How to Define a Rhythm using a Scientific Methodology by Jean de Prins ,

Observation of all living beings allows us to note that manifestation of activity does not remain constant in time. A simple study can be done by measuring a variable such as temperature, blood pressure, or activity and graphically representing the values measured over the course of time. Often enough, the observer will detect a motif that is repeated more or less uniformly, and will deem it reasonable to qualify it as a "biological rhythm". The duration of this motif may vary within very broad limits, from a few hundredths of a second to a year. Very quickly, many questions are raised for the experimenter. Is the form of this motif uniform? What is the average amplitude of this motif? Is the motif synchronized by the exterior world? Is the rhythm exogenous or endogenous? To reply to all these questions, the observer will wish to characterize the phenomenon with the assistance of objective parameters that provide a maximum amount of information. The estimation of these parameters requires the perfecting of a methodology, based on clear and precise concepts. These concepts will allow the realization of algorithms (or calculation procedures) that, from the values measured, should give a relevant evaluation of the chosen parameters. This procedure will constitute the "treatment of the data", it being understood that the values measured form what is inappropriately known as "data". In fact, these values not raw data and, on the contrary, are often subject to caution. Criticism of them is part of procedure.

I. TREATEMENT OF THE DATA

Treatment of data is a delicate procedure because concepts and algorithms are interdependent. Even if the presentation calls upon mathematical language and is based on mathematical theories, the treatment of the data does not constitute a deductive mathematical theory. In fact, it does not suffice to formulate a hypothesis, even if it is required to be plausible in the selected experimental context. As we will note, difficulties will appear at the time of application. Certain hypotheses, formulated in the strict framework of the mathematics, are neither verified nor verifiable in the experimental circumstances, which are marred by approximations and of "fuzzy" positions. This problem is often avoided, and when it is, results are published as being "mathematically" proven, while no one is capable of establishing that the theory used is relevant. Consider the conditions typically met with in biology. The experimenter observes a variable according to time. This variable measures a number N of pairs of values (y_i, t_i) where y_i is the nth value of the variable measured at the moment t_i. The values of t_i can be judged to be exact. Notice nevertheless that no measure is realized instantaneously. In fact, every value y_i will be linked to a measurement taken over a length of time centered on the moment t_i. In biology, the values y_i will often not be very precise. We will say that they are marred by significant errors. It is convenient to write that $y = Y. + e_i$ or Y_i is the true value, measured by an ideal instrument. E_i, given by the difference between the value measured and the true value, is the error linked to our imperfect methodology. The error e_i can have various origins. Take, for example, the measurement of the temperature of a human subject during a 24 hour period. We can distinguish the following errors:

(a) **Instrumental Errors**. Let us consider the case of a measurement of temperature in a probe whose resistance varies according to its temperature. The value measured will depend on fluctuations of the reference currents and tensions, of the rounded off value selected for its numerical value, etc. In a well conceived instrument, these errors will be chance, of zero average, independent from one measurement to another and comparatively weak. On the other hand, even in a well conceived device the error may vary according to the value of Y_i. For example, in certain cases the magnitude of the error will be proportional to Y_i.

(b) **Errors of transposition**. We wish to measure the "central" temperature of our subject. To this effect, we will place our probe in a given position, while assimilating the temperature of the

probe to our central temperature concept . This transposition is not necessarily adequate: the place may be poorly chosen because it does not correspond to our central temperature concept, or because it is sensitive to parasitic influences; the probe does not follow the temperature variations properly, etc. The resultant error may be significant, fluctuating or systematic.

(c) Errors linked to Disruptive effects. During the 24 hour period, the subject eventually undergoes stress that temporarily increases his central temperature. If the goal of our manipulation is to study a "normal" subject , we may consider that these variations constitute abnormalities. This decision is always difficult, for it is difficult to remain objective in all circumstances.

All determination of parameters implies the usage of a well defined algorithm that, after a number of logical or mathematical operations, gives us one or several numerical values. For these parameters to be relevant, it is necessary to respect certain conditions:

(d) Plausible Hypotheses: : the hypotheses formulated must correspond to plausible suppositions in the observed experimental conditions. Too often the hypotheses are formulated to ensure ease of calculation! It some is for example thus when the observer puts that the independent errors and distributed according to a "normal" (i.e. Gaussian, bell curve) law . An elementary precaution consists in verifying, within the limits of possibility, the hypotheses formulated.

