PERGAMON

Psychoneuroendocrinology 28 (2003) 916–931

PNEC

www.elsevier.com/locate/psyneuen

Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change

Jens C. Pruessner [a,*], Clemens Kirschbaum [b], Gunther Meinlschmid [c], Dirk H Hellhammer [c]

[a] *McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada*

[b] *Institute of Experimental Psychology II, University of Düsseldorf, Düsseldorf, Germany*

[c] *Department of Psychobiology, University of Trier, Trier, Germany*

Received 19 April 2002; received in revised form 21 August 2002; accepted 23 August 2002

Abstract

Study protocols in endocrinological research and the neurosciences often employ repeated measurements over time to record changes in physiological or endocrinological variables. While it is desirable to acquire repeated measurements for finding individual and group differences with regard to response time and duration, the amount of data gathered often represents a problem for the statistical analysis. When trying to detect possible associations between repeated measures and other variables, the area under the curve (*AUC*) is routinely used to incorporate multiple time points. However, formulas for computation of the *AUC* are not standardized across laboratories, and existing differences are usually not presented when discussing results, thus causing possible variability, or incompatibility of findings between research groups. In this paper, two formulas for calculation of the area under the curve are presented, which are derived from the trapezoid formula. These formulas are termed 'Area under the curve with respect to increase' (AUC_I) and 'Area under the curve with respect to ground' (AUC_G). The different information that can be derived from repeated measurements with these two formulas is exemplified using artificial and real data from recent studies of the authors. It is shown that depending on which formula is used, different associations with

* Corresponding author. Tel.: +1-514-398-8330; fax: +1-514-398-8948.
E-mail address: jens@bic.mni.mcgill.ca (J.C. Pruessner).

doi:10.1016/S0306-4530(02)00108-7

other variables may emerge. Consequently, it is recommended to employ both formulas when analyzing data sets with repeated measures.

Keywords: Area under the curve; Repeated hormone measures; Total concentrations; Changes over time

1. Introduction

The computation of the area under the curve (*AUC*) is a frequently used method in endocrinological research and the neurosciences to comprise information that is contained in repeated measurements over time. Depending on the nature of the study, it serves a variety of different purposes. In clinical trials, the *AUC* can be employed to monitor the effects of a specific medication over the trial period. In endocrinological studies, the *AUC* is used to estimate ultradian and circadian changes of hormones, and to assess the overall secretion over a specific time period. In pharmacological studies, the *AUC* is useful to evaluate dose/response relationships (Ghizzoni et al., 1994; Maes et al., 1994; O'Brien et al., 1996). The computation of the *AUC* allows the researcher to simplify the statistical analysis and increase the power of the testing without sacrificing the information contained in multiple measurements.

However, despite the proven usefulness of this method, its application across laboratories is limited, for a number of reasons. First, different formulas are used by different laboratories to derive the *AUC* from any given dataset, thus compromising the comparability of their findings. Second, the different formulas used are usually not explicitly elaborated or listed in the papers, thus making it impossible to compare the computation of the *AUC* itself. It is maybe because of these reasons that despite the usefulness of the method, some research groups do not refer to *AUC* but instead refer to some key time points for evaluation of their data (Gormley et al., 1992; Tucci et al., 1996). It can be assumed that researchers might be more easily convinced to employ *AUC* in their statistical analysis if a standardized, easy to apply formula was available.

Deriving formulas for computation of the *AUC* also depends on the information the researcher is interested in. A data set comprised of repeated measurements over time contains at least two different sorts of information. First, it contains the information whether any *changes* occurred over time (was there a change in the events being quantified in the dependent variable during the observation period?). Second, each data set also allows assessing the *overall intensity* at which the recorded events occurred. It is easy to imagine studies with repeated measures where either one, or both parameters are of main interest for the researcher (e.g.: Did the medication have an overall effect compared to the control group? Did a habituation of the effect of the medication occur during the trial period?)

These two different sets of information can best be described with two different formulas, which are outlined in this manuscript. One is tentatively called 'Area under the curve with respect to ground' (AUC_G), whereas the second formula is termed 'Area under the curve with respect to increase' (AUC_I). By using an artificial dataset,

together with a recently acquired endocrinological dataset, the benefits of employing both formulas for statistical analysis are demonstrated.

2. Theory and Methods

The different formulas for the area under the curve can be derived from the trapezoid formula (Reinhardt and Soeder, 2001). A typical trapezoid separated into triangles and rectangles is illustrated in Fig. 1.

The information needed in order to calculate the formula consists of (a) the measurements themselves and (b) the time distance between the measurements. In this example with a group containing six repeated measures, the measurements have been named m_1 through m_6, and the time distances between these measures have been named t_1 through t_5. Formula (1) serves as an index of the area under the curve. It is the formula that calculates the total area under the curve of all the measurements as the area of interest. It thus takes into account the difference between the single measurements from each other (i.e., the change over time) and the distance of these

