Improving melatonin circadian phase estimates

Eus J.W. Van Someren a,b,*, Elsbeth Nagtegaal c

a Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands
b VU University Medical Center, Departments of Neurology, Clinical Neurophysiology and Medical Psychology, Amsterdam, The Netherlands
c Meander Medical Center, Department of Pharmacy, P.O. Box 1502, 3800 BM Amersfoort, The Netherlands

Received 25 July 2006; accepted 2 March 2007
Available online 9 May 2007

Abstract

The quality and quantity of sleep is to a large extent determined by whether the sleep period is in alignment with the most favorable circadian time window for sleep. Misalignment results in compromised sleep. In order to determine this circadian time window, the 24-h profile of melatonin secretion is generally considered to provide the most optimal estimate. Melatonin secretion occurs only during the night, and several methods to determine its onset and offset markers have been proposed. In spite of the usefulness of determining circadian phase estimates from melatonin, its feasibility is somewhat restricted because the required number of repeated measurements comes at a high cost for compliance and laboratory assays. In addition, the complexity of some of the previously proposed methods to analyze data and obtain phase estimates may require a statistician.

We here propose a set of novel functions to better describe the typical melatonin profile, which usually has a rather fixed baseline level during the day, has differences in the steepness of its rising and falling limbs, and may have a nocturnal plateau or even two peaks instead of one during the night. The functions can easily be fitted, even to incomplete or noisy melatonin data, with the most common statistical software packages, and the resulting parameters give direct information on the mentioned characteristics, which provide important additions to complete the usual restricted information on phase and amplitude.

We show that the proposed curves fit better than single- to three-harmonic cosine curves to the typical melatonin profiles of both healthy subjects (n = 13) and subjects diagnosed with Delayed Sleep Phase Syndrome (DSPS, n = 27), Disorders of Initiating and Maintaining Sleep (DIMS, n = 9), or sleep complaints not otherwise specified (n = 7). Of note, because the functions provide a parsimonious description of the melatonin profile, phase estimates derived from them are more reliable (i.e., robust for noise and data loss). We illustrate that phase estimates deviate on average only by about 10 min in case of the loss of some of the data points and in case of the addition of noise.

Finally, we introduce a sparse-sampling schedule tailored to capture the most important aspects of the melatonin curve. It is shown that such schedule – reducing the number of samples by more than 50% – in combination with the proposed functions results in reliable melatonin onset phase estimates, deviating only about 10 min from estimates based on 24 samples.

The proposed methods strongly contribute to the feasibility, in terms of both cost and analysis availability, for researchers and clinicians to include the most reliable marker of the circadian timing system in their diagnosis and treatment evaluations.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Melatonin; DLMO; Circadian phase; Methodology; Phase marker; Sleep; Circadian rhythm; Curve fitting; Mathematical functions; Sparse sampling; Noise; Data loss

1. Introduction

Most, if not all, body functions show characteristic endogenously generated 24-h rhythms, orchestrated by an internal circadian timing system. Misalignment of these rhythms with the effective light–dark cycle, which
is set both by the environment and the sleep period, has considerable repercussions for human well-being (for definitions of terminology, see list at the end of the paper). Examples of such misalignment are jet lag, the delayed and advanced sleep phase syndromes, and circadian rhythm disorders due to shift work [e.g., 1–3]. In principle, these disturbances can be restored by use of the appropriate zeitgebers, notably bright light and melatonin, and possibly also scheduled food intake, physical activity and temperature pulses. The effect of these zeitgebers follows a very specific phase–response curve. For example, the application of bright light shortly after the minimum of the endogenous core body temperature cycle results in a phase advance, whereas the very same stimulus induces a phase-delay when applied just a few hours earlier. Thus, a few hours off the optimal treatment time may make the difference between efficacy, no effect, or even adverse effects. Consequently, any treatment aimed at amelioration of misalignment symptoms by accelerating phase-angle restoration or supporting entrainment, requires first and foremost an appropriate estimate of the phase of the endogenous circadian timing system. As has been argued before [e.g., 4], evaluation of circadian phase is an important step in the diagnosis and treatment of patients with suspected circadian rhythm sleep disorders. In addition to the circadian phase, the amplitude and daytime levels of the melatonin rhythm may be of interest in elderly subjects and appear to be affected by light exposure as well [reviewed in 5].

There is consensus that the best phase marker presently available can be derived by sampling the 24-h rhythm in plasma or saliva melatonin levels. Daytime melatonin levels are low, and the rapid increase towards the elevated nocturnal levels has shown to be a reliable phase marker. Whereas other phase estimates (e.g., the minimum of the core body temperature rhythm) are highly sensitive to confounding effects of posture, activity and arousal state, melatonin levels are quite robust to such confounders with the exception of light exposure (note, however, that effects of posture [6] and physical activity [7] have been demonstrated). Melatonin may in fact contribute to the circadian rhythms in core body temperature and sleep by promoting heat loss through selective vasodilatation in the skin of the extremities [8]. The resulting much elevated nocturnal skin temperature [9] has been proposed to promote sleep through activation of thermosensitive neurons in sleep-regulating brain areas [10–14]. The endogenous increase of melatonin levels is suppressed by exposure to bright light. Consequently, it is generally accepted that samples should be taken in dim light during the period when melatonin levels are expected to be elevated.

