comparisons; the precise mathematical definition is given in ref. 11.

In this study, we calculated ApEn(m,r) values for all data sets, $m = 1$ and $r = 20\%$ of the SD of the individual subject's hormone time-series.** Normalizing r to each time-series SD gives ApEn a translation- and scale-invariance to absolute serum concentration levels (14). ApEn is a relative measure of process regularity, and can show significant variation with changing background noise characteristics. Because ApEn generally increases with increasing process noise, it is appropriate to compare data sets with similar assay coefficients of variation, as we do here.

Previous studies that included both theoretical analysis (16, 17, 25) and clinical applications (13–15, 18, 21–24) have demonstrated that the input parameters indicated above produce good statistical validity (reproducibility) for ApEn applied to time-series of the lengths considered here. The ApEn application with $m = 1$ estimates the rate of entropy for a

first-order ($m = 1$) approximating Markov Chain to the underlying true process (26). Further technical discussion of mathematical and statistical properties of ApEn, including robustness to noise and artifacts, mesh interplay, relative consistency of (m,r) pair choices, asymptotic normality under general assumptions, statistical bias, and error estimation for general processes can be found elsewhere (16, 25).

Cross-ApEn. To quantify asynchrony (conditional irregularity), we use cross-ApEn, as introduced in ref. 12, definition 5. As noted there, cross-ApEn can be employed to compare sequences from two distinct yet intertwined variables in a network, herein applied to the joint LH–T time-series. The precise definition is thematically similar to that for ApEn:

Let $u = (u(1), u(2), \ldots u(N))$ and $v = (v(1), v(2), \ldots v(N))$ be two length-N sequences. Fix input parameters m and r. Form vector sequences $x(i) = (u(i), u(i + 1), \ldots u(i + m - 1))$ and $y(j) = (v(j), v(j + 1), \ldots v(j + m - 1))$ from u and v, respectively. For each $i \leq N - m + 1$, set $C_i^m(r)(v \parallel u) =$ (number of $j \leq N - m + 1$ such that $d[x(i), y(j)] \leq r)/(N - m + 1)$, where $d[x(i), y(j)] = \max_{k = 1, 2, \ldots, m} (|u(i + k - 1) - v(j + k - 1)|)$, i.e., the maximum difference in their respective scalar components. The $C_i^m(r)$'s measure *within a tolerance r* the regularity, or frequency, of (v-) patterns similar to a given (u-) pattern *of window length m*. Then define $\Phi^m(r) (v \parallel u)$ as the average value of $\ln C_i^m(r) (v \parallel u)$, and finally, define cross-ApEn$(m,r,N)(v \parallel u) = \Phi^m(r) (v \parallel u) - \Phi^{m+1}(r) (v \parallel u)$.

For this study, we applied cross-ApEn with $m = 1$ and $r = 0.2$ to standarized LH ($= u$) and testosterone ($= v$) time-series data, i.e., for each subject, we applied cross-ApEn(1, 0.2) to the $\{u^*(i), v^*(i)\}$ series, where $u^*(i) = (u(i) - \text{mean } u)/\text{SD } u$ and $v^*(i) = (v(i) - \text{mean } v)/\text{SD } v$. This standardization, in conjunction with the choice of m and r, ensures good replicability properties for cross-ApEn for the data lengths studied.

RESULTS

All statistical comparisons below employ the two-sided t test, except for the ApEn(LH) vs. ApEn(T) comparisons within each of the younger and older cohorts, for which we employed the paired t test. Results are given as mean $\pm$ SD.

Inspection of serum hormone concentration profiles suggests that clear pulse identification is a nontrivial endeavor, especially for the aged subjects' T series (see Fig. 4, ref. 8). For LH, aged subjects had greater ApEn values (1.525 $\pm$ 0.221) than younger subjects (1.207 $\pm$ 0.252), $P < 0.003$. For T, aged subjects also had greater ApEn values (1.622 $\pm$ 0.120) than younger subjects (1.384 $\pm$ 0.228), $P < 0.004$. In Fig. 1, scatterplots of mean LH level vs. ApEn(LH), and of mean T level vs. ApEn(T) visually confirm this statistical distinction. The decision rule that associates ApEn(LH) values greater than 1.445 with aged subjects has a specificity of 93% and a sensitivity of 82%, whereas the decision rule that associates ApEn(T) values greater than 1.60 with aged subjects has a specificity of 100% and a sensitivity of 64%.

