1. Introduction

Given that gravity is the driving force of convective heat transfer, a lack of gravity impairs the natural share of convective heat transfer from the body surface in a sustained manner. Consequently, under microgravity conditions astronauts/cosmonauts can experience substantial thermal discomfort, especially during extravehicular activities (Nicogossian et al., 1994). At present, air temperatures are regulated at between 18 and 29 °C with a maximum humidity of 72% and normobar–normoxic conditions. Nevertheless, findings from the space station MIR show that cosmonauts also complain about thermal discomfort at the extremities within thermoregulated environments. Hence, a straightforward global indicator to explore thermoregulation in astronauts/cosmonauts would be to continuously monitor core body temperature during spaceflights.

Usually, in research settings core body temperature is measured by inserting a thermosensor in the esophagus, nasopharynx, rectum, or tympanum/auditory meatus. However, these methods are neither really applicable during daily routines on ground nor in space. This is due to the fact that the requirements for a method serving to measure core body temperature are demanding: the thermosensor (i) has to be non-invasive, (ii) should be easy to operate, (iii) must fulfill basic hygiene standards, (iv) must not be biased towards various environmental conditions, (v) should be sensitive enough to quantitatively reflect minor changes in arterial blood temperature, and (vi) its response time should be as short as possible (Cooper et al., 1964; Shiraki et al., 1986). Very recently, we therefore presented a new method called Double Sensor, combining a skin surface temperature sensor with a heat flux sensor, to achieve this goal under various physical and environmental conditions (Gunga et al., 2005, 2008). Based on this experience we decided to use the Double Sensor during long-term bed-rest to establish whether rectal temperature recordings in humans could be replaced by a non-invasive skin temperature sensor combined with a heat flux sensor (Double Sensor) located at the forehead to monitor core body temperature changes due to circadian rhythms. Rectal and Double Sensor data were collected continuously for 24 h in seven men undertaking strict head-down tilt bed-rest. Individual differences between the two techniques varied between −0.72 and +0.55 °C. Nonetheless, when temperature data were approximated by cosinor analysis in order to compare circadian rhythm profiles between methods, it was observed that there were no significant differences between mesor, amplitude, and acrophase (P > 0.310). It was therefore concluded that the Double Sensor technology is presently not accurate enough for performing single individual core body temperature measurements under resting conditions at normal ambient room temperature. Yet, it seems to be a valid, non-invasive alternative for monitoring circadian rhythm profiles.

© 2009 Elsevier B.V. All rights reserved.
2. Methodology

2.1. Berlin BedRest Study 2 (BBR2-2)

The study was conducted by the Centre of Muscle and Bone Research and performed at the University Hospital Charité Campus Benjamin Franklin in Berlin during the years 2007–2008. The present validation study was a prerequisite for the evaluation of the interaction of thermoregulation and different types of strength training interventions during simulated micro-g conditions of the BBR2-2. These results as well as the distinct responses of skin and core temperature as well as heat flux will be discussed in a separate paper and are therefore omitted here. The BBR2-2 was divided into four consecutive campaigns of six subjects each. Subjects underwent 60 days of 6° head-down tilt bed-rest. Rectal and Double Sensor temperature recordings were obtained for a period of 36 h on the following head-down tilt bed-rest days 6–8 (HDT6–8), 20–22 (HDT20–22), 34–36 (HDT34–36), and 48–50 (HDT48–50) (except for the first campaign in which only three testing periods could be obtained: on HDT6–8, HDT34–36, and HDT48–50). The measurements started at Saturday evening at 19:30 h and lasted until Monday morning at 06:30 h (except for the first testing period of the first campaign where the measurement started at 13:00 h until 01:00 h). In order to analyse circadian rhythms for a period of 24 h, data were extracted between 19:30 and 19:00 h the following evening. Furthermore, for each subject only data from a single period of 24 h (either HDT6–8, HDT20–22, HDT34–36, or HDT48–50) was included in final statistical analyses (see Section 2.2 for details).