(e) Stability of calculation: it is obvious that the values of the parameters obtained are influenced by the experimental errors. For certain algorithms, all small variations of the initial data generate large variations in the estimated parameters. This position is unacceptable and mathematicians will say that it is a matter of a "poorly put problem". This is the case when we wish to calculate the derivative of a function known by a cloud of pairs of points. All the same, an estimation of the differences between successive experimental values might produce inexact values.

The choice of treatment of data will be comparatively easy when we possess an adequate mathematical model and when measurement errors are weak and of known distribution. Chronobiology that studies biological rhythms does not generally possess any supporting model of the rhythm. What is more, the errors are most often significant. Given this situation, the treatment of data will present many difficulties.

II. HISTORICAL SUMMARY

Although biological rhythms were revealed as early as antiquity, the first rigorous works are situated in the 19th century. Intensive studies date from only a few decades (1)
Let us consider first of all Bünning significant book (2) that appeared in 1958. The words "rhythm" and "oscillation" are often used, but the corresponding concepts are not explicitly defined. On the other hand, many figures represent the experimental results, and make the reader implicitly understand what the terms of "rhythm" and "oscillation" cover. It is a question of an apprenticeship similar to that of the small child who is learning colours. Information is furnished with the assistance of examples, while designating what must be noticed, without justifying it by the usage of a supporting procedure. Nevertheless, it refers to concepts defined in mathematics and physics. Thus the notion of "period" is used, with no argument as if it were obvious. It is useful to know that this notion is based on the mathematical definition of the periodic function. It is a function such that for all values of the time t, we have $f(t + P) = f(t)$. In this equation, P is the value of the period. In everyday language, this relation implies that the repetitive motif always remains the same as itself. For many phenomena in physics or in astronomy, this is indeed the case for long periods of time, and it seems therefore legitimate to use the period concept. But how is one to define a period when variability is significant from one motif to the next? Now that position is the one that we encounter frequently in chronobiology, and is the one that Bünning deals with! We may suppose that an evaluation is realized, while spotting on the graphs the length of time D necessary to observe the number R of repetitive motifs. It may always be done by the relation d/r. But what is the meaning of this magnitude? The notion of "phase" is also introduced by Bünning. It proves significant in chronobiology. Its definition in everyday language is "each of the states of a thing in evolution" (transl.of an entry in the Petit Robert dictionary) Biologists often use it in this sense to discuss the phases of the ovarian cycle or mitosis. But what becomes of this concept in

the case of a temperature cycle or of diastolic pressure? To solve this problem, Bünning refers to the definition of Aschoff, Klotter and Wever: "instantaneous state of an oscillation inside a period, represented by the value of the variable and all its derivatives" This definition is unusable. It is enough to recall that, in current experimental conditions, the calculation of derivative is a "poorly put problem". We need all definitions here! What is more, this definition is understandable with difficulty, and seems to imply the usage of units that are not very current for the measurement of the phase. In fact, this definition is never used. Bünning refers essentially to characteristic observable features of the repetitive motif (beginning, maximum, etc.) to determine the phase in a conventional way. We note, therefore, for the purpose of researchers in the fifties, an apprehension that is more qualitative than quantitative, more intuitive than definite. We have to acknowledge clearly that the most important one, in the eyes of biologists, is to know and understand the origin and the role of the rhythms, the concrete nature of the endogenous or exogenous oscillations and the coupling property between these oscillators and the environment (synchronisation). As soon as attention was attracted to the rhythms, the greatest difficulty is certainly not seeing them!