Notably, there was no difference in *mean* serum LH levels between the younger (2.409 $\pm$ 0.658 units/liter) and aged subjects (2.830 $\pm$ 1.064 units/liter) levels, $P = 0.26$; and there was no difference in mean T levels between the younger (459 $\pm$ 148 ng/dl) and aged subjects (415 $\pm$ 115 ng/dl) levels, $P = 0.41$.

Aged subjects had greater cross-ApEn values (1.961 $\pm$ 0.121) than younger subjects (1.574 $\pm$ 0.249), $P < 10^{-4}$. Importantly, there was nearly complete separation of younger and older subject cross-ApEn values, as observed in Fig. 2, with all younger subjects' cross-ApEn values smaller than all but a single older subject's value. The decision rule that associates cross-ApEn values greater than 1.85 with aged subjects has a specificity of 100% and a sensitivity of 91%. In counterpoint, cross-correlation (Pearson "R"), reveals no significant differences, either in the Pearson R values directly, older subjects (0.078 $\pm$ 0.210) vs. younger subjects (0.030 $\pm$ 0.284), $P = 0.629$; or in the magnitude of the correlation, assessed by |Pearson R|,

---

**As a consequence of the sleep protocol, the measured time-series show a modest variation in length, about a $N = 168$ point mean. In comparing ApEn of different data set lengths, we neglect a small statistical bias in the estimator ApEn(m,r,N), as a function of N. As indicated in ref. 16, the expected value of ApEn(m,r,N) increases asymptotically with N to a limit ApEn(m,r) for virtually all processes. If we were to incorporate a bias or length correction, we would first observe that the younger group data set lengths were, on average, 7% larger than those of the aged subjects. Accordingly, the ApEn values of younger subjects should be *reduced* by a correction term that is a function of an unknown process, to accommodate an average 7% reduction in data set length for an unbiased comparison of group ApEn values between younger and aged subjects. Because ApEn values for the younger subjects are seen below to be significantly smaller than those for aged subjects, a reduction in ApEn values for the younger subject values would enhance the degree of younger-aged group separation. Thus, the reported findings are slightly conservative as to the extent of younger vs. aged group distinctions shown via ApEn. Furthermore, we established an upper bound of 0.06 for such a bias correction, for the range of data set lengths encountered here, so that any correction would have minimal qualitative effect on our results. A similar observation applies to the cross-ApEn analysis.
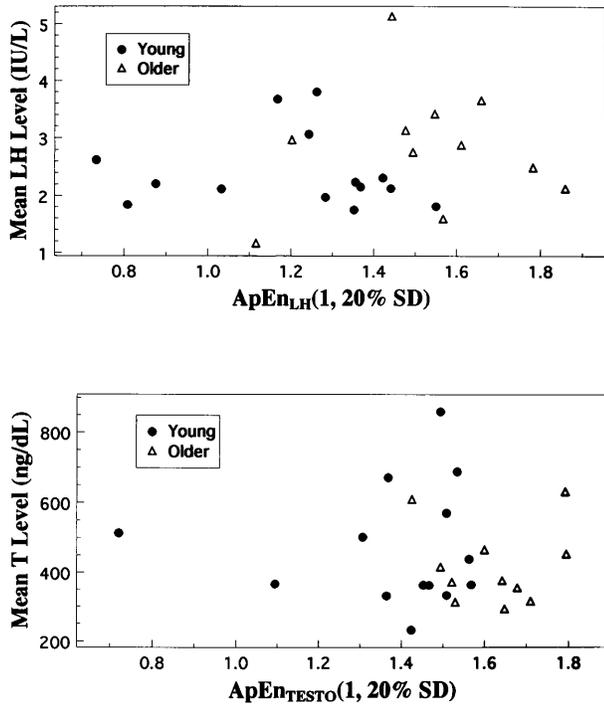
FIG. 1.   Age contrast in ApEn values for LH and T time-series considered singly. (*Upper*) Individual subject $ApEn_{LH}(m = 1, r = 20\%$ SD) values vs. mean serum LH concentrations. (*Lower*) Individual subject $ApEn_{TESTO}(m = 1, r = 20\%$ SD) values vs. mean serum T concentrations.

FIG. 2.   Individual subject cross-ApEn values vs. cross-correlation (Pearson R), applied to the joint LH–T time-series in healthy young vs. older men.

older subjects (0.150 ± 0.162) vs. younger subjects (0.231 ± 0.155), $P = 0.220$.