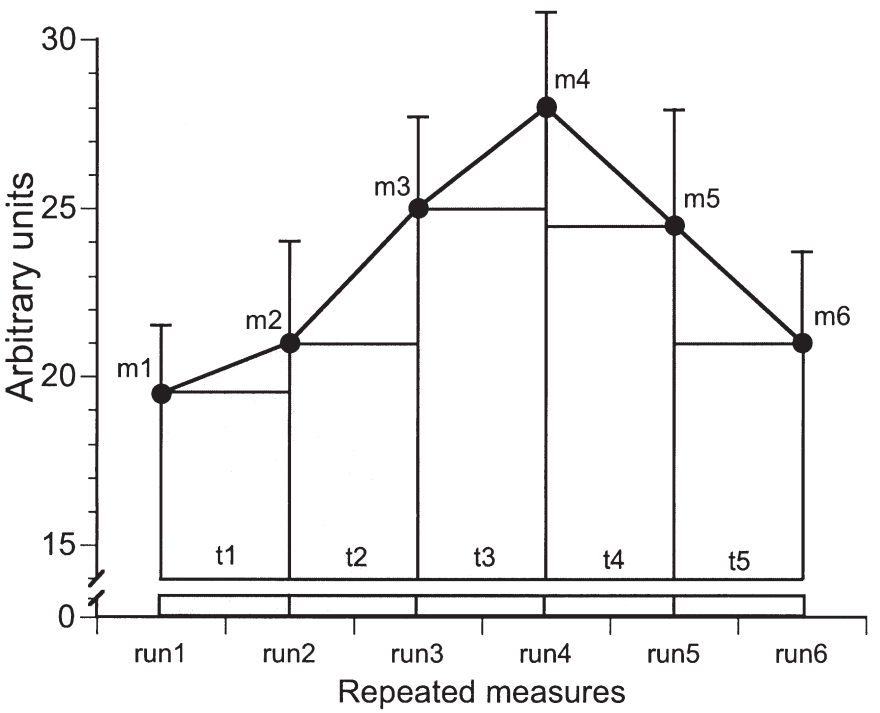

Fig. 1. Time course of an artifical dataset with six measurements; the triangles and rectangles illustrate the composition of the area under the curve with respect to ground (AUC_G). m_1 to m_6 denote the single measurements, and t_1 to t_5 denote the time interval between the measurements. Note that although in this example, the time interval between the measurements is identical for all observations, individual time intervals can vary depending on the study.

measures from the ground, or zero (i.e., the level at which the changes over time occur). This formula is the summation of the trapezoids as illustrated in Fig. 1. Since it calculates the area under the curve with respect to the ground, it has been named AUC_G.

$$AUC_G = \frac{(m_2 + m_1)\cdot t_1}{2} + \frac{(m_3 + m_2)\cdot t_2}{2} + \frac{(m_4 + m_3)\cdot t_3}{2} \\ \frac{(m_5 + m_4)\cdot t_4}{2} + \frac{(m_6 + m_5)\cdot t_5}{2} \quad (1)$$

with t_1 to t_5 denoting the distance between the measurements and m_1 to m_6 representing the single measurements.

Note that formula (1) remains valid independent of the changes over time of the measurements, i.e. independent of increases or decreases over time. Formula (1) can be summarized and becomes formula (2):

$$AUC_G = \sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i)\cdot t_i}{2} \quad (2)$$

with t_i denoting the individual time distance between measurements, m_i the individual measurement, and n the total amount of measures.

Formula (2) is independent of the total number of measurements and can be used with any number of repetitions. Given the case that the time distance between the measurements is identical (e.g., six measurements over fifty minutes with ten-minute intervals between measurements), formula (2) can be further reduced. Being constant, the time distance measure can be removed from the equation, leading to formula (3):

$$AUC_G = \sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i)}{2} \quad (3)$$

with m_i denoting the individual measurement, and n the total number of measurements.

It has to be noted that the result of formula (3) is no longer a true area under the curve, since in this formula the time distance is set to '1' in order to simplify it. In cases where the time interval is constant but not '1' (seconds, minutes, etc.), the resulting values must be regarded as a linear transformation of the area under the curve. However, its use for further statistical analysis is of course unrestricted, since it correlates to 1 with the true area under the curve derived from formula (2).

Besides AUC_G, there is another area of interest that can be calculated from repeated measurements. Figure 2 illustrates this area, again with the measurements found in group 1. In this case, the area under the curve is calculated with reference to the first value. In contrast to AUC_G, it ignores the distance from zero for all measurements, thereby emphasizing the changes over time. This area has therefore been named AUC_I (Area under the curve with respect to the increase).

The formula can be derived from the formula for AUC_G, since it is identical to AUC_G except for the removal of the area between ground and the first measure (baseline) for all time points. Hence, it can be expressed as in formula (4):