There is, however, less consensus as to which characteristic of the 24-h melatonin profile provides the most robust marker of the endogenous circadian timing system. Even the most frequently used phase marker, known as the Dim Light Melatonin Onset (DLMO) time, has been calculated in many different ways, resulting in different estimates. Examples are the calculation of the interpolated time point associated with the upward crossing of melatonin levels above a fixed threshold [15], or calculation of the interpolated time point where the melatonin level exceeds the mean of a number of previous samples by two of their standard deviations [16]. These methods have as major drawbacks that samples should be taken frequently – for example, half-hourly [e.g., 17] – and that the outcome is sensitive to the peak level reached and to measurement error in single samples surrounding the calculated onset time.

To overcome these drawbacks, it has been proposed to first fit an appropriate curve to the time-series of melatonin samples, and to derive phase estimates from the parameters describing this curve, for example, the time point where the curve crosses one-fourth of the fitted amplitude [18]. Since such curve takes into account all available data points, phase estimates are supposedly more robust and less dependent on error variation of just a few data points. The proposed curves to be fitted to the data range from the very elementary 24-h single cosine to more refined multi-harmonic cosine models and the even more complex models applying locally weighed polynomial regression [19] or Bayesian techniques and pharmacokinetics to estimate differential equations [20,21]. Whereas the simpler models are relatively easy to implement, they are not doing justice to the characteristic profile of melatonin levels. For example, the more or less constantly low daytime baseline melatonin level is poorly fitted by a single cosine curve. Although the addition of harmonics can account for this flat daytime level in a better way, it comes at the cost of introducing artificial side-lobes at the onset and offset of the melatonin peak [see Fig. 2 in Ref. 22]. On the other hand, whereas the application of the more complex models mentioned above results in better description of the melatonin profile, they require a high sampling rate and number of samples due to the large number of parameters to be estimated and are not easily implemented. Consequently, their use is limited to clinicians and researchers with mathematical back-up.

A complicating factor is that especially in field studies, but even in well-controlled laboratory studies, several factors may introduce noise and missing data. Data may be missing due to non-compliance and may contain contaminated samples if abstinence from food and drinks and rinsing of the mouth, is not feasible. In addition, it may be highly unworkable to require hourly or even half-hourly sampling throughout a 24-h period. Given these practical limitations, it is of utmost importance to utilize data-analysis techniques that are robust for noise and missing data. Moreover, such data...
analysis should ideally give a robust estimate if sampling is limited to a much reduced but well-chosen number of time points.

The aim of the present report is threefold. First, we introduce a novel curve, based on previous work of Batschelet [23] and Ruf [24], and demonstrate that it captures the typical melatonin profile with better precision yet more parsimoniously than previously applied curves. Importantly, the curve can be obtained easily using the nonlinear curve fitting procedures available in most standard statistical software packages including SPSS (see Appendix). Second, we demonstrate that phase estimates derived from such fitted curves are more robust to missing data and measurement error than previously applied methods. Finally, we introduce a sampling schedule that is much reduced to a well-chosen number of time points and show that the schedule yields a reliable phase estimate if the proposed curves are fit to such limited but well-chosen number of samples. We believe that these advances could benefit the cost-effectiveness and feasibility of the investigation and treatment of circadian rhythm disturbances in field studies and in less cooperative or frail subjects.

2. Methods

2.1. Subjects

Subjects (n = 56, 36 ± 2 mean ± standard error of the mean [s.e.m.] years of age, 27 males and 29 females) were recruited through general practitioners and advertisements for melatonin profile determination as reported in our previous studies. Subjects were either without complaints (n = 13), or suffered from sleep complaints (n = 41). According to the International Classification of Sleep Disorders, the latter were diagnosed with Delayed Sleep Phase Syndrome (DSPS, n = 27), Disorders of Initiating and Maintaining Sleep (DIMS, n = 9), or sleep complaints not otherwise specified (n = 7). Exclusion criteria were as follows: age under 12 years, prior use of melatonin, liver diseases, renal failure, severe neurological or psychiatric disorders and pregnancy [25]. Of importance to the generalizability of our evaluation, the subjects covered a wide range of melatonin profiles. Individual amplitudes (peak–trough) ranged from 8.7 to 150.5 pg/ml, and peak levels were observed between about 1:00 and 9:00 h.

2.2. Protocol

Twenty-four hour curves of saliva levels of melatonin were assessed under continuous bed rest conditions. Ambient light intensity was less than 100 lux from 08:00 to 18:00 h, and less than 20 lux from 18:00 until 8:00 h. Room temperature was kept within the range of 19–21 °C. Sleep was allowed. Saliva was collected hourly by asking the subject to chew on a cotton Salivet (Sarstedt, Nümbrecht, Germany) for 1 min. The subjects did not smoke or brush their teeth during the study, and did not drink or eat from 15 min before until the end of each sample time.

2.3. Melatonin assay

Saliva samples were kept at 4 °C until the end of assessment, centrifuged at 1000g for 2 min and stored at −20 °C until assay. Aliquots of 400 μl were added directly to the radioimmunoassay (RIA) kit tubes of Buhlmann Laboratories, Basel, Switzerland. The detection limit of this assay was 0.5 pg/ml, with intra-assay and interassay coefficients variation of ±9% and ±12%, respectively.