As another perspective on changes with aging in the joint LH–T variable system, we ascertained that for younger subjects serum T concentration time-series were more irregular than the corresponding LH concentration series, $P < 0.02$, whereas this per subject distinction vanished in the aged group, with no significance in the pairwise ApEn(LH) and ApEn(T) values, $P > 0.28$.[††]

DISCUSSION

Summary and Resultant Biological Questions. Summarizing the primary statistical results, for each of LH and T, older males have consistently and significantly more irregular serum reproductive-hormone concentrations than younger males. The distinction between ApEn(T) and ApEn(LH) indicating greater irregularity of the former in young men was lost in older men. Furthermore, cross-ApEn quantitatively supports a mechanistic hypothesis, a loss of synchrony with aging in the coupled LH–T system. The cross-ApEn finding reinforces the utility of studying network aspects, in addition to single-variable or nodal aspects, of hormone systems, both in statistical analysis and in modeling, and ultimately, in evaluating therapies. The determinations that mean serum LH and T concentrations in the young and older males were not significantly different, nor were linear cross-correlations, further suggest the need for the distinct perspectives assessed by quantification of irregularity and (a)synchrony.

Our inferences in the aging male reproductive axis (above) are in agreement with findings of greater growth hormone irregularity of release with increased aging (24). Thus, we hypothesize that greater secretory irregularity, and possibly greater asynchrony, with increased aging may be a more general paradigm for many hormones, potentially indicating a diminution of subsystem integrity (20) and/or of (synchronous) control.

It seems worthwhile to compare the results for the male to corresponding findings for the female, although sex-steroid levels decline more markedly in postmenopausal individuals than in aging men. Any comparisons between male and female evolution of "reproductive" hormone secretion as a function of increasing age are at best partial, given the cessation of female reproductive capacity in the aged, in contrast to continued, albeit diminished male fertility in advanced age. However, the above findings clearly indicate distinct quantitative shifts in male hormonal secretory dynamics with aging. Thus, the question arises as to how mechanistically greater individual signal irregularity (in LH or T release) or joint signal asynchrony are linked directly or causally to clinical changes in, e.g., spermatogenetic function or libido, as commonly seen in elderly males. Moreover, does the increased LH–T irregularity and asynchrony in older males occur gradually, at a relatively constant rate throughout life, or instead develop rather abruptly during a relatively shorter time-frame of months or years, the latter analogous to estrogen transitions in females across the menopause? For this last point, we hypothesize a more gradual evolution based on the somewhat analogous determination of a modest, slow continuous decrease in mean total T serum concentrations with increasing age seen, e.g. by Zumoff *et al.* (6) in a study of normal men 21–85 years of age.

In principle, there are several possibilities for the source of the erosion of LH–T synchrony quantified above. These include: (*i*) decreased multi-synaptic modulation and/or synchrony of the hypothalamic gonadotropin-releasing hormone (GnRH) neuronal network that produces the GnRH drive to pituitary LH synthesis and secretion; (*ii*) altered feedback control of individual and/or coupled GnRH–LH secretory activity by gonadal hormones, via a disrupted feedback signal, e.g., of T itself, or deficient responsiveness to the feedback signal; (*iii*) decreased GnRH and non-GnRH-dependent paracrine or autocrine coordination of LH secretion by gonadotroph cells; and/or (*iv*) disruption of effective (LH–T) stimulus-secretion coupling at the level of the Leydig cell in the testis. Further physiological studies will be required to clarify the precise basis of this change. Nonetheless, because there is increased ApEn of LH release after short-term ketoconazole

---

[††]It is appropriate to apply the statistic ApEn(T) − ApEn(LH), based on the serum concentration time-series, only to establish that there are significant differences between two subgroups, here younger vs. aged, based on the joint [LH–T] dynamics. Given that LH and T were measured via different assays, and have distinct clearance rates, we do not compare LH secretory irregularity to T secretory irregularity for either young or old.
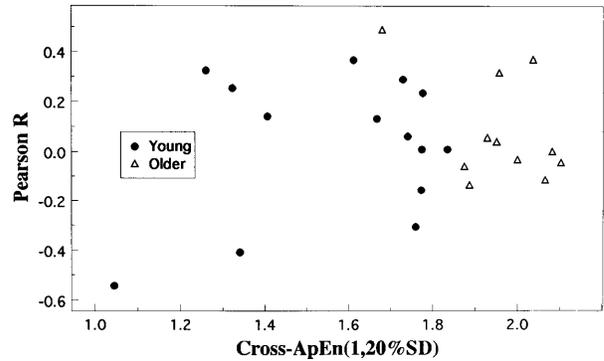
Medical Sciences: Pincus *et al.*

Proc. Natl. Acad. Sci. USA 93 (1996) 14103

treatment in young men when T secretion falls,‡‡ and increased ApEn of GH release with fasting as IGF-1 falls (21), we favor decreased feedback signal strength, or diminished GnRH–LH system responsiveness to feedback signal intensity, as a unifying hypothesis.

