# Pooling ambulatory saliva cortisol samples over consecutive days – as reliable as arithmetic means


Peter Garvin, John Carstensen and Margareta Kristenson





Original publication:

Peter Garvin, John Carstensen and Margareta Kristenson, Pooling ambulatory saliva cortisol samples over consecutive days – as reliable as arithmetic means, 2008, Scandinavian Journal of Clinical and Laboratory Investigation, (68), 6, 508-512.
http://dx.doi.org/10.1080/00365510701832229.

# Pooling ambulatory saliva cortisol samples over consecutive days – as reliable as arithmetic means.

Running head: pooling cortisol reliable

Peter Garvin[1,*], John Carstensen[2], Margareta Kristenson[1].

[1] Division of Social Medicine and Public Health Science. Department of Medicine and Health Sciences. Linköping University, Sweden.

[2] Division of Health and Society. Department of Medicine and Health Sciences, Linköping University, Sweden.

[*] Corresponding author: Peter Garvin

Department of Medicine and Health Sciences.

Division of Social Medicine and Public Health Science.

Linköping University.

581 85 Linköping.

Sweden

Email address: peter.garvin@ihs.liu.se

Telephone: +46 13 22 88 75

Fax: +46 13 22 18 65

Word count abstract: 198

Full word count: 3 118, references included

## Abstract

*Objective:* When cortisol measurements are to be studied in large populations, cost-effective analyses are needed. This study aimed at testing if one pooled cortisol value over three consecutive days is as reliable as using the arithmetic mean of the samples from the same measure points.

*Methods*: 30 participants aged 45 to 69 collected saliva in salivettes immediately after awakening ($t_1$), thirty minutes after awakening ($t_2$) and in the evening ($t_3$) during three consecutive days. A fixed volume from each of the samples ($t_1$, $t_2$ and $t_3$) was pooled prior to laboratory analysis. Mean levels over three days for $t_1$, $t_2$ and $t_3$ were compared to corresponding levels of pooled vials. Cortisol levels were analysed using a radio immunoassay.

*Results*: All measures tested had high correlations between mean values and pooled samples, exemplified with diurnal deviation $r_{dif\ t2\text{-}t3}$=0.974 (CI 0.946;0.987), and awakening response $r_{dif\ t2\text{-}t1}$ =0.982 (CI 0.963;0.991). There were no statistical differences between the pooled values and the arithmetic means.

*Conclusions*: Pooling samples gave as reliable results as arithmetic means did. Pooling samples prior to laboratory analysis provides a cost-effective method for measuring general diurnal cortisol variation in field research projects.

## Keywords

Cortisol, HPA-axis, measurements, methodology, saliva, stress.

## Introduction

Given the facetted role of cortisol, it has been suggested that dysregulation of this hormone is associated with a number of pathophysiological conditions [1-3]. In particular, cortisol has been suggested to play a vital role linking psychosocial stress to somatic response [4-7]. As cortisol in saliva has been shown to reflect the levels in serum with good precision [8,9], the assessment of cortisol could be easily performed in a non-invasive procedure using saliva sampling. It is well accepted in both standardised laboratory approaches and ambulatory saliva sampling that dynamic responses give more relevant information than static measures [7]. Hence, at least two different time points of measuring are needed in order to estimate cortisol dynamics.

Comparing consecutive days, the correlation between cortisol levels have been reported to be around r=0.5, and the intra-individual coefficient of variance (CV) to be around 20 % [10-13]. Samples from more than one day are commonly used in order to reduce the potential state variables that may occur in everyday life, and thus achieve a more reliable and valid assessment. Calculations are then typically performed on mean values of two or three consecutive days, for each of the different time point of measuring [14-16]. However, the combination of a need for dynamic measures and the use of consecutive days in study designs lead to a large set of samples. The high total number of samples required is a major challenge in all study designs based on a large number of participants, as the analyses become very expensive. Thus, a more cost-effective analysis would be highly appreciated.