Nevertheless, this rather imprecise vision does not necessarily satisfy scientists. It will generate the wish to use quantitative statistical methods. Even if the chronobiology must resolve a special and complex problem, it is not the only discipline confronted with oscillations. Many biologists have therefore drawn from the arsenal of the analytic methods available, while if need be adapting these to their specific needs. We develop this aspect in the next section. With the perfecting of the "Cosinor" method in the sixties, a special effort has been made by F. Halberg and his assistants. Algorithms have been published, and a terminology that is meant to pass as rigorous has been proposed. The principle is to adapt a cosinusoidal function with a set period to the data using the method of least squares. The procedure furnishes an estimation of three parameters with their confidence interval, the average level (Mesor), the amplitude and the acrophase. The latter value is the angular phase (in the sense it is used with in physics) maximum of the tailored cosine curve. Since his creation, the Cosinor often was used. The advantages are clear. Biological rhythms are assimilated to cosine functions, which are exemplary periodic functions. All concepts of period , phase and amplitude are directly applicable. The confidence intervals allow us to decide, by using the classical laws of statistics, the presence of a rhythm or not. In the spirit of its authors, the method of Cosinor is always applicable, and therefore constitutes a norm that drives to a standardisation of the results. The representation of these is codified. This normalization allows the inventory of the detected rhythms, and their comparisons. The method is presented as an essential progress that allows a microscopic approach, a lot richer one than the simple visual qualified examination "of macroscopic inspection of the data" (3). It is unfortunately necessary to temper this enthusiasm. In fact this algorithm implies that the observed values are the result of a phenomenon varying in a cosinusoidal manner in time and marred by independent errors and following a distributed "normal" law. Now more often than not, the form of the rhythm is not in the form of a cosine curve. Consequently, either the amplitude, or the acrophase does not constitute a relevant parameter. Thus, the acrophase can be very distant from the maximum of the rhythm being studied, and this can be happen without the confidence interval revealing it. Besides, the interpretation of these intervals is not well defined. Most users imagine that they indicate, among other things, the instabilities both of amplitude and of phase in the rhythms being studied. In reality, these intervals characterize adjustment of the cosine curve, which is very different. Thus, rhythms presenting great variations in length of the successive motifs will be able to give an acrophase interval that is small in comparison with these fluctuations. In short, the confidence intervals will most often be erroneous and misleading. The Cosinor method has its supporters and its detractors. Subjected to many criticisms, the method was improved. Several usage modalities are foreseen (4), and these have the result reduces the advantages of normalization. And despite the improvements, the disadvantages identified previously persist.

III. METHODS OF OSCILLATION ANALYSIS

We will limit ourselves in this section to a reminder of the essential property of the methods used in chronobiology. An article of about one hundred pages was devoted to this subject (5) and readers desiring detailed information itself may refer to it.

a) Fourier Analysis

Fourier analysis is based on the theorem of Fourier that shows that all periodic functions can be represented by the sum of a constant term and of sinusoidal functions of periods P, P/2, P/3,,

p/n....Each of these functions will be known as a "line". The one for period P is the baseline; the others are the harmonics. The set of lines constitutes the spectrum. In practices, an algorithm (Discret Fourier Transforms) allows us to evaluate the spectrum from the experimental values, every line being characterized by its period, its amplitude and its phase. Imagine data corresponding to a rhythm, observed on some strictly repetitive motifs. Used under appropriate conditions, the analysis will give us a spectrum behaving like a baseline whose period corresponds to that of the rhythm, and some harmonics. If we introduce variations in the lengths or in the amplitudes of the successive motifs, we will see the number of lines increase, and the initial spectrum will become more "fuzzy". If, what's more, measurement is marred by errors, we will see a forest of lines emerge. Of course, certain lines will dominate the spectrum, and eventually will be associated with the motive. However, often the harmonics, of weaker amplitudes - but which are important for knowledge of the form - will disappear in the forest of parasitic lines. The interpretation of the results will therefore be very delicate. The major reason for difficulties is obvious: the theory is based on the notion of periodic function while most biological rhythms have a great deal of variability.

Nevertheless, of very numerous based methods on this approach are used (periodogram, spectral covariance, etc.)

b) "Modern" Spectral Analysis

This analysis is based on the adjustment to the data of a recursive equation. Such an equation makes it possible to calculate a term in a sequence of values, based on the knowledge of the preceding values of the set and on a sequence of random values. The successive values calculated in this way will constitute a "random process". For the analysis, we will do the opposite. We will consider that our data constitute a random process, and we wish to estimate the recursive equation generator. Knowledge of it will allow us to establish pertinent parameters. This time, in principle, the rhythms are applied irregularly. But they are applied too much! Thus, the phase concept has totally disappeared. This is very annoying for the study of rhythms synchronized, for example, by the sequence of night and day. Think of the activity of rats, very irregular, but always well greater at night. Nothing, in the selected mathematical approach allows us to characterize, or even to detect this important synchronisation. For the non-synchronized rhythms, the analysis may seem perfect, unfortunately its results will depend in to a great extent on choice of the recursive equation and adjustment criteria. The interpretation will always be delicate, and it is imperative to compare the results based on different recursive equations. Notice that with respect to the biological rhythms it analyzes, Fourier asks for a regularity that is too rigorous, while the modern spectral methods apply an irregularity that is too great. In these cases, it is prudent to use both techniques, to raise the agreements and disagreements to approach the interpretation.

c) Filtering

At the time of filtering, we submit our measurements to sorting. We will envisage two types of filtering: smoothing and the narrow band filtering.