Complementarity of Present Findings to Previous LH and T Age-Related Changes. A number of age-related changes have been established earlier for both LH and T secretion. Our findings provide an entirely distinct and complementary perspective to previously identified differences in means or amplitudes of suitable physiological variables, so that secretory typicality can be assessed quantitatively both on the basis of mean and amplitude level of output and on the basis of orderliness of serial output. Mathematically, we observe a primary difference between regularity measures, such as ApEn, and moment statistics (e.g., means, standard deviations); namely, moment statistics and their nonparametric counterparts are computed without regard to the *order* of the series to which they are applied. For ApEn, the serial data order is the crucial factor. Additionally, to the best of our knowledge, a direct statistical assessment of joint LH–T network characteristics of either younger or aged men has not previously been accomplished, which cross-ApEn now addresses.

The relative clarity of the young/old separation by ApEn and cross-ApEn takes on enhanced importance in light of reassessment of age-related changes in mean reproductive hormone levels, especially in the case of T, for which there is no clear consensus. Touitou (27) elucidates the controversy for T, in part due to diurnal variations, and time-of-day sampling; e.g., Vermeulen *et al.* reports a decrease in plasma concentrations in the aged (4), based on morning sampling, whereas Harman and Tsitouras (2) show unchanged T levels in the aged group, based on early afternoon studies. Bremner *et al.* (1) showed an overall 24-hr decrease in T concentrations ($P < 0.05$) in the aged, although none between mid-afternoon and late evening. Zumoff *et al.* (6) concluded that total T levels decrease in the aged, though there was considerable overlap between younger and older mean levels (Fig. 1, ref. 6). Nankin and Calkins (3) reported that mean serum total and free T were similar in young and older groups, whereas the mean absolute non-sex hormone-binding globulin-bound T level, as an index of bioavailable T, was significantly lower in the older group. And above, we deduced that for nocturnal observations, mean serum T concentrations are not necessarily distinguishable between very healthy young and older males. For immunoreactive LH, the inference of no overall mean level changes between young and older subjects has been verified by several studies (3, 5, 6), as well as by our analysis. Even so, Warner *et al.* (5) and Urban *et al.* (28) suggested that whereas mean serum immunoreactive and bioactive LH concentrations are age-invariant, the ratio of biological to immunological LH activity decreases in the aged basally or after stimulation, respectively.

Our findings also augment, from a distinct perspective, previous analyses of mean frequency and amplitude characteristics of LH and T release episodes, where elderly men exhibit more frequent (low amplitude) LH secretory bursts, and amplitude-attenuation of T secretory bursts (8). Indeed, the present appraisal by ApEn and cross-ApEn offers a clearer young/old group separation than the aforementioned significant frequency and amplitude differences, especially for T (compare figures 5 and 6 of ref. 8 to figures 1 and 2 herein). In addition, we observe that for parameters such as pulse characteristics and irregularity, a rapid sampling protocol is crucial in some settings (e.g., the 2.5-min sampling paradigm employed herein) to obtain a fine probabilistic description of the contiguous measurement series, whereas for mean level analysis such as those described in the previous paragraph, the

20-min sampling protocol used by many researchers is generally sufficient to provide an accurate estimate (of the mean). In instances in which mean secretory burst frequency and amplitude differences are relatively subtle (e.g., young vs. old LH and T characterizations), inferences from the finer sampling can actually differ qualitatively from those based on significantly coarser sampling (e.g., compare results from ref. 8 to those from coarser sampling protocols employed in refs. 7 and 9).

It is crucial to note the counterpoint between the perspectives of irregularity and that of diurnal variation. Bremner *et al.* (1) observed a clear loss of nyctohemeral T variation in older men, which, as indicated above, was critically noted in establishing the experimental protocol in this study. From a broader perspective, as summarized by Copinschi and van Cauter (29), changes in circadian rhythms toward lower amplitude and/or phase advance with increasing age have been established for the peripheral levels of many other hormones (30–33). However, this change in circadian variation is a very different notion of rhythm change from that of a change in irregularity, as quantified here. Statistically, the extent of 24-hr variation [e.g., quantified by Bremner *et al.* (1) as the highest point minus the lowest point] is basically a measure of overall day-night amplitude. The nyctohemeral changes with age reflect an evolution from a decidedly nonstationary time-series in younger subjects, with pronounced day-night secretory differences, toward a more stationary output, with blunted overall variation. Thus, the very real attenuation of circadian variation, while linguistically often labeled as a "loss of rhythmicity," is more precisely an appropriately quantified change in a notion of overall amplitude or variation, typically applied as a measure of the extent of time-series non-stationarity over a relatively long time period (24 hr). This is juxtaposed with the quantification of changes in *irregularity* or *disorderliness* of serial data seen above and elsewhere (21, 23), inasmuch as irregularity and amplitude measure epistemologically distinct concepts. Nonetheless, the generality of the finding of consistently blunted overall daily circadian variation with increasing aging is most convincing, and the interpretation that this is at least partially due to changes in central nervous system control (29) is thematically consistent with our hypothesis of an age-related increased asynchrony/network dissociation in broad classes of hormonal networks.