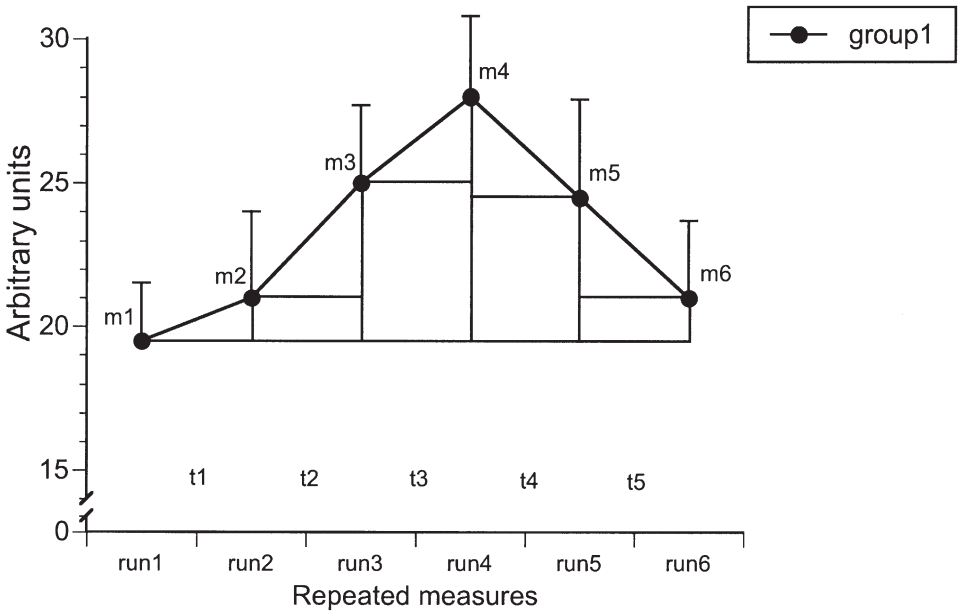

Fig. 2. Time course of group 1 over the six measurements; the triangles and rectangles illustrate the composition of the area under the curve with respect to the increase (AUC_I). m_1 to m_6 denote the single measurements, and t_1 to t_5 denote the time interval between the measurements. Note that although in this example, the time interval between the measurements is identical for all observations, individual time intervals can vary depending on the study.

$$AUC_I = AUC_G - m_1 \cdot \sum_{i=1}^{n-1} t_i \tag{4}$$

with m_1 denoting the first measurement and t_i denoting the time distances between measurements.

Replacing AUC_G with the variables in formula (2) leads to formula (5):

$$AUC_I = \left(\sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i) \cdot t_i}{2} \right) - \left(m_1 \cdot \sum_{i=1}^{n-1} t_i \right) \tag{5}$$

with m_i denoting the single measurements, t_i denoting the time distance between the measurements, and n denoting the total amount of measurements.

Formula (5) is also universally applicable to any number of repeated measurements. As noted before, with the time distance between measurements being constant, formula (5) can be further simplified by removing the time distance variable t_i and allows us to define formula (6):

$$AUC_I = \left(\sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i)}{2} \right) - (n-1) \cdot m_1 \tag{6}$$

with m_i denoting the single measurements and n denoting the total amount of measurements.

It has to be noted that the final AUC_I formula (6), although derived from area calculations, bears a difficulty. Given the case that the repeated measurements show a stronger decrease than increase over time, the result of this formula could become negative, since it is based on the reference to the first value. Therefore, in all cases with negative results of this formula, with a particular subject showing no increase but decrease, the area would have to be set to 0, avoiding negative areas and denoting the fact that no increase was seen in the particular subject. However, this results in a potential loss of information, since it might also be interesting to learn how strong the decrease in a subject was. Therefore, it is suggested to continue the statistical analysis even with negative values, but in these cases the result must be regarded as an 'index of decrease' rather than an area. This is no consideration for the AUC_G, since the reference to zero results in a true area in all cases.

The two *AUC* formulas were evaluated by using them in a one-factor ANOVA design (group) and comparing the results with those obtained from a two-factor mixed design ANOVA (group by time) with all measurements in one artificial and one original dataset. Examples for the calculation of the AUC variables with different measurements are given in Appendix A.

3. Study 1: Artificial data set

3.1. Artificial data creation

A computer program (HyperCard®, Apple Computer, Cupertino, CA, USA) was employed to generate random numbers varying around a fixed value. The program was designed so that these numbers follow a normal distribution around the fixed value. Datasets can be created this way representing possible measurements over time. In order to highlight the different information provided by the two formulas presented in this study, four different groups of data were created. Each group was assigned six time-points to represent measurements over time. Each 'time point' was filled with 20 random numbers varying around the fixed mean value. The mean values were chosen deliberately to illustrate the usefulness of the formulas of the *AUC*. Figure 3 shows the mean values and standard deviation for the four groups.

In this dataset, group 1 shows a significant increase over time for the first four measurements, and declines again for the last two measurements. While group 2 shows a similar course of increase, this increase takes place at a different level, approximately 10 units higher than group 1. The mean values of group 3 vary around the same level as those of group 2; but in contrast to group 2, it shows no increase over time; instead, the mean values do not vary much over the six measurements. Lastly, group 4 shows a u-shaped time course, with values decreasing during the first four measurements and recovering close to the original value towards the end of the six measurements.