### Specific aim

This study aimed at testing if one pooled cortisol value over three consecutive days is as reliable as using the arithmetic mean of the samples from the same measure points.

## Methods

### Design.

Participants were recruited from the LSH-study (Life conditions, Stress and Health), randomly drawn from a normal population. See Hollman and Kristenson for details [17]. Data were collected from late 2003 to early 2004, and participants, aged 45-69 years, were evenly distributed by sex and age. 30 individuals were randomly chosen within the first 1,000 participants in the LSH-study to this sub-study, without any stratification criteria.

### Samples and laboratory settings.

The participants collected saliva samples during three consecutive days in salivettes (Sarstedt AB, Sweden) immediately after awakening ($t_1$), thirty minutes after awakening ($t_2$) before breakfast and just before going to bed in the evening ($t_3$). The time points were chosen to capture the peak in the morning and the restitution in the evening [11]. The salivettes were kept in a refrigerator during and immediately after the sample days, and were thereafter brought to the closest PHC centre, where they were centrifuged. 150 μl from each salivette or as much as the salivette with least content contained if amounts less than 150 μl, were pooled; one vial for awakening samples (pool- $t_1$), one for samples taken thirty minutes after awakening (pool- $t_2$) and one for evening samples (pool- $t_3$). If one salivette was empty, saliva from the other two were pooled together.

Remaining saliva samples were transferred to separate vials, giving in total 12 vials per participant. After centrifugation and pooling, vials were stored in $-70^{o}$ C in approximately six months before analysis. Earlier studies have demonstrated that this storage time would not have an impact on detectable cortisol levels [18]. All samples were thawed and analysed in duplicates in the same run in a radioimmunoassay method (Diagnostic Products Corporation,

Los Angeles. CA). Intra-assay coefficient of variation was 9.4 % at 5 nmol/l and 9.8 % at 12 nmol/l.

**Statistical analysis**

Arithmetic mean levels for $t_1$, $t_2$ and $t_3$ were compared to levels of pooled vials (pool- $t_1$, pool- $t_2$ and pool- $t_3$). Correlations and 95 % confidence intervals using the Fischer z to r transformation were calculated for cortisol levels at different time points ($t_1$, $t_2$ and $t_3$) at each of the three days versus each other, for pooled values versus cortisol levels at $t_1$, $t_2$ and $t_3$ each day, and for pooled values versus arithmetic mean values of $t_1$, $t_2$ and $t_3$. Partial correlations were adjusted for the following: Sex, age, deviation from self reported normal awakening time (n=5, 18 % reported one awakening time that differed more than one hour in comparison to the other days of assessment), non-adherence to the protocol regarding taking sample 30 minutes after awakening (n=8, 9 % reported other than taking the sample 30 minutes after $t_1$: four samples after 20 mins, two after 33 mins, one after 40 mins and one after 82 mins, respectively) and outliers (n=2; 2 % with $t_3$ being higher than $t_1$ and $t_2$). Participants were excluded in specific sub-analyses of partial correlations when information from questionnaires was not sufficient. Those were limited to a few cases (n=0 to n=4).
Both absolute levels and measures of diurnal variation were studied. Pairwise t-tests were used to investigate if there were any significant differences when pooled values were compared to arithmetic means.

## Results

Of the 30 randomly chosen individuals within the LSH-study (mean age 57.6 yrs, SD 7.1), there were 12 men (mean age 58.4 yrs, SD 6.7) and 18 women (mean age 57.1 yrs, SD 7.5). Eight participants (n=8, 27%) reported an education attainment of college level, thirteen (n=13, 43%) reported attained ten to twelve years, and nine participants (n=9, 30%) reported

an education attainment of nine completed years or less. The overall mean of all unpooled samples were 10.6 nmol/l (n=263, SD 9.5), with an absolute range from 0.31 to 59.0 nmol/l . Arithmetic means over three days for $t_1$, $t_2$ and $t_3$ were 11.7 (n=86 SD 6.5), 17.5 (n=89 SD 10.1) and 2.3 nmol/l (n=88 SD 2.4), respectively, with an intra-individual CV of 29 %. Corresponding mean levels of the pooled samples ($t_1$-pool, $t_2$–pool and $t_3$ –pool) were 11.5 (n=30 SD 5.4), 17.2 (n=30 SD 8.1) and 2.2 (n=30 SD 1.3), respectively.