Potential Applications. More generally, quantification of signal regularity of both LH and T release, as well as of their mutual relationship and synchrony, could be employed to evaluate a variety of clinical disorders and the efficacy of medical interventions. Furthermore, if a disorder is most prominently characterized by diminution of synchrony, means to restore synchrony may require putatively novel therapeutic strategies. From an experimental perspective, studies will be required ultimately to specify the source(s) determining synchrony, e.g., from the possibilities indicated above, and to perturb this source directly. However, even prior to this identification, one could attempt to restore synchrony obliquely, by providing dual, synchronous administration of agents that respectively induce LH and T production. The point is that if a disorder is biologically determined by an overall system decoupling, a recoupling or reestablishment of temporal concordance may be required to restore physiological function, rather than any means of perturbing a single target node.

Appendix: Mathematical and Statistical Considerations for ApEn and Cross-ApEn

Replicability of Cross-ApEn. To establish a theoretical statistical validity of cross-ApEn as employed here, we studied a range of two-variable vector AR(2) processes, and several types of coupled two-variable analogs of the "variable lag"

---

‡‡Zwart, A. D., Iranmanesh, A. & Veldhuis, J. D., 77th Annual Endocrine Society Meeting, Washington, DC, June 14–17, 1995, abstract OR29-4.

process described below, for each of which we applied cross-ApEn(1, 0.2) to standardized time-series (x–y pair) outputs, 50 replicates of $n = 150$ point data lengths per process. For each process studied, SD (cross-ApEn) was ≤ 0.06, the SD calculated from the cross-ApEn values from the 50 replicates; this imparts reasonable replicability properties similar to that given elsewhere for ApEn (16, 25). This degree of reproducibility is not unexpected, because cross-ApEn is a parameter that aggregates *low*-order, two variable joint distributions at a moderately coarse resolution (determined by r).

Complementarity of ApEn and Cross-ApEn to Correlation and Spectral Analyses. Mathematically, the need for ApEn, and particularly for cross-ApEn, is clarified by considering alternative parameters that might address similar concepts. In comparing two distinct signals or variables (e.g., to assess a degree of synchrony), primary parameters that one might employ include the cross-correlation function, and the cross-spectrum (34), with single variable counterparts the auto-correlation function and the power spectrum. Evaluation of these parameters often is insightful, but with relatively small length data sets statistical estimation issues are nontrivial and, moreover, interpretation of the sample cross-correlation function is highly problematic, unless one employs a model-based prefiltering procedure (ref. 34, p. 139). Furthermore, "standard" spectral estimation methods such as the fast Fourier transform can be shown to be inconsistent and/or so badly biased that findings may be qualitatively incorrect, especially in the presence of outliers and nonstationarities. This is vividly demonstrated by Thomson (35), who recently developed a superior multiple-data-window technique with major advantages compared with other spectral estimation techniques (35, 36). These difficulties are mirrored in the cross-spectrum, in addition to an often serious bias in estimation of coherency in short series.

Most importantly, the autocorrelation function and power spectrum, and their bivariate counterparts, are most illuminating in linear systems, e.g., SARIMA models, for which a rich theoretical development exists (37). For many other classes of processes, these parameters often are much less effective at highlighting certain model characteristics, even apart from statistical considerations. To illustrate this point consider the following simple model, which we denote as a "variable lag" process: this consists of a series of quiescent periods of variable length duration, interspersed with identical positive pulses of a fixed amplitude and frequency. Formally, we recursively define an integer time-valued process denoted VarLag whose i^{th} epoch consists of (a quiescent period of) values $= 0$ at times $t_{i-1} + 1, t_{i-1} + 2, \ldots, t_{i-1} + \text{lag}_i$, immediately followed by the successive values $\sin (\pi/6)$, $\sin (2\pi/6)$, $\sin (3\pi/6)$, $\sin (4\pi/6)$, $\sin (5\pi/6)$, $\sin(6\pi/6) = 0$ at the next 6 time units, where lag_i is a random variable uniformly distributed on (randomly chosen between) the integers between 0 and 60, and t_{i-1} denotes the last time-value of the $(i\text{-}1)^{\text{st}}$ sine pulse. Fig. 3*A* displays representative process output, with Fig. 3*B* a closer view near time $t = 400$. The power spectrum and autocorrelation function calculations shown in Fig. 3 *C* and *E* were calculated from a realization of length $n = 100,000$ points. (The coarse pulse sampling in the above definition was chosen to approximate typical sampling resolution in clinical studies.)