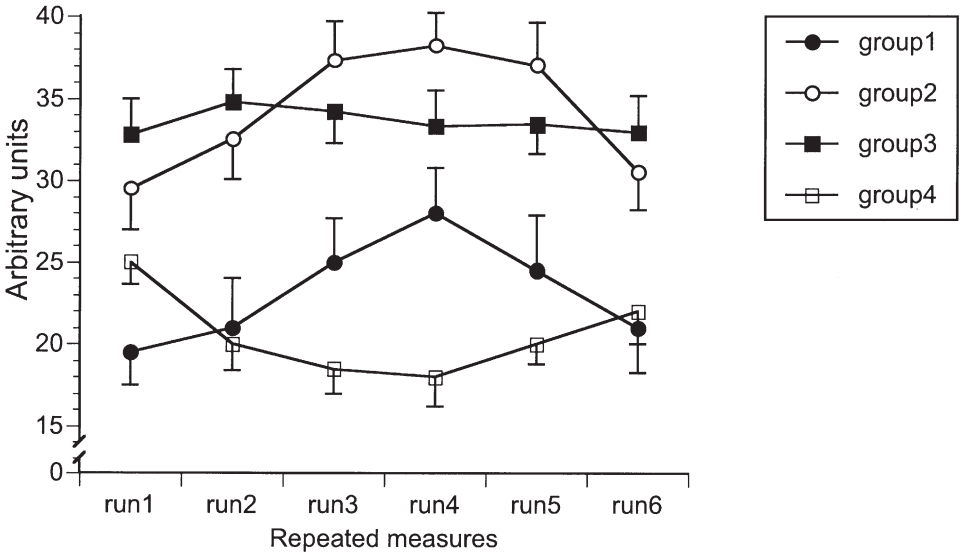

Fig. 3. The four groups over the six runs shown in a line graph. Error bars shown are standard deviations. All groups consist of n=20; total n=80.

3.2. Statistical analysis

The formulas for AUC_G and AUC_I were calculated by assuming identical one minute intervals between measurements, that way allowing the application of formulas (3) and (6). The original values were entered into a two factor (group by time) *within* ANOVA to calculate differences between the four groups. A Scheffe post hoc test was then employed to determine the direction of the differences between the groups. These results were then compared with a one factor (group) ANOVA that was calculated separately for the AUC_I and AUC_G values. Again, a Scheffe post hoc test was applied to determine the nature of the differences between the groups.

3.3. Results

Table 1 shows the correlation matrix of the two *AUC* measures with the single values of all four groups.

Several observations can be made when comparing the correlation coefficients of the two area measures with the single measures. First, the association between the single measures and AUC_G is higher than AUC_I. Second, the correlation between the first two measures and the AUC_I is particularly weak. This observation can be explained when recalling the composition of the two formulas; the first value constitutes the level at which the subsequent increase takes place, and this information is *taken out* from the AUC_I value. Consequently, the correlation with the first value can be expected to be weak in AUC_I. The generally weaker correlation between the single measurements and AUC_I can be explained in a similar fashion: Every single

Table 1
Pearson correlation coefficients (level of significance) of the two *AUC* measurements and the single measures; time 1 to time 6: individual time points of the dataset

	run 1	run 2	run 3	run 4	run 5	run 6
AUC_I	$r=-0.23$ $p=0.04$	$r=0.33$ $p=0.002$	$r=0.63$ $p<0.001$	$r=0.71$ $p<0.001$	$r=0.60$ $p<0.001$	$r=0.22$ $p=0.04$
AUC_G	$r=0.69$ $p<0.001$	$r=0.92$ $p<0.001$	$r=0.94$ $p<0.001$	$r=0.90$ $p<0.001$	$r=0.94$ $p<0.001$	$r=0.81$ $p<0.001$

Table 2
Results of the two-factor within ANOVA with the single measures

Effect	df (effect)	MS (effect)	df (Error)	MS (error)	*F*	*p*
Group	3	5698.6	76	7.94	692.58	<0.001
Run	5	129.5	380	7.52	12.97	<0.001
group×run	15	183.24	380	7.52	19.61	<0.001

measure can be regarded as consisting of two pieces of information: its distance from its neighbor and its distance from the ground (or zero). Since the distance from the ground is taken out from the composition of AUC_I, it can be expected that the correlation with the single time points are generally smaller than compared with AUC_G.

The results of the two-factor (group by time) within ANOVA with the single measurements are shown in Table 2.

While it is not surprising to see significant *F*-values for all main and interaction effects (since the datasets were specifically created to serve that purpose), the result of the Scheffe post hoc test contains additional information. It reveals that for the main effect of *group*, the difference between group 1 and 4 is not significant ($p=0.83$). Also, the differences between groups 2 and 3 turn out to be not significant ($p=0.85$). This is best illustrated when recalling the overall mean values for the four groups: although all groups show a different time course, groups 1 and 2 and groups 2 and 3 have almost identical overall mean values.

Next, two one-factor (group) ANOVAs with the two *AUC* values were computed. Table 3 shows the results for AUC_I and Table 4 shows the results for AUC_G.

Table 3
Results of the one factor (group) ANOVA with AUC_I as dependent variable

Effect	df (effect)	MS (effect)	df (Error)	MS (error)	*F*	*p*
Group	3	7359	76	205.15	36.05	<0.001

Table 4
Results of the one factor (group) ANOVA with AUC_G as dependent variable

Effect	df (effect)	MS (effect)	df (Error)	MS (error)	*F*	*p*
Group	3	24913	76	38.44	648.18	<0.001

The results from the Scheffe post hoc test revealed that for AUC_I, group 1 was significantly different from group 4 ($p<0.001$), while the difference to groups 2 and 3 did not reach statistical significance ($p=0.11$ in both cases). Group 2 showed a statistical significant difference to groups 3 and 4 ($p<0.001$), and group 3 was also different from group 4 ($p<0.001$).

