Test-Retest Reliability of Heart Rate Variability and Respiration Rate at Rest and during Light Physical Activity in Normal Subjects

Alida M. Guijt, Judith K. Sluiter, and Monique H.W. Frings-Dresen
Coronel Institute of Occupational Health, Academic Medical Center, Universiteit van Amsterdam, Amsterdam, The Netherlands

Received for publication April 6, 2006; accepted July 21, 2006 (ARCMED-D-06-00126).

**Background.** A variable that remains stable over repeated measurements (in stable conditions) is ideal for tracking modifications of the clinical state. The aim of the present study is to examine test-retest reliability of time-domain heart rate variability and respiration rate measurements using a portable device on normal subjects during rest and light physical activity.

**Methods.** Twenty-six normal subjects [18 females and 8 males aged 28 ± 6 years and 34 ± 12 years (mean ± SD), respectively] underwent two measurements for time-domain heart rate variability (SDNN and RMSSD) and respiration rate, with 7 days in between. Measurements took place under three conditions: lying down in a laboratory, cycling in a laboratory and sleeping in an ambulatory surrounding. Reliability was assessed statistically by calculating intra-class correlation coefficients (ICC).

**Results.** Reliability was found to be good to excellent for both time-domain heart rate variability (SDNN: ICC values between 0.74 and 0.85, RMSSD: ICC values between 0.75 and 0.98) and for respiration rate (ICC values between 0.77 and 0.96).

**Conclusions.** Both time-domain heart rate variability and respiration rate can be reliably assessed. However, we advise reliability research in a clinical setting before using the device for tracking modifications in a clinical state. © 2007 IMSS. Published by Elsevier Inc.

**Key Words:** Heart rate variability, Respiration rate, Autonomic nervous system, Time-domain, Supine position, Reproducibility.

---

**Introduction**

Experimental evidence for an association between cardiovascular diseases and signs of either increased sympathetic and/or reduced parasympathetic activity has encouraged the development of quantitative markers of autonomic activity such as heart rate variability (1). Heart rate variability is used as a non-invasive method that reflects the functioning of neural modulation and cardiac activity (2). In addition to being a way to diagnose cardiovascular diseases (1), heart rate variability can be used to observe changes in neural modulation and cardiac activity after a treatment plan has started. It is important to know the variation within subjects in repeated measurements when changes in heart rate variability due to interventions are observed (2,3). A variable that remains stable over repeated measurements (assuming a similar condition over time) is ideal for tracking modifications of the clinical state because changes of the variable have a higher probability of truly representing either alterations in such states or an effect of the experimental condition (4).

Cardiovascular processes operate in interaction with respiration in order to meet the highly variable metabolic demands of the organism and in order to maintain homeostasis (5). Respiration is assumed to modulate the activity of the cardiac vagal nerve (6). It is even suggested that there exists a central integration of certain respiratory and cardiovascular processes, such that some central neurons serve both functions (7).

Most of the published studies analyzed the reproducibility of heart rate variability obtained from 24-h measurements...
and/or in cardiovascular patients (8, 9). In general, it has been accepted that 24-h measurements of heart rate variability indices have adequate reproducibility (8). Unfortunately, little is known about the reproducibility of heart rate variability recordings that are of short duration (10). The studies available reported heterogeneous results, in part because of varying statistical analyses and the different populations studied (2). Available studies of short duration heart rate variability recordings have mostly used non-portable recorders, which make 24-h recordings impossible (11–13). Furthermore, Sandercock et al. recently reviewed current literature and concluded that future research to accurately assess the reliability of heart rate variability is necessary (14). The European and North American Task Force (1) reported specific guidelines for the assessment of heart rate variability. Recall that heart rate variability is also modulated by respiration rate, and although a great deal of therapy in patients with stress-related disorders includes breath regulation, there are no recent studies available that investigate reproducibility of respiration rate. Compared to the assessment of heart rate variability, assessment of respiration rate is relatively easy. This is why it would be worthwhile to investigate if respiration rate can be assessed reliably.