Processes consisting of alternatingly quiescent and active periods seem reasonable for biologists to consider, as they appear to model a wide variety of phenomena. However, within mathematics, such processes with a variable quiescent period are not commonly studied. To the endocrinologist, output from the above model would be considered smoothly pulsatile, especially with the identical pulses; the variable lag process would be most readily distinguished from its constant lag counterpart (for which $\text{lag}_i = 30$ time units for all i) via a decidedly positive SD for the interpulse duration time-series, in the variable lag setting, as opposed to SD $= 0$ (constant interpulse duration) in the constant lag setting. The essential point here, however, is that for VarLag,
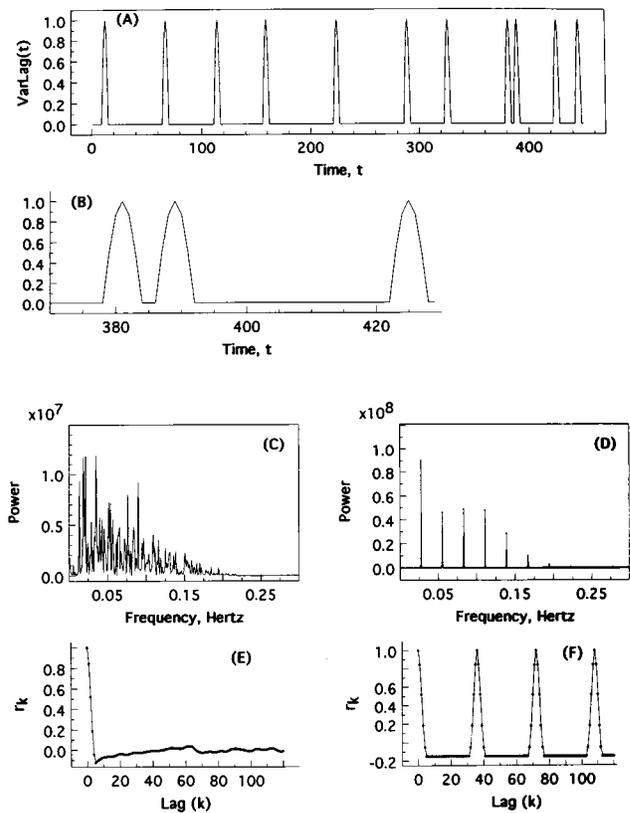

FIG. 3. (*A*) Representative time-series for a "variable lag" sine wave process denoted VarLag (see text for formal definition). (*B*) Close-up view of *A*, near time $t = 400$. (*C*) Power spectrum for VarLag. (*D*) Power spectrum for a constant (fixed) lag analog of VarLag. (*E*) Autocorrelogram corresponding to *C*. (*F*) Autocorrelogram corresponding to *D*. *C*–*F* are all derived from time-series of length $n = 100,000$ points.

the power spectrum and autocorrelation function somewhat confound, as seen in Fig. 3 *C* and *E*. Based on these figures alone, the pulsatile nature of the time-series realizations is hardly evident, and for all $k \geq 6$, the autocorrelation coefficient r_k at lag k is insignificantly different from 0. In contrast, the power spectrum and autocorrelation function confirm the periodicity of the constant lag analogue, shown in Fig. 3 *D* and *F*, as expected. Significantly, the issues here are in the parameters, rather than statistical inadequacies based on an insufficiently long output, or on artifacts (outliers), since Fig. 3 *C*–*F* were derived from calculations based on 100,000 points from a purely theoretical model.

Similar limitations of the spectra and autocorrelation function are inherent to wide classes of mathematical processes. We can construct large classes of variable lag processes simply by considering point processes (38), in which we replace the "point" occurrence by a pulse occurrence, the pulse itself of either a fixed or variable form. The associated counting process could be of any character, and need not be so special as Poisson or renewal (as in the above example). Also, variable lags between events to be compared are the normative case in nonlinear differential equations, in Poisson clumping models (39), and in output variables in typical (adaptive) control theory models and queueing network models. Notably, for many two-dimensional analogs of variable lag processes, and for many two-dimensional systems in which no small set of dominant frequencies encapsulates most of the total power, the cross-spectrum and the cross-correlation function often will similarly fail to highlight episodicities in the underlying model and data, and thus fail to highlight concomitant changes to such episodic components.