2.4. Analysis I: Comparing functions to describe the melatonin profile

The first aim of the study was to investigate whether novel cosine-based functions tailored to include aspects of the typical melatonin profile like a fixed baseline, differences in the steepness of its rise and fall, and bimodal peaks, could compete with previously proposed curves. To do so, the goodness of fit (R²) and Akaike Information Criterion [AIC, 26] were calculated. Curves were fitted using the constrained nonlinear regression features of SPSS 10 for Macintosh software (SPSS Inc., Chicago, IL, USA). Since the fitted functions required constraints on the allowed range of parameter values, the Levenberg–Marquardt algorithm could not be applied and the sequential quadratic programming algorithm used instead. R² represents the fraction of the variance in the data points accounted for by the function. It will continue to increase towards 1 if parameters are added to the regression model. However, adding parameters comes at the cost of losing robustness of the functions: if one data point changes, parameter estimates may change considerably. Moreover, the interpretation of parameters may become increasingly complex. The AIC is a method to compare the goodness of fit of several functions while taking the number of parameters into account. The AIC is proportional to the sum of squared residual errors (RSS) but includes a penalty for every parameter additionally included. It is calculated as follows: \( \text{AIC} = n \times \log(\text{RSS}/n) + 2K \), with \( n \) = number of data points and \( K \) = number of parameters. A smaller AIC indicates that less parameters are needed to account for a certain amount of variance in the data.

Four previously proposed functions were fit to all 56 melatonin time-series: (1) the three-parameter, single 24-h cosine function, also known as ‘cosinor'; (2) the five-parameter, two-harmonic cosine function, including
the 24- and 12-h component; (3) the seven-parameter, three-harmonic cosine function, including the 24-, 12- and 8-h component; (4) the baseline cosine function (BCF), originally described by Ruf [24] and applied to human melatonin data by Zhou et al. [27]. In addition, three novel functions were fit. The functions included components of previously described functions [23,24]. Although the functions may appear complex at first sight, their strength lies in the repeated use of the same parameters within the formula, resulting in a parsimonious set of parameters to be estimated, respectively, 5, 5 and 6, for the three functions described below, which is similar to the two-harmonic cosine function for the first two, and still one parameter less than the three-harmonic function for the third, most complex function.

First, the skewed baseline cosine function (SBCF) combines a fixed baseline with skewness of the cosine function, allowing differences in the steepness of the rising and falling limb of the melatonin peak. This can be accomplished by adding a single parameter ($v$) to the baseline cosine function as follows:

$$Y[x] = b + \frac{H}{2(1-c)}(\cos(x - \phi + v\cos(x - \phi)) - c) + |\cos(x - \phi + v\cos(x - \phi)) - c|$$

Here $Y$ represents the predicted melatonin value at time $x$, where $x$ is expressed in radians ($0$ to $2\pi$), representing $0$ to $24$ h. The parameters are directly proportional to the baseline level ($b$), the height ($H$) (i.e., peak level), the width ($c$), phase ($\phi$, in radians) and skewness ($v$) of the peak. The width-parameter ($c$) must be estimated with constraints, allowing it to vary between $-1$ for a regular cosine without any part ‘cut off’ to baseline, through $0$ for a cosine of which the 12 lowest hours are set to the baseline level, up to near $1$ for a cosine with a very narrow peak on a fixed baseline of near $24$-h duration. Likewise, the parameter–parameter ($v$) has to be estimated with constraints, allowing it to vary between $-0.5$ to describe a curve with a slow rise and steep fall, and $0.5$ to describe a curve with a steep rise and slow fall. The function is shown in Fig. 1, second panel.

Second, the bimodal baseline cosine function (Bimodal BCF) combines a fixed baseline with the possibility for a broader and flattened peak or even two peaks in the cosine function to describe evening and morning peaks of melatonin secretion that might represent the separate expressions of rhythms of two oscillators [28]. The bimodality can be accomplished by adding a single parameter ($v$) to the baseline cosine function, as follows:

$$Y[x] = b + \frac{H}{2(1-c)}(\cos(x - \phi) + m\cos(2x - 2\phi - \pi) - c) + |\cos(x - \phi) + m\cos(2x - 2\phi - \pi) - c|$$

The variables and parameters are defined as described above, now with the addition of the parameter ($m$), to allow for bimodality of the peak, which is proportional to the value of $m$. The bimodality-parameter ($m$) has to be estimated with the constraint that it is larger or equal to zero. The function is shown in Fig. 1, third panel.

Third, the bimodal skewed baseline cosine function (Bimodal SBCF) combines the two functions mentioned above, now allowing for both bimodality and skewness:

$$Y[x] = b + \frac{H}{2(1-c)}(\cos(x - \phi + v\cos(x - \phi)) + m\cos(2x - 2\phi - \pi) - c) + |\cos(x - \phi + v\cos(x - \phi)) + m\cos(2x - 2\phi - \pi) - c|$$

The function is shown in Fig. 1, fourth panel.

### 2.5. Analysis II: Comparing robustness of phase markers for data loss and noise

The second aim of the study was to investigate the robustness of several melatonin onset and offset estimates in case of (1) a random loss of data points or (2) random noise added to the samples. To do so, we first calculated several previously proposed phase markers on the original 24 data points of each subject. Subsequently, the same phase markers were calculated after omitting four randomly chosen data points from the original 24. Random selection was done per subject, resulting in different missing time points for each subject. For all phase markers, absolute differences were then calculated between the original estimates and estimates from the incomplete datasets. By comparing the mean absolute difference for all phase markers, it could thus be evaluated how robust they are, that is, how accurate the estimates are in case of data loss. A similar method was used to evaluate robustness of the phase markers for noise on the sampled melatonin values. Noise is likely to occur both due to contamination of saliva samples by previously taken food and drinks, as well as due to the limited accuracy of melatonin assays. Artificial noise was added by scaling every data point with a randomly generated factor between 0.8 and 1.2. As with the random missing data described above, absolute differences were then calculated between the original estimates and estimates from the datasets with added noise. By comparing the mean absolute difference for all phase markers, it could thus be evaluated how robust they are, that is, how accurate the estimates are in case of noisy data. In addition, the reliability (intraclass correlation coefficient) of the estimates from the incomplete and noisy datasets as compared to the original datasets were calculated.