In the present study, reproducibility of heart rate variability and respiration rate during three maneuvers in normal subjects are determined. Heart rate variability and respiration rate are determined using the Co²ntrol (Decon Medical Systems, Weesp, The Netherlands), a chest strap that provides a signal to a wireless receiver. The aim of the present study is to examine test-retest reliability of heart rate variability and respiration rate measurements using the Co²ntrol in normal subjects during rest and light physical activity. Short duration heart rate variability assessment studies are scarce and their results heterogeneous; moreover, reproducibility of respiration rate has not been studied for many years and we had no basis for expecting particular results beforehand.

Materials and Methods

Subjects

All participants in the study had to be healthy and aged from 18 to 55 years. We wanted to assess reliability of heart rate variability and respiration rate in a heterogeneous sample of subjects without health complaints. To be certain that performing exercise was safe, participants’ health status was assessed on the basis of the physical activity readiness questionnaire (15). After reading about the experimental procedure, 26 subjects signed informed consent to participate in the study. Participants were students and office workers. The group included 18 females and 8 males with an average age of 28 ± 6 years and 34 ± 12 years (mean ± SD), respectively. A power analysis was performed (nQuery Advisor) and 23 subjects were needed to be able to find intraclass correlations with a 95% CI between 0.80 and 0.95, with a power of 0.80 and an α level of 0.05. nQuery Advisor calculates sample size by means of a patterned correlation model as recommended by Muller. The model is discussed by Murray et al. (16).

Study Protocol

Between April and June 2005 all subjects underwent evaluations of heart rate variability and respiration rate on two occasions, separated by 7 days. Subjects were asked to select two comparable days for the assessments. On each occasion a laboratory and an ambulatory measurement were performed. For the first assessment the participants visited the Academic Medical Center between 07:00 a.m. and 11:30 a.m. Exactly 7 days after the first assessment, the subjects again visited the Academic Medical Center to repeat the protocol at approximately the same time as the first assessment. After arrival, the test leader attached the Co²ntrol (Decon Medical Systems) and gave a short explanation of the procedure. A 35-min laboratory protocol was then performed consisting of the subject sitting for 5 min for acclimatization purposes; lying in a supine position on a bed in a noise-free environment for 10 min; performing light exercise for 15 min, cycling on a bicycle ergometer using a single load of 50 W with a pedal frequency of 60/min (the posture of the subjects was exactly the same on both measurement days); and sitting for 5 min to download and check the data gathered. After the laboratory procedure, the Co²ntrol was programmed to start measuring again during the evening and night for the ambulatory measurement. The participants were free to leave after the laboratory protocol to resume their usual daily pattern. They wore the Co²ntrol until the following morning. All subjects were asked to report the time they went to sleep.

Parameters

Evaluations of heart rate variability and respiration rate were performed in subjects following the protocol described above. The maneuvers (lying in a supine position, cycling on a bicycle ergometer, and sleep) were selected because they are used extensively as stimuli for cardiovascular control research. It is known that they induce specific changes in the sympatho-vagal balance and may reach a stable heart rate (17). Data were recorded using the Co²ntrol. The Co²ntrol is built according to the guidelines of the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1) and is able to store a large amount of data, making accurate 24-h measurements possible. The Co²ntrol uses a Polar HR “detection board” (PCBA receiver) for registration of RR intervals. The QRS detection timing accuracy and detection reliability of the detector system was tested with an artificially generated ECG signal. The tests indicated timing
errors < 1 msec that can be detected in real measurements even under noisy conditions (18). The device is attached to an elastic belt. The belt contains a stable case with heart rate electrodes and a polar HR transmitter (Polar T31 transmitter, Polar Electro, Almere, The Netherlands). The Co2ntrol defines the ‘normal to normal’ (NN) intervals, intervals between adjacent QRS complexes, with a sampling rate of 1 msec. The QRS detection algorithm uses optimized pre-filtering in conjunction with a matched filter and dual edge threshold detection (18). To measure respiration rate, the inhale and exhale times assessed by means of the chest extension were logged every 3 sec during the laboratory procedure and every 5 sec during the ambulatory measurements. The amplitude resolution of the Co2ntrol recording analogue to digital conversion is 10 bits or, i.e., 1024 points.