In contrast to the autocorrelation function and spectral differences between the above variable lag and constant lag processes, respective ApEn(1, 20% SD) values for the two processes are in close agreement: mean ApEn = 0.195 for the variable lag process, while ApEn = 0.199 for the constant lag setting. This agreement in ApEn values manifests the primary requirement of matching (sub)patterns within data, while relaxing the requirement of a dominant set of frequencies at which these subpatterns occur. The two-variable analogue of ApEn, given by cross-ApEn, similarly enables one to assess synchrony in many classes of models. It thus should not be surprising that in this study cross-correlation (Pearson R) does not show significant group differences, whereas cross-ApEn does.

We emphasize, nonetheless, that Fig. 3 *C–F* neither invalidate spectral power and (lagged) autocorrelation calculations, nor do they violate a properly oriented intuition. The broadbanded spectrum in Fig. 3*C*, and the negligible lagged autocorrelation in Fig. 3*E* for lag $\geq$ 6 time-units, primarily reflect the independent, identically distributed, relatively broad distribution of the variable lag$_i$. Visually, this conforms to viewing Fig. 3*A* from afar, in effect (nearly) ignoring the nature of each pulse, instead de facto primarily focusing on the "random" timing of the peaks as the process of interest. The viewpoint taken by ApEn is thus complementary to the spectrum and correlogram, more de facto focusing on (close-up) similarities between active pulses, e.g., from the perspective given in Fig. 3*B*, while in effect nearly ignoring the nature of the quiescent epoch aspect of the process. The putative utility of ApEn and cross-ApEn to endocrinologists is based on the recognition that in many settings, changes in the episodic character of the *active* periods within pulsatile secretory time-series appear to mark physiologic and pathophysiologic changes; thus, there is a concomitant need for quantitative methods that primarily address this perspective, e.g., ApEn and cross-ApEn.

Simple variable lag processes similar to VarLag would not confound pulse-identification statistics widely used within endocrinology (19). In many settings, such pulse-identification methods are quite sufficient to characterize and distinguish distinct physiologic states. However, in other settings, clear pulse identification appears to be a challenging endeavor, e.g., clinically, for GH time-series in acromegalics (21) and for healthy female rats (23), as well as for the T data studied here and, theoretically, for several distinct classes of mathematical processes (17, 20). Crucially, as well, there has been no quantification of two-variable synchrony from a pulse-identification perspective, except to determine whether or not co-pulsatility of discrete events within concomitant hormone series is nonrandom (40).