Both raw data-based and curve-based melatonin *onset* phase markers were calculated as previously proposed by others:
Fig. 1. Examples of raw data and fitted curves. Note that of the 24 melatonin samples the 11 open circles indicate the samples included for the dedicated sparse-sampling evaluation. The upper panel shows the baseline cosine function fitted to the data of subject 12. The second panel shows, for the same subject, that the fit improves by including a skewness parameter. Typical of most melatonin profiles, the rising limb is steeper than the falling limb. The third panel shows an example of how the bimodal baseline cosine function better accommodates the profile of subject 39. The fourth panel shows an example of the bimodal skewed baseline function fitted to the profile of subject 56. The fifth panel shows that the three-harmonic cosine function fitted to the same data of subject 56 results in spurious fluctuation between 9:00 and 21:00 h.
(1) On the raw data, a DLMO (DLMO-fixed) was determined using interpolation to find the time that a fixed threshold of 3.5 pg/ml was reached, as originally proposed to be 10 pg/ml for plasma levels by Lewy et al. [15] and adapted to the ±threefold lower levels in saliva [16,29].

(2) On the raw data, a second DLMO (DLMO-mean ± 2SD) was determined using interpolation to find the time that a subject-specific threshold [16,30,31] was reached, defined as the mean ± 2 standard deviations of the three daytime samples taken at 17:00, 18:00 and 19:00 h.

(3) For the fitted curves, the up-cross times as proposed by Klerman et al. [18] were calculated (i.e., the time after which the curve exceeds one-quarter of its amplitude).

Both raw data-based and curve-based melatonin offset phase markers were calculated as previously proposed by others:

(1) On the raw data, SynOff was determined as the last point after which levels began to decline. This marker has been proposed to reflect the offset of melatonin production [32]. We automated its determination by taking the first value in the time-series that exceeded all next five values.

(2) For the fitted curves, the down-cross times as proposed by Klerman et al. [18] were calculated (i.e., the time after which the curve drops below one-quarter of its amplitude).

2.6. Analysis III: Evaluating the robustness of a dedicated sparse-sampling schedule for melatonin

The first aim of the study was to investigate whether novel cosine-based functions tailored to include aspects of the typical melatonin profile like a fixed baseline, differences in the steepness of its rise and fall, and bimodal peaks, could compete with previously proposed curves. Table 1 shows the average $R^2$ and AIC for the different curves. With the exception of the single cosine (‘cosinor’) model, the goodness of fit of all curves was excellent; 92–96% of the variance in the data points was captured by the fitted functions. The AIC indicated that all functions that include a baseline needed considerably less parameters to obtain a certain goodness of fit than the first three functions, which do not include a baseline parameter. The best fitting yet most parsimonious function was the skewed bimodal baseline cosine. A repeated-measures analysis of variance (ANOVA) demonstrated a highly significant overall effect of model choice ($p < 0.0001$), and subsequent t-tests indicated that its $R^2$ was significantly better (higher) than all other curves ($p < 0.01$), and its AIC better (smaller) than all other curves ($p < 0.02$) except for the skewed baseline cosine. The second best fitting curve was the skewed baseline cosine, with significantly better $R^2$ and AIC than all other curves except for the bimodal baseline cosine, which performed almost equally well.

2.7. Analysis IV: Comparing robustness of phase markers for data loss and noise

The second aim of the study was to investigate the robustness of several melatonin onset and offset markers derived from a sparse-sampling schedule, specifically designed to capture the characteristics of the melatonin profile with a limited but well-chosen number of data points. The schedule included four hourly samples taken around the expected onset and increasing limb and three samples taken around the expected decreasing limb and offset. The remaining daytime and nighttime periods are each covered by two more widely spaced sampling times.

The time points are chosen such as to minimally disrupt the normal sleep–wake pattern of most subjects, except for two nocturnal samples. For the present dataset, the following time points were chosen: 20:00, 21:00, 22:00, 23:00, 01:00, 04:00, 7:00, 8:00, 9:00, 13:00 and 16:00 h. As described above, curves were fit and phase markers derived from these limited datasets. For all phase markers, absolute differences were then calculated between the original estimates and estimates from the incomplete datasets. In addition, the reliability (intraclass correlation coefficient) of the estimates from the incomplete datasets as compared to the original datasets were calculated.

3. Results

3.1. Analysis I: Comparing functions to describe the melatonin profile

The first aim of the study was to investigate whether novel cosine-based functions tailored to include aspects of the typical melatonin profile like a fixed baseline, differences in the steepness of its rise and fall, and bimodal peaks, could compete with previously proposed curves. Table 1 shows the average $R^2$ and AIC for the different curves. With the exception of the single cosine (‘cosinor’) model, the goodness of fit of all curves was excellent; 92–96% of the variance in the data points was captured by the fitted functions. The AIC indicated that all functions that include a baseline needed considerably less parameters to obtain a certain goodness of fit than the first three functions, which do not include a baseline parameter. The best fitting yet most parsimonious function was the skewed bimodal baseline cosine. A repeated-measures analysis of variance (ANOVA) demonstrated a highly significant overall effect of model choice ($p < 0.0001$), and subsequent t-tests indicated that its $R^2$ was significantly better (higher) than all other curves ($p < 0.01$), and its AIC better (smaller) than all other curves ($p < 0.02$) except for the skewed baseline cosine. The second best fitting curve was the skewed baseline cosine, with significantly better $R^2$ and AIC than all other curves except for the bimodal baseline cosine, which performed almost equally well.