Partial correlations between single measurements at specific time points at different days of sampling and corresponding results for single days versus pooled values are shown in Table 1. The compliance for collecting all samples requested was n=26 (87%) for $t_1$, n=29 (96%) for $t_2$ (96 %) and n=28 (93%)for $t_3$, respectively. The correlations were adjusted for sex, age, deviation from self reported estimated normal awakening time, non-adherence to the protocol regarding taking sample 30 minutes after awakening and outliers in night sample, $t_3$ (two outliers in the first night samples; concentration 14.0 nmol/l and 16.6 nmol/l, respectively). As the effect of adjustment for sex and age was marginal, they are not presented.

The correlation coefficients using full model when correlating samples of one of the consecutive days to another were in the range from r= 0.46 (p=0.033) to r= 0.77 (p<0.001). The correlation coefficients using full model when correlating pooled values to any of the consecutive days were generally higher, ranging from r= 0.74 (p<0.001) to r= 0.87 (p<0.001). A descriptive plot over pooled samples and arithmetic mean values is shown in Figure 1. Correlations and pairwise t-tests of mean values and pooled samples are presented in Table 2a for absolute levels, and in Table 2b for measures of diurnal variation, i.e. differences and ratios of $t_1$, $t_2$ and $t_3$, and area under curve [19]. The correlation coefficients for measures of diurnal variation were in the range from $r_{ratio\ (t1/t3)}$=0.803 (CI 0.623;0.902) to $r_{difference\ (t2-t1)}$=0.982 (CI 0.963;0.991). The paired t-tests performed could not reveal any significant difference when comparing arithmetic means to pooled values. The lowest p-value derived

*Table 1. Partial correlation coefficients, r, between cortisol levels in saliva at each of the days versus each other, and for each day versus pooled values. 95 % CI in parenthesis, n= 26 to n=30. Stepwise adjustment for sex, age, deviation from self reported normal awakening time, non-adherence to taking sample 30 minutes after awakening, and outliers defined by a cortisol level at night, t3, that is higher than levels of both t2 and t1. Two participants show this pattern day 1.*