1. Bremner, W. J., Vitiello, M. V. & Prinz, P. N. (1983) *J. Clin. Endocrinol. Metab.* **56,** 1278–1281.
2. Harman, S. M. & Tsitouras, P. D. (1980) *J. Clin. Endocrinol. Metab.* **51,** 35–40.
3. Nankin, H. R. & Calkins, J. H. (1986) *J. Clin. Endocrinol. Metab.* **63,** 1418–1420.
4. Vermeulen, A., Ruben, R. & Verdonck, L. (1972) *J. Clin. Endocrinol. Metab.* **34,** 730–735.
5. Warner, B. A., Dufau, M. L. & Santen, R. J. (1985) *J. Clin. Endocrinol. Metab.* **60,** 263–268.
6. Zumoff, B., Strain, G. W., Kream, J., O'Connor, J., Rosenfeld, R. S., Levin, J. & Fukushima, D. K. (1982) *J. Clin. Endocrinol. Metab.* **54,** 534–538.
7. Deslypere, J. P., Kaufman, J. M., Vermeulen, T., Vogelaers, D., Vandalem, J. L. & Vermeulen, A. (1987) *J. Clin. Endocrinol. Metab.* **64,** 68–73.
8. Mulligan, T., Iranmanesh, A., Gheorghiu, S., Godschalk, M. & Veldhuis, J. D. (1995) *J. Clin. Endocrinol. Metab.* **80,** 3025–3031.
9. Veldhuis, J. D., Urban, R. J., Lizarralde, G., Johnson, M. L. & Iranmanesh, A. (1992) *J. Clin. Endocrinol. Metab.* **75,** 707–713.
10. Vermeulen, A., Deslypere, J. P. & Kaufman, J. M. (1989) *J. Clin. Endocrinol. Metab.* **68,** 68–72.
11. Pincus, S. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2297–2301.
12. Pincus, S. M. & Singer, B. H. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2083–2088.
13. Fleisher, L. A., Pincus, S. M. & Rosenbaum, S. H. (1993) *Anesthesiology* **78,** 683–692.
14. Pincus, S. M., Cummins, T. R. & Haddad, G. G. (1993) *Am. J. Physiol.* **264,** R638–R646.
15. Pincus, S. M., Gladstone, I. M. & Ehrenkranz, R. A. (1991) *J. Clin. Monit.* **7,** 335–345.
16. Pincus, S. M. & Huang, W.-M. (1992) *Commun. Stat.-Theory Methods* **21,** 3061–3077.
17. Pincus, S. M. & Keefe, D. L. (1992) *Am. J. Physiol.* **262,** E741–E754.
18. Pincus, S. M. & Viscarello, R. R. (1992) *Obstet. Gynecol.* **79,** 249–255.
19. Urban, R. J., Evans, W. S., Rogol, A. D., Kaiser, D. L., Johnson, M. L. & Veldhuis, J. D. (1988) *Endocr. Rev.* **9,** 3–37.
20. Pincus, S. M. (1994) *Math. Biosci.* **122,** 161–181.
21. Hartman, M. L., Pincus, S. M., Johnson, M. L., Matthews, D. H., Faunt, L. M., Vance, M. L., Thorner, M. O. & Veldhuis, J. D. (1994) *J. Clin. Invest.* **94,** 1277–1288.
22. Siragy, H. M., Vieweg, W. V. R., Pincus, S. M. & Veldhuis, J. D. (1995) *J. Clin. Endocrinol. Metab.* **80,** 28–34.
23. Pincus, S. M., Gevers, E. F., Robinson, I. C. A. F., van den Berg, G., Roelfsema, F., Hartman, M. L. & Veldhuis, J. D. (1996) *Am. J. Physiol.* **270,** E107–E115.
24. Veldhuis, J. D., Liem, A. Y., South, S., Weltman, A., Weltman, J., Clemmons, D. A., Abbott, R., Mulligan, T., Johnson, M. L., Pincus, S., Straume, M. & Iranmanesh, A. (1995) *J. Clin. Endocrinol. Metab.* **80,** 3209–3222.
25. Pincus, S. M. & Goldberger, A. L. (1994) *Am. J. Physiol.* **266,** H1643–H1656.
26. Pincus, S. M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4432–4436.
27. Touitou, Y. (1995) *Horm. Res.* **43,** 12–19.
28. Urban, R. J., Veldhuis, J. D., Blizzard, R. M. & Dufau, M. L. (1988) *J. Clin. Invest.* **81,** 1020–1029.
29. Copinschi, G. & van Cauter, E. (1995) *Horm. Res.* **43,** 20–24.
30. Ho, K. Y., Evans, W. S., Blizzard, R. M., Veldhuis, J. D., Merriam, G. R., Samojlik, E., Furlanetto, R., Rogol, A. D., Kaiser, D. L. & Thorner, M. O. (1987) *J. Clin. Endocrinol. Metab.* **64,** 51–58.
31. Iranmanesh, A., Lizarralde, G. & Veldhuis, J. D. (1991) *J. Clin. Endocrinol. Metab.* **73,** 1081–1088.
32. Tenover, J. S., Matsumoto, A. M., Clifton, D. K. & Bremner, W. J. (1988) *J. Gerontol.* **43,** M163–M169.
33. van Coevorden, A., Mockel, J., Laurent, E., Kerkhofs, M., L'Hermite-Baleriaux, M., Decoster, C., Neve, P. & van Cauter, E. (1991) *Am. J. Physiol.* **260,** E651–E661.
34. Chatfield, C. (1989) *The Analysis of Time Series: An Introduction* (Chapman & Hall, London), 4th Ed.
35. Thomson, D. J. (1990) *Philos. Trans. R. Soc. London A* **330,** 601–616.
36. Kuo, C., Lindberg, C. & Thomson, D. J. (1990) *Nature (London)* **343,** 709–714.
37. Box, G. E. P. & Jenkins, G. M. (1976) *Time Series Analysis, Forecasting and Control* (Holden–Day, San Francisco).
38. Bremaud, J. P. (1981) *Point Processes and Queues: Martingale Dynamics* (Springer, New York).
39. Aldous, D. (1989) *Probability Approximations via the Poisson Clumping Heuristic* (Springer, Berlin).
40. Veldhuis, J. D., Johnson, M. L. & Seneta, E. (1991) *J. Clin. Endocrinol. Metab.* **73,** 569–576