3.2. Analysis II: Comparing robustness of phase markers for data loss and noise

The second aim of the study was to investigate the robustness of several melatonin onset and offset estimates in case of a random loss of data points or random noise added to the samples. Table 2 gives an overview of the average absolute deviation from the originally

<table>
<thead>
<tr>
<th>Function (parameters)</th>
<th>$R^2$</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single cosine (3)</td>
<td>0.76 ± 0.10</td>
<td>47.5 ± 14.2</td>
</tr>
<tr>
<td>Two harmonic cosine (5)</td>
<td>0.92 ± 0.10</td>
<td>38.0 ± 13.4</td>
</tr>
<tr>
<td>Three harmonic cosine (7)</td>
<td>0.93 ± 0.09</td>
<td>39.1 ± 13.7</td>
</tr>
<tr>
<td>Baseline cosine (4)</td>
<td>0.94 ± 0.09</td>
<td>31.7 ± 14.3</td>
</tr>
<tr>
<td>Skewed baseline cosine (5)</td>
<td>0.95 ± 0.09</td>
<td>30.7 ± 14.9</td>
</tr>
<tr>
<td>Bimodal baseline cosine (5)</td>
<td>0.95 ± 0.06</td>
<td>31.2 ± 14.4</td>
</tr>
<tr>
<td>Skewed bimodal baseline cosine (6)</td>
<td>0.96 ± 0.05</td>
<td>30.0 ± 14.5</td>
</tr>
</tbody>
</table>
calculated phase onset markers in case of random loss of four data points and in case of random noise. In case of random datapoints, the BCF onset phase estimate remained closest to the original, closely followed by the SBCF onset phase estimate. Paired t-tests did not reach significance in the comparison between the BCF-based and non-BCF-based onset phase estimates regarding their resemblance to the original. However, the reliability of all BCF-based onset estimates was higher than the upper limit of the 95% confidence interval of the reliability of all non-BCF-based onset estimates.

In case of increased noise in the data, the SBCF and Bimodal BCF onset phase estimate remained closest to the original. Paired t-tests showed a significantly worse resemblance to the original for the F3-based onset phase estimate as compared to the BCF- ($p < 0.01$), SBCF- ($p < 0.001$) and Bimodal BCF- ($p < 0.005$) based onset phase estimates, and tended to do worse than the Bimodal SBCF as well ($p < 0.07$). The mean ± 2SD DLMO onset phase estimate showed a significantly worse resemblance to the original than the SBCF did ($p < 0.03$) and tended to do worse than the Bimodal BCF ($p < 0.08$) as well. The fixed threshold DLMO onset phase estimate was also rather robust for an increase in noise, with the notable exception of a few cases, hence the much larger standard deviation of the deviation. The reliability of all BCF-based onset estimates was higher than the lower limit of the 95% confidence interval of the reliability of all non-BCF-based onset estimates except for the comparison between the reliabilities of the F3-based onset phase estimate and the Bimodal SBCF-based onset estimate.

In a similar way, Table 3 provides the average absolute deviations from the originally calculated phase offset markers. In case of random loss of data points, the BCF offset phase estimate remained closest to the original, closely followed by the SBCF offset phase estimate. Paired t-tests, however, did not reach significance in the comparison between the BCF-based and non-BCF-based onset phase estimates regarding their resemblance to the original. Reliability coefficients did not differ either.

In case of increased noise in the data, the SBCF offset phase estimate remained by far the closest to the original estimate. Paired t-tests indicated significantly (all $p < 0.0001$) worse resemblance to the original for SynOff as compared to all other offset phase estimates. In addition, the F3-based offset phase estimate had a significantly ($p < 0.02$) worse resemblance to the original than the SBCF-based offset estimate did and tended to do worse than the Bimodal BCF ($p < 0.08$). The reliability of the SynOff estimates was lower than the lower limit of the 95% confidence interval of the reliability of all other offset estimates.

3.3. Analysis III: Evaluating the robustness of a dedicated sparse-sampling schedule for melatonin

The third aim of the study was to investigate the robustness of melatonin onset and offset estimates in