For the Scheffe post hoc test that was applied to the ANOVA with AUC_G, the results indicated that group 1 was significantly different from groups 2 and 3 ($p<0.001$) but not from group 4 ($p=0.18$). Group 2 was significantly different from groups 1 and 4 ($p<0.001$), but not group 3 ($p=0.09$). Finally, group 3 turned out to be significantly different from group 4 ($p<0.001$). Due to the strong main and interaction effects in the artificial data set, the same results were observed using other available post hoc tests (data not shown). The results from the post hoc test are summarized in Table 5.

The results from group 1 illustrate that the two AUC variables reveal two different kinds of information. While group 1 is significantly different from groups 2 and 3 when employing the AUC_G measure, it is not different when using the AUC_I measure. The effects are opposite when comparing groups 1 and 4, where only the AUC_I measure reveals a statistical significant difference.

4. Study 2: Cortisol levels in a population of chronically stressed high-school teachers

4.1. Subjects and study design

Endocrinological data from a recent study with 69 teachers were chosen to validate the usefulness of the AUC formulas. In this study, 69 teachers from elementary and

Table 5
Results of the one Scheffe post hoc tests with AUC_G and AUC_I as dependent variables in the ANOVA design

	Group1		Group2		Group3	
	AUC_I	AUC_G	AUC_I	AUC_G	AUC_I	AUC_G
Group1	–	–				
Group2	n.s.	$p<0.001$	–	–		
Group3	n.s.	$p<0.001$	$p<0.001$	n.s.	–	–
Group4	$p<0.001$	n.s.	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$

high schools in the region of Trier, Germany, sampled saliva for cortisol analysis at three separate days at the time of awakening and 15, 30 and 60 minutes thereafter. At the night before day three, all teachers took 0.5 mg dexamethasone (PO) to test suppression of the hypothalamic–pituitary–adrenal (HPA) axis the next morning by this agent. Furthermore, all subjects underwent a psychological and medical screening including assessment of work stress (Schulz and Schlotz, 1999) and assessment of number and intensity of physical complaints (Fahrenberg, 1994). The results of this study have been reported elsewhere (Pruessner et al., 1999). Here, only the endocrinological data from day 3 (after dexamethasone suppression), the number of physical complaints, and the perceived workload are used for further analysis. The associations between these variables and the *AUC* measures have not been presented before.

4.2. Statistical analysis

AUC variables of increase versus total release were calculated from the four endocrinological measures on day three (0, 15, 30 and 60 minutes after awakening, following overnight dexamethasone suppression). Since the time interval between the measurements was not identical in this study, formula (2) was used for computation of AUC_G, whereas formula (5) came into effect for computation of AUC_I. These two variables were then used subsequently in two cluster analyses using the *k*-means method (Wishart, 1998) with the objective to define groups of subjects with similar endocrinological patterns. In the first analysis, only the AUC_G variable was used to find two groups of subjects who show opposite patterns in total release of cortisol after awakening. In a second analysis, only the AUC_I variable was used to find two groups of subjects who show opposite patterns of increase of cortisol after awakening. Since the cluster analysis method makes no restrictions in terms of group size, numbers of subjects in each group are usually not balanced. Each analysis was performed specifying two cluster solutions. Given the nature of the two formulas, it was expected to build groups that either were most distinct with respect to the increase after awakening (when employing AUC_I in the cluster analysis), or the total release after awakening (when employing AUC_G in the cluster analysis). Two new variables were created from the two cluster analyses, indicating whether an individual belonged to group 1 or 2 in the AUC_I cluster analysis (AUC_I group), or to group 1 or 2 in the AUC_G cluster analysis [AUC_G group].

Next, these new group variables were used as independent variables for a two way *mixed design* ANOVA (*AUC* group by time) with the raw cortisol levels as dependent variables. The ANOVA was performed twice, once with the AUC_I group variable, and once with the AUC_G group variable as independent variable. Also, the two group variables were used in two independent sample *t*-tests with the number of physical complaints and the perceived work stress as dependent variables, respectively. Since both the independent *t*-tests as well as the ANOVAs were performed twice with different independent variables, the alpha level of significance had to be adjusted to 2.5% (two-sided) to account for the number of tests. Finally, Pearson correlation coefficients were calculated with the *AUC* variables and the psychological

and physical variables to complement the information that could be derived from the two group variables and the endocrinological data.

4.3. Results

The two groups created from the cluster analysis employing the AUC_I variable differed most significantly in their increase, as expected. Group 1 consisted of twelve subjects, who showed a very strong cortisol increase after awakening, as compared to 57 subjects in group 2 who showed no such increase (Fig. 4). The difference between the two groups was demonstrated by the results from the two-way *mixed design* ANOVA with the AUC_I group variable as independent variable and the four endocrinological measures as dependent measures, showing a highly significant main effect of group (F[1,68]=55.1, $p<0.001$).