**Data Reduction and Analysis**

To define the heart rate variability parameters the raw data were transferred to the specially developed software named *Lifestylemanager* (Decon Medical Systems). The last 7 min of the 10 min of lying down were selected to define the laboratory rest values, the last 10 min of the 15 min of cycling were selected to define the laboratory light physical activity values and the second and fourth hours of sleep were selected to define the ambulatory rest values. For all four selected time periods the raw data were manually selected and, when present, disturbed signals were filtered out. Then the heart rate variability parameters, the standard deviation of the NN intervals (SDNN [msec]) and the square root of the mean of the sum of squares of differences between adjacent NN intervals (RMSSD [msec]) were defined with the *Lifestylemanager*. For the respiration rate values, respiration rate was sampled every 30 sec for the laboratory data and every minute for the ambulatory data in the Co2ntrol software (Decon Medical Systems). The mean of the sum per maneuver was used for analysis.

Reliability is one of two ways to quantify reproducibility. Measures of reliability refer to the variance in variation between patients in relation to the total variance of the measurements. Measures of agreement, the second way to quantify reproducibility, refer to the absolute measurement error that is associated with one measurement taken from one individual. Agreement provides information on whether a measurement device is able to achieve the same value in a subject over repeated measurements whereas reliability provides information on whether a measurement device can distinguish between persons (19). In this study we used the intra-class correlation coefficient (ICC) as a way to quantify reliability and Bland-Altman plots and standard error of measurement (SEM values) as a way to quantify agreement.

First, the means and standard deviations for each selected time period for both heart rate variability parameters (SDNN and RMSSD) and respiration rate were calculated. Then, for each subject, the means and differences between both assessments for each parameter were defined. Bland-Altman 95% limits of agreement and Bland-Altman plots (20,21) were computed. About 95% of the observations should lie within about 2 SD of the mean (22). Next, the ICCs were computed for the four selected time periods (laboratory rest, laboratory light physical activity, ambulatory rest night second hour and ambulatory rest night fourth hour). The measurement error is represented by the SEM, SEM values (the square root of the error-mean-square) were also calculated (23). Respiration rate values were then defined again with another sample rate. The mean respiration rate for the 7 min of laboratory rest was calculated with a sample rate of 1 min, and for the ambulatory data the mean of 7 min, randomly chosen from the second hour, was calculated, also using a sample rate of 1 min. Finally, ICCs between the laboratory rest and ambulatory rest conditions were calculated for both heart rate variability and respiration rate. Model 3.1 was used for all ICCs, as this has been recommended for reliability analysis (24). Data were considered reproducible if the ICC ranged from 0.60 to 0.81. If the ICC was > 0.81, reproducibility was considered excellent (12,25,26). Statistical analyses were done using SPSS version 12.0 for Windows (SPSS Inc., Chicago, IL).

**Results**

Means and standard deviations of the first and second evaluations for both heart rate variability parameters (SDNN and RMSSD) are presented in Table 1. Further, the number of measurements that were used for analysis, ICC, ICC 95% limits of agreement and SEM values are shown in Table 1.

For lying down in the laboratory and ambulatory night second hour, the ICCs of the mean SDNN values were considered good. For laboratory cycling and ambulatory night fourth hour, the ICCs were both > 0.81 and thus considered excellent. The lower bounds of the ICC 95% limits of agreement for the SDNN values for lying down in the laboratory and ambulatory night second hour were considered poor, whereas the lower bounds of the ICC 95% limits of agreement for laboratory cycling and ambulatory night fourth hour were considered good. The ICC of the RMSSD values for lying down in the laboratory was considered good. The ICCs of the RMSSD values for laboratory cycling, ambulatory night second hour and ambulatory night fourth hour were considered excellent. The lower bound of the ICC 95% limits of agreement for the RMSSD value of lying down in the laboratory was poor. The lower bound of the ICC 95% limits of agreement for laboratory cycling was considered good and for the ambulatory night second hour and the ambulatory night fourth hour, excellent. The
absolute measurement errors (SEM) for the SDNN values for lying down in the laboratory, ambulatory night second hour and ambulatory night fourth hour were between 12 and 14 ms. For the condition laboratory cycling it was 5.05 msec. Absolute measurement error for the RMSSD values for the condition laboratory cycling was 2.01 msec, for lying down in the laboratory 10.37 msec, for the condition ambulatory night second hour 6.29 msec and for the condition ambulatory night fourth hour it was 3.62 msec. Table 2 shows the data for respiration rate just as Table 1 shows for heart rate variability.