---

### Table 2

<table>
<thead>
<tr>
<th>Onset phase markers</th>
<th>Random missing</th>
<th>Random noise</th>
<th>11 samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLMO fixed</td>
<td>0:13 ± 0:30 (0.87)</td>
<td>0:08 ± 0:23 (0.90)</td>
<td>0:15 ± 0:36 (0.85)</td>
</tr>
<tr>
<td>DLMO mean ± 2SD</td>
<td>0:16 ± 0:37 (0.92)</td>
<td>0:14 ± 0:32 (0.94)</td>
<td>0:19 ± 0:28 (0.93)</td>
</tr>
<tr>
<td>F3 1/4 amplitude upcross</td>
<td>0:14 ± 0:34 (0.93)</td>
<td>0:12 ± 0:11 (0.99)</td>
<td>0:15 ± 0:13 (0.83)</td>
</tr>
<tr>
<td>BCF 1/4 amplitude upcross</td>
<td>0:10 ± 0:14 (0.99)</td>
<td>0:10 ± 0:09 (0.99)</td>
<td>0:10 ± 0:09 (0.99)</td>
</tr>
<tr>
<td>SBCF 1/4 amplitude upcross</td>
<td>0:10 ± 0:19 (0.98)</td>
<td>0:09 ± 0:09 (1.00)</td>
<td>0:12 ± 0:24 (0.96)</td>
</tr>
<tr>
<td>Bimodal BCF 1/4 amplitude upcross</td>
<td>0:13 ± 0:23 (0.96)</td>
<td>0:08 ± 0:07 (0.99)</td>
<td>0:12 ± 0:17 (0.98)</td>
</tr>
<tr>
<td>Bimodal SBCF 1/4 amplitude upcross</td>
<td>0:11 ± 0:19 (0.97)</td>
<td>0:09 ± 0:12 (0.99)</td>
<td>0:13 ± 0:19 (0.97)</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Offset phase markers</th>
<th>Random missing</th>
<th>Random noise</th>
<th>11 samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SynOff</td>
<td>0:13 ± 0:41 (0.94)</td>
<td>0:53 ± 1:13 (0.74)</td>
<td>0:57 ± 1:05 (0.76)</td>
</tr>
<tr>
<td>F3 1/4 amplitude downcross</td>
<td>0:14 ± 0:26 (0.96)</td>
<td>0:11 ± 0:08 (0.99)</td>
<td>0:30 ± 0:56 (0.53)</td>
</tr>
<tr>
<td>BCF 1/4 amplitude downcross</td>
<td>0:12 ± 0:23 (0.96)</td>
<td>0:11 ± 0:10 (0.98)</td>
<td>0:22 ± 0:28 (0.93)</td>
</tr>
<tr>
<td>SBCF 1/4 amplitude downcross</td>
<td>0:13 ± 0:24 (0.96)</td>
<td>0:04 ± 0:04 (0.99)</td>
<td>0:23 ± 0:43 (0.85)</td>
</tr>
<tr>
<td>Bimodal BCF 1/4 amplitude downcross</td>
<td>0:16 ± 0:28 (0.95)</td>
<td>0:14 ± 0:13 (0.98)</td>
<td>0:25 ± 0:42 (0.87)</td>
</tr>
<tr>
<td>Bimodal SBCF 1/4 amplitude downcross</td>
<td>0:17 ± 0:44 (0.90)</td>
<td>0:13 ± 0:14 (0.98)</td>
<td>0:28 ± 0:44 (0.85)</td>
</tr>
</tbody>
</table>
case of their derivation from a dedicated sparse-sampling schedule of 11 instead of 24 hourly samples. The BCF onset (Table 2) and offset (Table 3) estimates remained closest to the original, closely followed by phase estimates based on the skewed BCF and bimodal BCF. Of note, the sparse-sampled onset estimates deviated much less from the original than the sparse-sampled offset estimates and had higher reliabilities as compared to the original estimates based on all 24 samples.

As to the onset phase, paired t-tests showed a significantly worse resemblance to the original for the mean ± 2SD DLMO-based estimates as compared to the BCF- (p < 0.02), Bimodal BCF- (p < 0.02) and Bimodal SBCF- (p < 0.04) based onset phase estimates. The BCF-based onset phase estimate also remained closer to the original as compared to the F3-based onset phase estimate (p < 0.01). The reliability of all BCF-based onset estimates was higher than the upper limit of the 95% confidence interval of the reliability of all non-BCF-based onset estimates.

As to phase offset, paired t-tests indicated significantly (all p < 0.01) worse resemblance to the original for SynOff as compare to all other estimates. The reliability of all BCF-based onset estimates was higher than the upper limit of the 95% confidence interval of the reliability of the F3-based offset estimate. Moreover, the reliability of the BCF- and Bimodal BCF-based onset estimates was higher than the upper limit of the 95% confidence interval of the reliability of the SynOff estimate.

4. Discussion

The melatonin rhythm has been recognized to provide one of the best accessible and reliable views on the circadian timing system. In the present paper, we first investigated how the rhythm could best be described in a function. Subsequently, we investigated how missing data and noise affect the reliability of several phase estimates of melatonin onset and offset. Finally, it was evaluated whether reliable phase estimates can be obtained using a cost-effective sparse-sampling protocol in combination with the newly proposed curve fitting functions.

As to the mathematical function best fitting the typical melatonin profile, we showed that the baseline cosine function as previously proposed by Ruf [24] provides a much better and more parsimonious description of the melatonin profile than a regular cosine with up to three harmonics can provide. An even better, yet parsimonious description can be obtained by including in the baseline cosine function parameters for skewness and bimodality. In contrast to the cumbersome interpretation of the parameters of a multi-harmonic cosine description, the skewed baseline cosine function here proposed provides straightforward parameters that are directly related to the typical properties of the melatonin profile that may be of importance to discriminate groups of subjects and treatment efficacy. First, the function provides a single peak phase estimate \( \phi \). Second, the baseline parameter \( b \) describes the average level during the daytime, when melatonin production is supposed to be absent. This parameter may prove of value in studies in aging and dementia, where increased daytime levels have been observed in several studies [reviewed in 5,33]. The baseline parameter provides an easy and precise quantification of the daytime level. Third, the parameter \( H \) directly describes the height of the melatonin peak above the baseline. Fourth, the parameter \( c \) is directly proportional to the width of the melatonin peak (i.e., the duration above baseline). The hours above baseline are simply calculated as \( 24 \times \arccos(c)/\pi \). Fifth, the skewness parameter \( r \) is directly proportional to the steepness-ratio of the increasing and decreasing limbs of the melatonin curve. It is conceivable that this parameter can provide insight into its metabolism: in case of a reduced metabolism, as may be the case with compromised liver and kidney function (e.g., at high age, the decrease to baseline may be expected to slow down).