Using the AUC_I group variable in an independent sample t-test with perceived work stress and physical complaints indicated a significant difference with regard to the perceived work stress, with subjects showing a high increase after awakening also reporting a higher stress load (t=5.3, $p<0.01$, adjusted). This is in line with previous findings reporting an association between stress load and cortisol increase

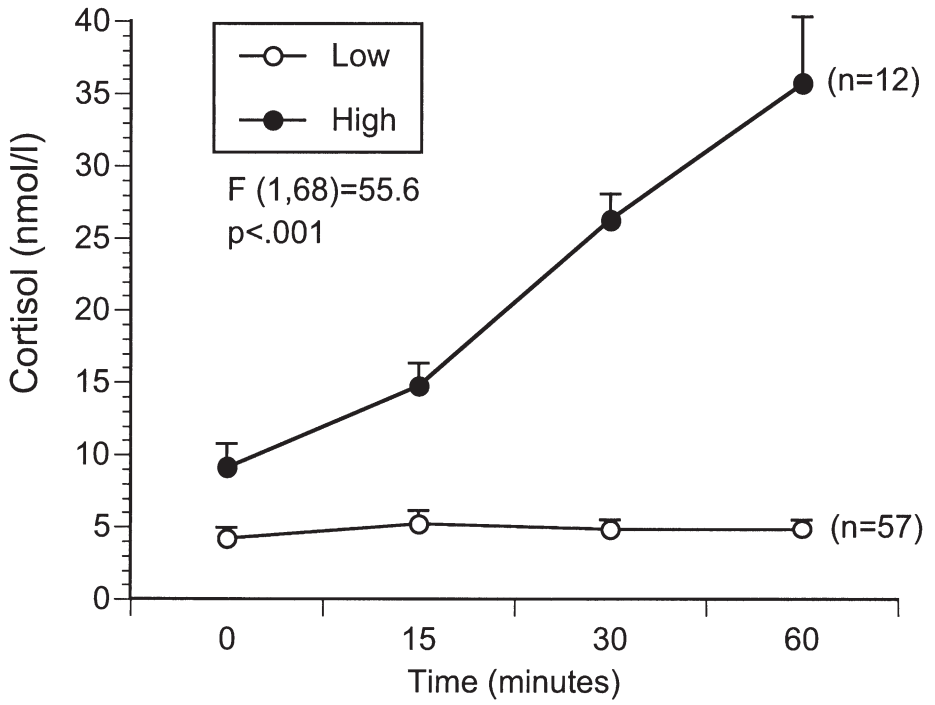

Fig. 4. Cortisol levels during the first hour after awakening in two groups of teachers with low and high increases as indicated by cluster analysis performed with AUC_I. Error bars shown are standard error of mean.

after awakening (Schulz et al., 1997). No effect could be seen for the number of physical complaints the subjects had reported (t=1.3; p>0.10). Also, the Pearson correlation between perceived stress and the AUC_I variable was significant (r=0.43, p<0.05), while no significant correlation with the number of physical complaints (r=0.17, p>0.20) could be observed.

The two groups of subjects created from the cluster analysis using the AUC_G variable resulted in a group consisting of 17 subjects, who showed a high total release of cortisol after awakening, as compared to 52 teachers who showed only a very small release of cortisol after awakening (Fig. 5). This was also reflected in the two-way *mixed design* ANOVA using the AUC_G group variable as independent variable and the single cortisol measurements as dependent variable, which showed a highly significant main effect of group (F[1,68]=492.5, p<0.001). In contrast to the cluster solution created from the AUC_I levels, there was no association with the stress load as indicated by t-test or Pearson correlation (t<1 and r=−0.13, both p>0.20). However, the t-test with the number of physical complaints as dependent variable now became significant (t=4.4, p<0.05, adjusted). This is in line with previous findings describing an association between the total release of cortisol in response to different stimuli and pain-sensitivity (Geiss et al., 1997). The Pearson correlation between AUC_G and the number of physical complaints was not significant (r=−0.15, p>0.20).

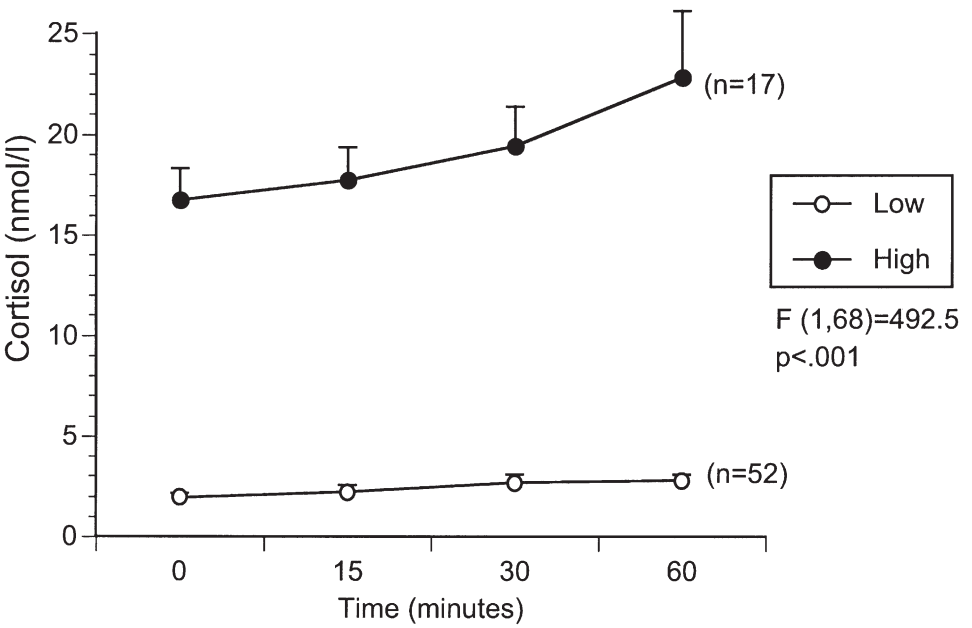

Fig. 5. Cortisol levels during the first hour after awakening in two groups of teachers with an overall low or high release as indicated by cluster analysis performed with AUC_G. Error bars shown are standard error of mean.