1. Smoothing

Let us assume our measurements are exact. The experimental points would generally be placed on a comparatively simple curve. In reality, the errors will provoke a dispersion of the values around this curve. We obtain a "cloud" of points. To the extent that the filter is well adapted, smoothing tends to re-establish, the original curve from this cloud. The danger lies in extensively modifying the original curve. The non linear filters, using the median, often make it possible to avoid this inconvenience (6). From smoothed values, we can generally study some of the characteristics of every motif (range, maximum, transitions, etc.) Using robust statistics allows us to quantify the dispersion of the estimated values. The advantage of this approach is its conceptual simplicity. We formulate no hypothesis concerning the property of the rhythm. On the other hand, it will often be difficult to distinguish between the clear fluctuations in the rhythm and the imprecision of the estimations.

2. Narrow Band Filtering

This filtering is equivalent to that operated in a radio set when, out of all the existing transmissions, the listener selects the band of frequencies used by the target broadcaster. In the same way, with the

assistance of a numerical filter, we can favour the baseline of a rhythm (demodulation and double demodulation algorithm), or even the baseline and its harmonics (averaging method) But, if the radio transmitters were conceived to favour narrow band filtering, it is not the same for biological organisms. Most often, filtering will diminish the fluctuations in the rhythms. Once more the characteristics obtained might be misleading, and interpretation must be prudent.

d) Chaos Analysis

We will mention the chaos analysis approach, which is currently being studied extensively. Work on non-linear dynamics and chaos has led to the notions of limiting cycle and strange attractor. The interpretation of phenomena presenting intervals of periodicity with many irregularities removed from them can be dealt with using this approach (7) In the current situation, it is essential to measure a huge number of cycles. Results have thus been obtained in the study of the cardiac rhythms. The interpretation is again very delicate, and controversies, numerous. This fascinating subject will certainly be clarified by work in progress.

CONCLUSIONS

On the one hand, the positions met with in the study of rhythms are highly varied and complex. On the other hand, we have access to an arsenal of methodologies, each with its domain of application, its advantages and drawbacks. Usually, no methodology is perfectly suited because we are situated outside its domain of strict validity. Notice in this respect that the errors and defects linked to the treatment of the data are rarely due to a lack of rigour in the mathematical methods used. They are generally caused by a lack of appropriateness to the experimental conditions of the hypotheses formulated. A prudent attitude consists in using several more or less acceptable methods jointly and in comparing in a very critical manner the different parameters obtained. If the results are coherent, we can accept them with relative confidence. In the current state of our knowledge, the question put by our title remains open. There is no single method to define a rhythm. When detection of a rhythm is tenuous, contradictory interpretations may arise. On the other hand, the prudent approach that we advise most often allows a reliable quantitative determination of relevant parameters.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

1. **A. REINBERG. et M. H. SMOLENSKY**, *Biological Rhythms and Medecine*; Springer-Vertag; 1983; pp. 1-12.
2. **E. BUNNING**, *The Physiological Clock*, Third English édition; Springer-Verlag; 1973.
3. **G. CORNÉLISSEN, F. HALBERG, S. S. de la Pena et W. JINYI**, *"The need for both macroscopy and microscopy in dealing with spectral structures*; Chronobiologia; Vol. 15; 1988, pp. 323-327.
4. **C. BINGHAM, B. ARBOGAST, G. CORNÉLISSEN, J.-K LEE et E HALBERG**, *"Inferential statistical methods for estimating and cornparing cosinor parameters*; Chronobiologia; Vol. 9; 1982; pp. 397-439.
5. **J. De PRINS, G. CORNÉLISSEN et W. MALBECQ**, *"Statistical Procedures in Chronobiology and Chronopharmacology"*, Annual Review of Chronopharmocology; Vol. 2; 1986; pp. 27-141.
6. **H. LEE et S. A. KASSAM**, *"Generalized Median Filtering and Related Nonlinear Filtering 'Techniques"* I.E.E.E. Transactions on Acoustics, Speech & Signal Processing; Vol-33; 1985; pp. 672-683.
7. **L. GLASS et M.C. MACKER**, *From Clocks to Chaos*, The Rhythms of Life, Princeton University Press; 1988.