ICCs for respiration rate for the three selected time periods (laboratory cycling, lying down in the laboratory and ambulatory night fourth hour) were all 0.81 and thus excellent. For the condition laboratory cycling it was 0.77 and thus good. The lower bounds of the ICC 95% limits of agreement for laboratory cycling and ambulatory night fourth hour are >0.81, thus excellent. For lying down in the laboratory, the lower bound is 0.65 and thus good and for ambulatory night second hour 0.46, and thus poor. SEMs were all between 0.62 and 1.58 msec.

Results from comparison between laboratory rest data and ambulatory rest data are presented in Table 3. It shows for both heart rate variability parameters (SDNN and RMSSD) and respiration rate, the mean (standard deviation), as well as ICCs and ICC 95% limits of agreement. We found the ICC was around 0.60 for the heart rate variability parameters SDNN and RMSSD and for the respiration rate, comparing the laboratory rest values with the ambulatory night second hour values which make them poor to good. The lower bounds of the ICC 95% limits of agreement are very poor, whereas the higher bounds are very large.

Bland-Altman plots for the means against the differences of heart rate variability parameters SDNN and RMSSD and of respiration rate values are shown in Figure 1. For the selected time period laboratory lying down, the assumptions of the limits of agreement are met. For the selected time period laboratory cycling the assumptions of the limits of agreement are met for SDNN and respiration rate. For RMSSD (92%) the assumptions of the limits of agreement are not met. For the selected time period second hour of sleep the assumptions of the limits of agreement are met for the RMSSD values. For SDNN and respiration rate the assumptions are not met for SDNN (90%) and for respiration rate (94%). For the selected time period fourth hour of sleep the assumptions of the limits of agreement are met for RMSSD and respiration rate. For SDNN (89%) the assumptions for the limits of agreement are not met.

### Table 1. Heart rate variability

<table>
<thead>
<tr>
<th></th>
<th>Laboratory cycling</th>
<th>Laboratory lying down</th>
<th>Ambulatory night second hour</th>
<th>Ambulatory night fourth hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>26</td>
<td>25</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Mean (SD) SDNN (msec)</td>
<td>25.9 (11.6)</td>
<td>54.8 (27.5)</td>
<td>61.6 (34.2)</td>
<td>62.9 (35.0)</td>
</tr>
<tr>
<td>Mean (SD) RMSSD (msec)</td>
<td>9.8 (4.4)</td>
<td>30.2 (24.3)</td>
<td>29.8 (27.4)</td>
<td>32.7 (27.3)</td>
</tr>
<tr>
<td>ICC (ICC 95% limits of agreement)</td>
<td>0.85 (0.70–0.93)</td>
<td>0.74 (0.50–0.88)</td>
<td>0.79 (0.54–0.91)</td>
<td>0.85 (0.66–0.94)</td>
</tr>
<tr>
<td>SEM SDNN</td>
<td>5.05</td>
<td>12.39</td>
<td>13.95</td>
<td>12.35</td>
</tr>
<tr>
<td>SEM RMSSD</td>
<td>2.01</td>
<td>10.37</td>
<td>6.29</td>
<td>3.62</td>
</tr>
</tbody>
</table>

Number of measurements (N) used for analysis, means (standard deviations) of the RMSSD (msec) and SDNN (msec) values for the first and second measurement, intra-class correlations (ICC), ICC 95% limits of agreement, and standard error measurement values (SEM).