| | Model adjusting for: | No other variables (unadjusted corr.) | Age + sex + deviation from estimated normal awakening time | + non-adherence to protocol, 30 mins after awakening time | + outlier | Full model |
|---|---|---|---|---|---|---|
| **Awakening** | day 1 vs day 2 | 0,45 (0.08;0.71) | 0,46 (0.08;0.73) | - | - | 0,46 (0.08;0.73) |
| | day 1 vs day 3 | 0,52 (0.18;0.76) | 0,52 (0.17;0.76) | - | - | 0,52 (0.17;0.76) |
| | day 2 vs day 3 | 0,69 (0.42;0.85) | 0,66 (0.38;0.84) | - | - | 0,66 (0.38;0.84) |
| | pool vs day 1 | 0,73 (0.50;0.88) | 0,79 (0.59;0.91) | - | - | 0,79 (0.59;0.91) |
| | pool vs day 2 | 0,84 (0.69;0.93) | 0,74 (0.51;0.88) | - | - | 0,74 (0.51;0.88) |
| | pool vs day 3 | 0,87 (0.75;0.95) | 0,82 (0.65;0.92) | - | - | 0,82 (0.65;0.92) |
| **30 mins** | day 1 vs day 2 | 0,69 (0.43;0.85) | 0,69 (0.38;0.84) | 0,61 (0.29;0.81) | - | 0,61 (0.29;0.81) |
| | day 1 vs day 3 | 0,50 (0.17;0.74) | 0,49 (0.14;0.74) | 0,46 (0.09;0.73) | - | 0,46 (0.09;0.73) |
| | day 2 vs day 3 | 0,62 (0.34;0.81) | 0,62 (0.33;0.82) | 0,62 (0.31;0.81) | - | 0,62 (0.31;0.81) |
| | pool vs day 1 | 0,77 (0.57;0.89) | 0,79 (0.59;0.90) | 0,76 (0.55;0.89) | - | 0,76 (0.55;0.89) |
| | pool vs day 2 | 0,86 (0.73;0.94) | 0,79 (0.59;0.90) | 0,81 (0.63;0.92) | - | 0,81 (0.63;0.92) |
| | pool vs day 3 | 0,87 (0.74;0.94) | 0,83 (0.66;0.92) | 0,87 (0.74;0.95) | - | 0,87 (0.74;0.95) |
| **Evening** | day 1 vs day 2 | -0,01 (-0.38;0.36) | 0,04 (-0.34;0.42) | - | 0,60 (0.30;0.80) | 0,60 (0.30;0.80) |
| | day 1 vs day 3 | 0,00 (-0.38;0.38) | 0,04 (-0.34;0.42) | - | 0,64 (0.34;0.83) | 0,64 (0.34;0.83) |
| | day 2 vs day 3 | 0,60 (0.31;0.80) | 0,77 (0.55;0.89) | - | - | 0,77 (0.55;0.89) |
| | pool vs day 1 | 0,30 (-0.08;0.61) | 0,13 (-0.25;0.49) | - | 0,83 (0.67;0.93) | 0,83 (0.67;0.93) |
| | pool vs day 2 | 0,66 (0.40;0.83) | 0,77 (0.57;0.90) | - | 0,76 (0.55;0.89) | 0,76 (0.556;0.887) |
| | pool vs day 3 | 0,73 (0.50;0.87) | 0,87 (0.75;0.95) | - | 0,87 (0.73;0.94) | 0,87 (0.73;0.94) |

Figure 1. Plots over arithmetic means and pooled values of cortisol in nmol/l. A to C show each time point separated. D is showing descriptive plot over all time points in one graph. Note that there is different scaling in different graphs due to high variation of levels.