For the bimodal cosine function and skewed bimodal baseline function, the parameter \( m \) (for bimodality) interacts with other parameters to describe the profile, and, therefore, the interpretation of the peak width, skewness and amplitude is less straightforward than for the skewed baseline cosine function. However, these values can relatively easily be retrieved by calculating the minute-by-minute values of the curve and their derivatives. A Microsoft Excel Sheet to do so is available on request.

A previously recognized issue in fitting a curve to a melatonin profile is that low amplitude profiles may be fitted less accurately than high amplitude profiles. This may lead to incorrect interpretations of, for example, increased variability in low secretors. It is, therefore, notable that the curves we recommend here fit equally well to low and high melatonin secretors. Since we included all datasets, no matter their amplitude, we had individual ranges (peak–trough) that varied from 8.7 to 150.5 pg/ml, allowing us to evaluate the relation between the peak–trough range and goodness of fit. In further support of the proposed methods, there was no significant correlation between the individual range and the \( R^2 \) of the recommended curves (BCF \( r = 0.15, p = 0.28; \) SBCF \( r = 0.13, p = 0.34; \) Bimodal BCF \( r = 0.19, p = 0.15; \) Bimodal SBCF \( r = 0.17, p = 0.21 \)). Thus, the BCF-based curves fit about equally accurately for low and high melatonin secretors. The dependency of the goodness of fit on the individual melatonin range tended to be slightly higher for the single, dual and triple-harmonic (non-BCF-based) cosine fits (respectively,
An important related issue is whether the proposed curves fit profiles of the different patient groups and controls equally well. We, therefore, made, for each curve type, comparisons of average goodness of fit between the different groups. An overall repeated-measures ANOVA suggested a trend for an effect of group on $R^2$ ($p = 0.09$). Post hoc $t$-tests indicated only marginal differences between two of the four groups: controls and DSPS subjects, on their average $R^2$ for the two-harmonic cosine ($p = 0.04$), three-harmonic cosine ($p = 0.04$), bimodal baseline cosine ($p = 0.03$) and skewed bimodal baseline cosine ($p = 0.02$). These curves seemed to fit slightly better in DSPS subjects than in control subjects. Closer examination of the reason for the rather surprising finding indicated a single very poor fit on most curves for one subject. Exclusion of this subject dismissed any group effect on $R^2$ ($p = 0.27$).

As to the effect of missing data and noise on the reliability of several phase estimates of melatonin onset and offset, we investigated two specific cases. We first applied noise on all data points by multiplying them with a factor that varied randomly between 0.8 and 1.2. Second, we forced a random 4 out of 24 data points to be missing. In both cases, the melatonin onset phase estimates were on average within 10 min from the original estimates on the 24 raw data points. Melatonin offset phase estimates deviated slightly more, with the notable exception of the robustness of the skewed baseline cosine function for added noise. The newly proposed functions generated offset phase estimates that were in general more robust for noise and missing data than previously used methods.

As to the evaluation of the reliability of phase estimates derived from the newly proposed curves fitted on 11 well-chosen samples instead of on all 24 samples, we demonstrated that melatonin onset estimates deviated on average slightly more than 10 min whereas offset estimates deviated about twice as much. The baseline cosine function performed best closely followed by its extension with either a skewness or a bimodality-parameter.

An important issue is the possibility that the curves based on the specific 11-sample schedule fitted differentially accurately for the different groups of subjects. An overall post hoc ANOVA on $R^2$'s, however, indicated that this was not the case ($p = 0.21$). A related issue is whether the sparse-sampling protocol is equally reliable in groups of subjects with different melatonin onset and offset phases. An overall ANOVA over all groups and phase estimate deviances indicated no group differences here either ($p = 0.73$), indicating that estimates from the sparse-sampling protocol approach estimates from a full sampling protocol equally well in controls and different patient groups. This robustness is advantageous from a practical point of view, since the actual curve (and phase) is not known a priori if a subject reports to the clinic with suspicion of DSPS. Thus, we think it is neither necessary nor desirable to move around the window of the sparse-sampling schedule given its high level of robustness to different profiles in different groups. The schedule in combination with the imposed restrictions on the waveform is apparently robust for different phases. This can be understood from the fact that a reduction of time points prior to onset or early after onset could be compensated by an increased number of time points to describe the rise, peak or drop area.

A further issue of the sampling schedule is that – even though most samples are taken during normal wakefulness – two nocturnal awakenings are still necessary. Since setting an alarm to obtain these samples disrupts sleep, combining sleep evaluation and melatonin evaluation in the same night would not be advised. Two alternatives seem feasible if both sleep and melatonin are of interest to the clinician or researcher. First, one could do a polysomnography on a first night, then start with daytime melatonin samples, and obtain the two nocturnal samples only on the second night. An alternative that in our hands has worked well at least in elderly subjects is not to ask them to set an alarm, but just take a saliva sample at any time they wake up at night, which is often the case at least once or twice for a toilet visit. Subjects enter the salivette in a dedicated in-house built container that automatically logs the sample time, thus freeing the subjects from the task of writing down the sample time. We ask subjects to do so during two subsequent nights and adhere to the scheduled sampling during the day in between. This approach has resulted in the availability of about three to four nocturnal samples in most subjects.