### Table 2. Respiration rate

<table>
<thead>
<tr>
<th></th>
<th>Laboratory cycling</th>
<th>Laboratory lying down</th>
<th>Ambulatory night second hour</th>
<th>Ambulatory night fourth hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>23</td>
<td>19</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mean (standard deviation)</td>
<td>20.7 (5.4)</td>
<td>14.1 (3.2)</td>
<td>15.1 (3.6)</td>
<td>14.1 (3.1)</td>
</tr>
<tr>
<td>Respiration rate (p/m)</td>
<td>20.7 (5.4)</td>
<td>14.4 (3.9)</td>
<td>14.6 (3.2)</td>
<td>14.1 (3.2)</td>
</tr>
<tr>
<td>ICC (ICC 95% limits of agreement)</td>
<td>0.92 (0.81–0.96)</td>
<td>0.85 (0.65–0.94)</td>
<td>0.77 (0.46–0.91)</td>
<td>0.96 (0.89–0.99)</td>
</tr>
<tr>
<td>SEM</td>
<td>1.58</td>
<td>1.38</td>
<td>1.63</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Number of measurements (N) used for analysis, means (standard deviations) of the respiration rate (breath frequency per minute) for the first and second measurement, intra-class correlations (ICC), ICC 95% limits of agreement, and standard error measurement values (SEM).
Table 3. Laboratory rest vs. ambulatory night

<table>
<thead>
<tr>
<th>N</th>
<th>SDNN (msec)</th>
<th>RMSSD (msec)</th>
<th>Respiration rate (p/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Mean | Laboratory lying down | 54.0 (30.4) | 30.5 (27.1) | 14.2 (4.1) |
| SD  | Ambulatory night      | 61.6 (34.2) | 29.8 (27.4) | 13.8 (2.8) |
| ICC (ICC 95% limits of agreement) | 0.62 (0.25–0.83) | 0.57 (0.18–0.80) | 0.59 (0.14–0.83) |

Number of measurements (N) used for analysis, means (standard deviations) for laboratory lying down and ambulatory night second hour, intra-class correlations (ICC) and ICC 95% limits of agreement for heart rate variability (SDNN and RMSSD) and respiration rate.

Discussion

This study examined test-retest reliability of heart rate variability and respiration rate measurements. Both parameters were found to be reliable during rest and light physical activity.

For all the selected time periods the ICCs of the heart rate variability parameters (SDNN and RMSSD) were found to be good or even excellent. Also, ICCs for respiration rate were found to be good or even excellent. This means that both time-domain heart rate variability (SDNN and RMSSD) and respiration rate can be reliably assessed with the Co2ntrol. Reliability is a parameter to examine whether a measurement device can distinguish between persons. The Co2ntrol can indeed be used to discriminate between subjects because the ICCs are good to excellent, which means that the absolute measurement error is small in relation to the variability between subjects.

Comparison of the laboratory rest values with the ambulatory night second hour values yielded ICCs from poor to good. The lower bounds of the ICC 95% limits of agreement were very low, whereas the upper bounds were very high. This means that new measurements can yield varied results, and that laboratory rest values are not very comparable to sleeping rest values.

The agreement parameters Bland-Altman plots and SEM provide information about the agreement between the repeated measurements. Bland-Altman plots provide a visualization of the agreement between the repeated measurements. Bland and Altman (22) reported that the assumptions of agreement are met when 95% of the observations lie between the 95% limits of agreement. Four plots do not meet this assumption but we still believe that the agreement found was reasonable because only 2 observations at maximum were found outside the limits (data not shown). In addition, because of the high mean score of some observations, a higher difference between the 2 measurements can be expected. SEM gives information about the size of the changes that can be detected when using the Co2ntrol for heart rate variability and respiration rate measurements. SEM should be smaller than the expected improvement or deterioration. This is important to know when the device will be used to track modifications in the clinical state.

The Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1) advised recordings of 5 min to standardize different studies investigating short-term heart rate variability. Unfortunately, the software developed to calculate the heart rate variability values, SDNN and RMSSD, needs 300 data points to calculate these values. The amount of data points available depends on the heart rate frequency. For subjects with a heart rate < 60 beats/min we needed a data selection of > 5 min. For this practical reason we decided to analyze rest values gathered over a period of 7 min instead of 5 min. Furthermore, the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1) state that frequency domain methods analyses are preferred to the time domain methods when investigating short-term recordings. However, they also mention that time domain methods, especially the SDNN and RMSSD methods, can be used for short durations. The Co2ntrol is developed to perform short- and long-term measurements. Because time domain methods are ideal for the analysis of long-term recordings and at the same time SDNN and RMSSD can be used for short-term measurements, the time domain method was chosen.

As can be seen in Tables 1 and 2, the number of subjects used for analysis was not the same for the ambulatory measurements as for the laboratory measurements. Software problems did not allow us to use the ambulatory data for six of the subjects for the analysis. Heart rate variability values obtained for one subject during laboratory rest and for another subject during the fourth hour of the night could not be used for analysis because the signals were interfered with. Respiration rate data for three subjects during laboratory cycling, as well as data for six subjects during laboratory rest, could also not be used as the signals were indistinct. For the ambulatory second and fourth hours of the night, the respiration rate signals for six subjects were indistinct and could not be included in the analysis. The indistinct respiration rate signals were mainly owing to the way the Co2ntrol is fastened. We expect that adapting the software will be able to address these issues. Despite some software problems, none of the subjects dropped out, which tells us that the assessment days were not too onerous. This is very important to know when working with patients.
Figure 1. Bland-Altman plots for heart rate variability (SDNN [msec] and RMSSD [msec]) and respiration rate (RR [breath frequency per minute]) values for laboratory lying down, laboratory cycling, ambulatory night second hour, and ambulatory night fourth hour. Mean differences (solid lines), 95% limits of agreement (dashed lines).
The findings of the present study show ICCs for SDNN values comparable to most of those reported in similar studies with healthy subjects. Marks et al. (12) and Schroeder et al. (13) reported that SDNN values for lying in a supine position were reproducible for durations of 5 and 6 min, respectively. Sandercoc et al. (2) evaluated SDNN values for lying in a supine position using three devices. They reported ICCs of 0.77 and 0.78 for two devices. For the third device they reported an ICC of 0.57, which is not as comparable. Pitzalis et al. (26) reported an ICC for SDNN values for this condition of 0.56. This value is also much lower than the value we found in our study.

Also, the ICCs for the RMSSD values for lying in a supine position were comparable to those reported in earlier studies. Schroeder et al. (13) and Carrasco et al. (11) reported reproducibility of RMSSD values for lying for 5 or 6 min in a supine position. Pitzalis et al. (26) reported an ICC of 0.23, which is very low compared to this and other studies. In addition to recording RMSSD values for lying in supine position, Carrasco et al. (11) also evaluated RMSSD for cycling on a bicycle ergometer. The reported ICC of 0.79 is comparable to the ICC of 0.85 found in this study. Unfortunately, there are no comparable studies measuring SDNN or RMSSD values during sleep.

We conclude from our findings that measurements of heart rate variability and respiration rate with the Co2ntrol are reliable. Therefore, it is a suitable device to discriminate among subjects. The ICCs comparing laboratory rest values and ambulatory night sleep values showed that it is important to take care when assuming that rest values measured in a laboratory are truly representative of actual rest values. This can be important for research that uses rest as a baseline. The study also shows better reliability for the laboratory cycling condition than for the condition lying down in the laboratory. Perhaps measurements taken during exercise yield a better baseline. The absolute measurement error (SEM) is useful for further research in a clinical setting. Changes in the variable larger than the SEM values indicate a modification in the clinical state because they cannot be the result of measurement error. When the technical problems are solved, the Co2ntrol should prove an excellent device for performing research when information on heart rate variability and respiration rate values are required in a clinical setting. The results also showed good reliability for respiration rate, which means that further research should be performed to examine whether respiration rate can be evaluated to track modifications in the clinical state.

Acknowledgments
We wish to thank Stans van der Poel and Desiree Schoordijk (Decon Medical Systems, Weesp, The Netherlands) for their help and for providing the measuring devices. Furthermore, we thank all participants for their cooperation.

References