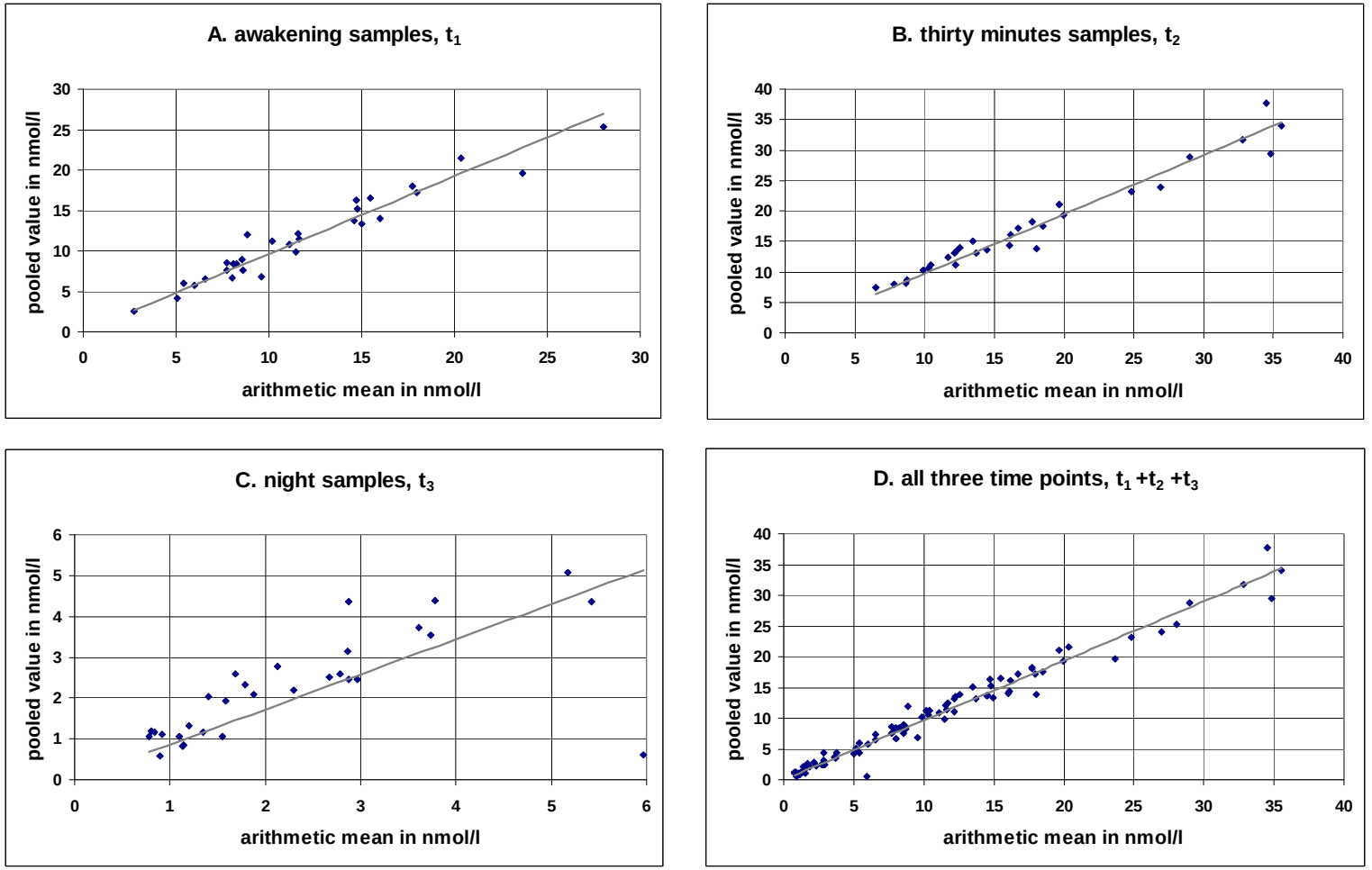

*Table 2. Correlation coefficients, r, and pairwise t-tests for the arithmetic mean versus the pooled value for each participant. 95 % CI in parenthesis, n=30. The three time points for measuring t1,t2 and t3 are showed in A (nmol/l), different measures on diurnal variation derived from these three are showed in B.*

**A. Time points**

| | | Paired t-test | |
|---|---|---|---|
| | correlation coefficient, *r* (95 % CI) | mean arith mean pool | difference arith-pool (95 % CI) |
| **Awakening samples $t_1$** | 0.97 | 11.84 | 0.31 |
| | (0.93;0.99) | 11.53 | (-0.23;0.86) |
| **30 mins. samples $t_2$** | 0.98 | 17.54 | 0.33 |
| | (0.95;0.99) | 17.21 | (-0.32;0.98) |
| **Evening samples $t_3$** | 0.66 | 2.30 | 0.09 |
| | (0.39;0.83) | 2.21 | (-0.33;0.51) |
| **All samples $t_1+t_2+t_3$** | 0.98 | 10.56 | 0.24 |
| | (0.97;0.99) | 10.32 | (-0.06;0.55) |