5. Discussion

This paper describes two formulas for computation of the *AUC* derived from the trapezoid formula. Transformation yielded formulas universally applicable to any number of repeated measurements in any experimental design. In addition, in cases where the time distance between measurements is identical throughout the experiment, a further simplification of the formulas for computation of *AUC* was presented.

Applying the formulas to a set of artificially created data (Study 1) with four groups and six repeated measurements in each group, it could be shown that the two formulas can reveal different information that is embodied in the repeated measurements. The names of the two formulas were thus chosen to represent this different information, i.e. 'Area under the curve with respect to increase' (AUC_I) and 'Area under the curve with respect to ground' (AUC_G). The use of *AUC* formulas simplifies the statistical analyses when the number of repeated measurements is high and there is a need for comprising the available information. It is also beneficial for the researcher to limit the amount of statistical comparisons between groups in order to minimize correction of the α-error probability. With the *AUC* variables, the number of repeated measurements is irrelevant and thus, the number of statistical comparisons only depends on the number of groups to be compared. With the two *AUC* formulas presented here, repeated measurements can be used to assess different aspects of the time course of the repeated measurements, without the need to go back to the original data. Although the use of a within design ANOVA with repeated measurements also allows the handling of a large number of repeated measurements quite easily, it is disadvantageous over the use of the formulas for two reasons: first, if the time interval between the measurements is not identical, the within design ANOVA has no method to correct for these differences; second, the analysis of the repeated measures is unable to differentiate between the two sorts of information comprised within each measure, thus it will most likely neglect some of the possibilities that have been demonstrated with the use of the formulas in this manuscript.

The usefulness of the two formulas was then further exemplified in Study 2, using an endocrinological dataset. Here, it could be shown that the resulting values from the two formulas might be differentially related to psychological variables, with one variable being more related to stress perception, and the other being more related to physical complaints. This practical example may thus demonstrate that it might be beneficial to build groups of subjects derived from the results of the *AUC* formulas to reveal associations with specific variables.

Whether the two formulas will reveal associations to different variables depends on the variable set, and the information of interest. With endocrinological data, it can be assumed that the use of the AUC_G will result in a measure that is more related to 'total hormonal output', whereas the use of AUC_I is more related to the sensitivity of the system, pronouncing changes over time.

However, since the two formulas are easy to apply, including them in statistical analyses with repeated measurements will only add to the information being extracted from the data. Furthermore, in studies with larger amounts of data due to either repeated sampling over several days (Pruessner et al., 1997), or due to the use of

different treatments in the same group of subjects with repeated measures (Yehuda et al., 1996; Altemus et al., 2001), the use of *AUC*s can significantly reduce the amount of data used for the statistical analysis. Finally, the emphasis on different characteristics (change over time versus overall intensity) of the two *AUC* variables provides the researcher with a first interpretation of the data, should differences appear between groups for one or the other *AUC* measure.

One additional aspect that needs to be briefly mentioned is the error distribution within the two *AUC* measures. Given that each of the single measures has the same (random) amount of error, AUC_I is characterized by accumulation of the error of the first measure m_1, since the formula is based on the difference between the first and the subsequent measures. In cases where the first measure contains a higher error, the resulting AUC formula will also contain a higher amount of error. In AUC_G, the first and the last measure contribute to a lesser extent to the area measure than the other values, the first value thus has a smaller impact. This can be of importance in studies where the first measure can be expected to contain an extra amount of error, due to the experimental design (e.g., catheter placement, nervousness of the investigator, etc.). Finally, another potential danger lies in the 'blind' application of the *AUC* measures in the statistical analysis, without prior visual inspection of the data. In cases where the repeated measurements show a change of pattern over time, these different patterns would be confounded if only one *AUC* measure was used for all measurements. In these cases, it might be advisable to form smaller sequences of repeated measurements to arrange the data differently. However, the decision on which sequences of data should be incorporated into an *AUC* measure should be made with respect to the content of the study, not on the data alone.

Appendix A

Examples for the calculation of ***AUC*** **variables with different measurements**

Case 1: Five repeated measures, identical time between measurements
Measure A: 3.5 (all measures with arbitrary units)
Measure B: 7
Measure C: 14
Measure D: 7
Measure E: 10

For the calculation of AUC_G, formula (3) applies:

$$AUC_G = \sum_{i=1}^{n-1} \frac{m_{(i+1)} + m_i}{2}$$

AUC_G=(B+A)/2+(C+B)/2+(D+C)/2+(E+D)/2

=(A+B+B+C+C+D+D+E)/2
=(A+E)/2 + B + C + D
=(10+3.5)/2 +7 + 14 + 7
=(6.75) +28
AUC_G=34.75

For the calculation of AUC_I, formula (6) applies:

$$AUC_I = \left(\sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i)}{2} \right) - (n-1) \cdot m_I$$