For the curve-based phase estimates, the up-cross and down-cross level were arbitrarily chosen at one-fourth of the amplitude of the curve, as proposed previously [18]. The baseline functions also allow for determination of the very time points where the curves start to deviate from the baseline, which may be of value. However, a post hoc investigation the robustness of these phase estimates for added noise and decimation of the dataset indicated that they were slightly less robust than the one-fourth up-cross and down-cross times. The latter thus appear a good choice.

In general, the reliability of the onset and offset markers derived from fitted functions was better than the reliability of markers based on interpolation (DLMO-fixed, DLMO-mean + 2SD, SynOff). This conclusion was previously reached also by Klerman et al. [18]. The present study adds to the work of Klerman et al. by demonstrating that further improvements can be reached by fitting baseline cosine functions rather than multi-harmonic cosine functions. A disadvantage of, for example, the
three-harmonic cosine function is that it has to generate ‘lobes’ over periods of time in which there is in fact no biological or empirical reason to believe they are present, as shown in Fig. 1, bottom panel, between 9:00 and 21:00 h. The same data appear better, yet more parsimoniously, described by the bimodal skewed baseline function as shown in the fourth panel of Fig. 1.

In summary, novel functions were proposed and validated to describe the melatonin profile in more detail than previously proposed functions. Importantly, the functions are more parsimonious, and consequently rather robust to noise and data loss. They fit the melatonin profiles of low and high secretors equally well. We demonstrated that it is even possible to obtain reasonable estimates if a dedicated sparse-sampling protocol is used, which greatly reduces the cost and increases the feasibility. For sparse-sampling studies, we primarily advise the five-parameter skewed baseline cosine function, with its easily interpreted parameters. It should be kept in mind that the proposed sparse-sampling protocol provides much more reliable melatonin onset phase estimates than melatonin offset phase estimates. For full-sampling studies, we advise the six-parameter bimodal skewed baseline cosine function, better describing separate evening and morning peaks or a flattened melatonin peak profile, as may be the case in studies on extended periods of sleep and darkness. The proposed functions are not only suitable for the robust determination of a melatonin onset or offset phase. Another advantage is that the estimated parameters may give important information on the melatonin curve other than the most often used melatonin onset only. Determination of the peak width, the steepness-ratio of the rising and falling limb, the baseline, and the bimodality may give new insights into circadian function in health and disease.

Definitions of terminology as used for the present paper

Phase advance: change towards an earlier timing – relative to the previous timing or to the fixed environmental 24-h clock time – of a behavioral or physiological process, e.g., a change towards an earlier occurrence of the midpoint of sleep or of the onset of the rise plasma melatonin level.

Phase delay: change towards later timing – relative to the previous timing or to the fixed environmental 24-h clock time – of a behavioral or physiological process; e.g., a change towards a later occurrence of the midpoint of sleep or of the onset of the rise plasma melatonin level.

Zeitgeber: literally ‘time-giver’. Any signal that acts on a biological clock as a time cue that has the capacity to advance or delay its rhythm and thereby to entrain it to a periodic occurrence of the signal. Known zeitgebers are light, melatonin and physical activity.

Entrainment: enforcing the fluctuations in behavioral or physiological processes to occur in synchrony with the periodicity by which zeitgebers occur.

Phase–response curve: graphical representation of how the phase change (hours of advance or delay) elicited by a zeitgeber depends on the time that the zeitgeber is given.

Phase angle: the relative timing of two processes (behavioral, physiological or environmental), e.g., the difference (in hours) between the midpoint of sleep and the onset of the rise plasma melatonin level.

Misalignment: a phase angle between two processes that does not optimally support the a physiological process or behavior, e.g., sleep maintenance is sub-optimal when the sleep period is chosen during the circadian phase that core body temperature is rising.

Acknowledgements

Funding by NWO-ZonMw, The Hague (projects SOW 014-90-001 and Innovation Grant 016.025.041); EU FP6 Sensation Integrated Project (FP6-507231).

Appendix A

Below, the SPSS syntax is given for the four baseline cosine functions described in the present paper. The following abbreviations are used for the variables and parameters: melatonin = the observed melatonin values, x = the corresponding sampling time, expressed in radians, b = baseline, h = height, f = phase (in radians), c = width-parameter, v = skewness-parameter, m = bimodality-parameter. For each model, the first line specifies the start values. In the examples, the start values result in optimization of the model after a first fit trial with a baseline level b of 1, a height h of 50, an acrophase f at 0.6 rad (2.17 h), a width c of 0.2 (10.27 h), a skewness v of 0.3 rad (steeper up than down), and a bimodality m of 0.3. It is advised to run curve fits with different sets of start values, to increase the probability to find the best fitting model. The second line specifies the model. Note that for the bimodal functions the value \( \pi \) is not directly available in SPSS, and should be calculated as \( 4 \ast \text{ARTAN}(1) \). The \( /\text{BOUNDS} \) line specifies the constraints.

* Baseline Cosine Function

MODEL PROGRAM b = 1 h = 50 f = 0.6 c = 0.2. 
COMPUTE PRED_ = b + h * (cos(x - f) - c + ABS(cos(x - f) - c))/(2 * (1 - c)).
CNLR melatonin
/PRED PRED. 
/BOUNDS c < 1; c > -1; h >= 0; b >= 0; f >=0 
/CRITERIA STEPLIMIT 2 ISTEP 1E + 20.
References