**B. Measures of diurnal variation**

| | | Paired t-test | |
|---|---|---|---|
| | correlation coefficient, *r* (95 % CI) | mean arith mean pool | difference arith-pool (95 % CI) |
| **Difference ($t_1$- $t_3$)** | 0.95 | 9.54 | 0.22 |
| | (0.90;0.98) | 9.32 | (-0.43;0.88) |
| **Difference ($t_2$- $t_3$)** | 0.97 | 15.23 | 0.23 |
| | (0.94;0.99) | 15.00 | (-0.50;0.97) |
| **Difference ($t_2$- $t_1$ )** | 0.98 | 5.69 | 0.01 |
| | (0.96;1.00) | 5.68 | (-0.58;0.61) |
| **Ratio ($t_1$/ $t_3$)** | 0.80 | 6.70 | -0.30 |
| | (0.62;0.91) | 7.00 | (-1.56;0.96) |
| **Ratio ($t_2$/ $t_3$)** | 0.81 | 10.38 | -0.43 |
| | (0.64;0.91) | 10.81 | (-2.36;1.50) |
| **Ratio ($t_2$/ $t_1$ )** | 0.97 | 1.70 | -0.06 |
| | (0.94;0.99) | 1.76 | (-0.15;0.05) |
| **Area under curve with respect to ground** | 0.97 | 33.39 | 0.70 |
| | (0.95;0.99) | 32.69 | (-0.49;1.87) |
| **Area under curve with respect to increase** | 0.98 | 9.70 | 0.07 |
| | (0.95;0.99) | 9.63 | (-0.85;0.99) |

from table 2 was found when comparing calculated mean for all samples to pooled values (p=0.113, n=90). All other of the significance tests had a p-value of 0.25 or higher.

## Discussion

The main results in Figure 1 and Table 2 demonstrate that pooling samples gave the same results as calculated mean levels. The correlation coefficients were generally high when comparing the two approaches, and there were no significant differences in the data set in the pairwise t-tests.

Table 1 illustrates that pooled values were a better proxy for the diurnal variation of any of the single days than using single day measurements as a proxy for the diurnal variation of any of the other days. Pooling two days did not provide as good proxy for any of the three days as pooling three days did (data not shown), but was still higher than using single day measurements.

Notably, this study was conducted in 10 PHC centres, and in several cases more than one technician per PHC were involved; i.e. around 15 individuals were involved in the pooling procedures. Thus, the feasibility is not restricted to studies where the analyses are conducted by one or e few laboratory workers, but also works in field studies involving a lot of personnel. Pooled values had a higher correlation to any of the days that constituted the pooled sample than the correlations between any two of the days studied.

A major limitation when pooling samples is that the data cannot be used if the research question is to explore causes of day-by-day variation, as e.g. Hellhammer et al [12] and Adam et al [20] have suggested. Pooling samples prior to laboratory analysis is only applicable when it comes to estimate a general HPA-capacity of response and restitution.

There was a general high compliance in this study. Apart from factors adjusted for in Table 1, only three participants (10%) did comment on their non-compliance to the protocol. Two of these reported that they had been eating shortly prior to one of the saliva sampling points. One reported intake of sedative medication prior to one of the points. Possibly due to the high compliance, there was only one adjustment that had a major impact on the partial correlations shown in Table 1, namely adjusting for the two outliers in the evening values.

As the samples were not re-analysed, it was not possible to conclude that the high levels were due to laboratory analysis, or if the salivettes were taken in the wrong order. In refined designs, it is recommended that every sample taken before pooling should be kept in a freezer until the pooled vials are analysed. This provides the opportunity to analyse the samples constituting a pooled sample if finding a abnormal level, thus being able to rule out if the value can be explained by a single sample, or if the abnormality are consistent throughout all samples.

## Conclusions

Pooling procedure gave as reliable results as arithmetic means over all measurements tested. Pooling samples prior to laboratory analysis provides a more cost-effective method when assessing diurnal cortisol variation in field research projects. In designs where arithmetic means are to be used to measure general diurnal variation of cortisol, it is suggested that pooling is performed prior to laboratory analysis to reduce laboratory costs without losing information on mean levels.