AUC_I=((B+A)/2+(C+B)/2+(D+C)/2+(E+D)/2)−4×A
=34.75−4×3.5
=34.75−14
AUC_I=20.75

Case 2: Six repeated measures, variable time between measurements
Measure A: 3.5 (all measures with arbitrary units)
t_1 (time to next measure): 10 minutes
Measure B: 7
t_2: 5 minutes
Measure C: 14
t_3: 15 minutes
Measure D: 7
t_4: 15 minutes
Measure E: 10

For the calculation of AUC_G, formula (2) applies:

$$AUC_G = \sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i) \cdot t_i}{2}$$

AUC_G=(B+A)/2×t_1+(C+B)/2×t_2+(D+C)/2×t_3+(E+D)/2×t_4
AUC_G=((3.5+7)/2×10+(7+14)/2×5+(14+7)/2×15+(7+10)/2×15)
AUC_G=(10.5/2×10)+(21/2×5)+(21/2×15)+(17/2×15)
AUC_G=52.5+52.5+157.5+127.5=390

For the calculation of AUC_I, formula (5) applies:

$$AUC_I = \left(\sum_{i=1}^{n-1} \frac{(m_{(i+1)} + m_i) \cdot t_i}{2} \right) - \left(m_I \cdot \sum_{i=1}^{n-1} t_i \right)$$

$$AUC_I=((B+A)/2\times t_1+(C+B)/2\times t_2+(D+C)/2\times t_3+(E+D)/2\times t_4)-(A\times(t_1+t_2+t_3+t_4\ ))$$
$$AUC_I=(390)-(3.5\times(10+5+15+15)$$
$$AUC_I=(390)-(3.5\times(45))=390-157.5=232.50$$

References

Altemus, M., Dale, J.K., Michelson, D., Demitrack, M.A., Gold, P.W., Straus, S.E., 2001. Abnormalities in response to vasopressin infusion in chronic fatigue syndrome. Psychoneuroendocrinology 26, 175–188.

Fahrenberg, J., 1994. Freiburger Beschwerdeliste FBL-G/R. Hogrefe, Göttingen.

Geiss, A., Varadi, E., Steinbach, K., Bauer, H.W., Anton, F., 1997. Psychoneuroimmunological correlates of persisting sciatic pain in patients who underwent discectomy. Neurosci. Lett. 237, 65–68.

Ghizzoni, L., Bernasconi, S., Virdis, R., Vottero, A., Ziveri, M., Volta, C., Lughetti, L., Giovannelli, G., 1994. Dynamics of 24-hour pulsatile cortisol, 17-hydroxyprogesterone, and androstenedione release in prepubertal patients with nonclassic 21-hydroxylase deficiency and normal prepubertal children. Metabolism 43, 372–377.

Gormley, G.J., Stoner, E., Bruskewitz, R.C., Imperato-McGinley, J., Walsh, P.C., McConnell, J.D., Andriole, J.L., Geller, J., Bracken, B.R., Tenover, J.S., et al. 1992. The effect of finasterine in men with benign prostatic hyperplasia. N. Engl. J. Med 327, 1185–1191.

Maes, M., Calabrese, J., Meltzer, H.Y., 1994. The relevance of the in- versus outpatient status for studies on HPA-axis in depression: spontaneous hypercortisolism is a feature of major depressed inpatients and not of major depression per se. Prog. Neuropsychopharmacol. Biol. Psychiatry 18, 503–517.

O'Brien, J.T., Ames, D., Schweitzer, I., Mastwyk, M., Colman, P., 1996. Enhanced adrenal sensitivity to adrenocorticotrophic hormone (ACTH) is evidence of HPA axis hyperactivity in Alzheimer's disease. Psychol. Med 26, 7–14.

Pruessner, J.C., Hellhammer, D.H., Kirschbaum, C., 1999. Burnout, perceived stress and cortisol responses to awakening in teachers. Psychosom. Med 61, 197–204.

Pruessner, J.C., Gaab, J., Hellhammer, D.H., Lintz, D., Schommer, N., Kirschbaum, C., 1997. Increasing correlations between personality traits and cortisol stress responses obtained by data aggregation. Psychoneuroendocrinology 22, 615–625.

Reinhardt, F., Soeder, H., 2001. dtv Atlas Mathematik. Deutscher Taschenbuch Verlag, München.

Schulz, P., Schlotz, W., 1999. Trier Inventar zur Erfassung von chronischem Streß (TICS): Skalenkonstruktion und teststatistische Überprüfung.

Schulz, P., Kirschbaum, C., Pruessner, J.C., Hellhammer, D.H., 1997. Increased free cortisol secretion after awakening in chronically stressed individuals due to work overload. Stress Med 14, 91–97.

Tucci, J.R., Tonino, R.P., Emkey, R.D., Peverly, C.A., Kher, U., Stantora, A.C., 1996. Effect of three years of oral alendronate treatment in postmenopausal women with osteoperosis. Am. J. Med. 101, 488–501.

Wishart, D., 1998. Clustan Graphics3: Interactive Graphics for Cluster Analysis. In: Gaul, W., Locarek-Junge, H. (Eds.), Classification in the Information Age. Proceedings of the 22nd Annual Conference of the Society for Classification. Springer, Berlin, pp. 268–275.

Yehuda, R., Levengood, R.A., Schmeidler, J., Wilson, S., Guo, L.S., Gerber, D., 1996. Increased pituitary activation following metyrapone administration in post-traumatic stress disorder. Psychoneuroendocrinology 21, 1–16.