## References


[1] Sterling P, Eyer J. Allostasis: A new paradigm to explain arousal pathology. In: Fisher, S. and Reason, J. (Ed) Handbook of Life Stress, Cognition and Health. John Wiley & Sons Ltd. New York, pp.629-649. 1988.

[2] Sapolsky RM, Romero LM, Munck AU. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. Endocr Rev 2000;21:55-89.

[3] Buckingham JC. Glucocorticoids: exemplars of multi-tasking. Br J Pharmacol 2006;147 Suppl 1:S258-68.

[4] Kirschbaum C, Wust S, Hellhammer D. Consistent sex differences in cortisol responses to psychological stress. Psychosom Med 1992;54:648-57.

[5] Levine S. Influence of psychological variables on the activity of the hypothalamic-pituitary-adrenal axis. Eur J Pharmacol 2000;405:149-60.

[6] Frankenhaeuser M. A biopsychosocial approach to work life issues. Int J Health Serv 1989;19:747-58.

[7] McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci 1998;840:33-44.

[8] Kirschbaum C, Hellhammer DH. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. Psychoneuroendocrinology 1994;19:313-33.

[9] Aardal E, Holm AC. Cortisol in saliva--reference ranges and relation to cortisol in serum. Eur J Clin Chem Clin Biochem 1995;33:927-32.

[10] Ice GH, Katz-Stein A, Himes J, Kane RL. Diurnal cycles of salivary cortisol in older adults. Psychoneuroendocrinology 2004;29:355-70.

[11] Sjogren E, Leanderson P, Kristenson M. Diurnal saliva cortisol levels and relations to psychosocial factors in a population sample of middle-aged Swedish men and women. Int J Behav Med 2006;13:193-200.

[12] Hellhammer J, Fries E, Schweisthal OW, Schlotz W, Stone AA, Hagemann D. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: state- and trait components. Psychoneuroendocrinology 2007;32:80-6.

[13] Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J, Perich C, Simon M. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59:491-500.

[14] Harris A, Ursin H, Murison R, Eriksen HR. Coffee, stress and cortisol in nursing staff. Psychoneuroendocrinology 2007;32:322-30.

[15] Thorn L, Hucklebridge F, Evans P, Clow A. Suspected non-adherence and weekend versus week day differences in the awakening cortisol response. Psychoneuroendocrinology 2006;31:1009-18.

[16] Edwards S, Evans P, Hucklebridge F, Clow A. Association between time of awakening and diurnal cortisol secretory activity. Psychoneuroendocrinology 2001;26:613-22.

[17] Hollman G, Kristenson M. The prevalence of the metabolic syndrome and its risk factors in a middle-aged Swedish population - Mainly a function of overweight? Eur J Cardiovasc Nurs 2007.

[18] Garde AH, Hansen AM. Long-term stability of salivary cortisol. Scand J Clin Lab Invest 2005;65:433-6.

[19] Pruessner JC, Kirschbaum C, Meinlschmid G, Hellhammer DH. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 2003;28:916-31.

[20] Adam EK, Hawkley LC, Kudielka BM, Cacioppo JT. Day-to-day dynamics of experience--cortisol associations in a population-based sample of older adults. Proc Natl Acad Sci U S A 2006;103:17058-63.

## Acknowledgement

We would like to thank all the nurses and biomedical staff on the participating PHC centres. We also like to thank Iva Soderström and Agneta Berg at Division of Clinical Chemistry at University hospital of Linköping, who performed the laboratory analyses. Further, we would like to thank Professor Holger Ursin for valuable comments on the work and on the manuscript. This work was supported with grants from the Swedish Research Council (2004-1881) and the Swedish Heart and Lung Foundation (20040530). Grants from Linköping University and the County of Östergötland were provided via the two joint strategic programmes CIRC, Cardiovascular Inflammation Research Centre and Department of Health and Society.