2.2. Subjects

Each individual campaign of the BBR2-2 included two male subjects in the control group and two intervention groups, respectively. In the present paper data are reported for the control group only. Since one subject in the fourth campaign could not undergo the intervention of the experimental group, it was opted to include him into the control group as well. Thus, a total of nine healthy Caucasian men ranging in age from 21 to 44 years were included in the control group. Exclusion criteria relevant to the current work were any cardiovascular or metabolic diseases, regular consumption of medication or orthopedic problems. Following a screening of their medical history and after the purpose, procedures, and known risks of the tests had been explained each participant gave written informed consent. The following limitations took place leading to non-obtainable data: during the first measurement (HDT6–8) of the first campaign only one of the two subjects agreed to use the rectal probe; during the third measurement (HDT34–36) of the second campaign one system malfunctioned; during the fourth measurement (HDT48–50) of the third campaign one subject was ill. Additionally, several data recordings were incomplete due to poor subject compliance. For the final sample, out of the total of 36 data sets of 24-h periods (9 subjects × 4 HDT) only one 24-h period per subject was considered for final analyses. Thus, based on an inclusion-threshold of 90% of data availability for a 24-h period, complete data were obtained for a single subject on HDT6–8, for two subjects on HDT20–22, for one subject on HDT34–35, and for two subjects on HDT48–49, resulting in a total sample of seven subjects (corresponding to seven 24-h periods). All procedures were approved by the ethics committee of the Charité University Hospital Berlin.

2.3. Technical devices

2.3.1. Rectal temperature sensor

Rectal temperature (TREC) was recorded at a depth of 50 mm past the anal sphincter using 4-mm NTC-thermal sensors (YSI 400 compatible, BlueTemp® products, bluepoint medical GmbH & Co. KG, Selmsdorf, Germany, accuracy: ±0.1 °C between 25 and 50 °C). Due to the rectal probes’ natural bending, no further fixation was necessary. The subject could wear their normal underwear/casual clothes. Data were collected at a frequency of 2 Hz, stored into flash memory systems (data logger, Heally-Sat Koralewski Industrie Elektronik, Hambühren, Germany), and subsequently transferred to a personal computer. Data for final analysis were limited to measurements obtained at consecutive 30-min intervals between 19:30 and 19:00 h the following evening.

2.3.2. Combined skin and heat flux sensor (Double Sensor)

Temperature recordings employing the Double Sensor (TD) were performed at the forehead on the vertical line above the eye directly underneath the hairline (Fig. 1B). Details of the underlying biophysical model are given in Gunga et al. (2008). In contrast to similar methodological attempts in the past (Fox and Solman, 1971; Smith et al., 1980; Taylor et al., 1998), the zero heat flux sensor principle of the Double Sensor (Patent Draegerwerk No. DE 100 38 247, DE 101 39 705, 2003) has been miniaturized and used without extra heating, and has been specially sealed (Fig. 1A).

As opposed to our previous study (Gunga et al., 2008), the Double Sensor was not integrated into the straps of a helmet. Instead the sensor was applied to the subject’s forehead using a two-sided adhesive ring tape that left the sensor area open to contact the forehead skin. Another adhesive tape securely fixed the sensor to the subjects head (Fig. 1B). Data collection and handling were performed in accordance with rectal temperature processing.

2.4. Statistics

All statistical analyses were performed using the SPSS software (Version 16.00, SPSS Inc., Chicago, IL, USA). Descriptive data are
reported as means and standard deviations. Measurements >38.5 °C and <36.0 °C were considered as artefacts and deleted from further analysis. Additionally, the remaining data was examined for outliers by identifying cases more distant than >1.5 interquartile ranges from the 75th and 25th quartile. Subsequently, scatter plots and correlation analysis was used to examine the relationship between rectal and Double Sensor temperature recordings. Additionally, the concordance correlation coefficient (CCC) was computed as suggested by Lin (1989, 2000) to obtain a measure of association that includes a correction factor which indicates how far the regression line deviates from the line of identity. Agreement between the two methods was further assessed by plotting their differences against their means (Bland and Altman, 1999). Limits within ±0.5 °C were defined a priori as clinically acceptable. Bias was tested for statistical significance using a paired t-test, while correlation analysis was employed to examine any trends for over- or underestimation of the reference method. Subsequently, cosinor analysis was performed within subjects for each temperature sensor to quantify circadian rhythm. Specifically, each individual time series data was fit to the following cosine curve (TableCurve 2D, V. 5.01, Systat Software Inc., Point Richmond, CA, USA):

\[ y(t) = M + A \times \cos(\omega t - \phi) \]

where \( t \) is the time of measurement, \( M \) the mean level of the cosine curve and termed mesor (midline estimating statistic of rhythm), \( A \) is the amplitude, \( \omega \) is the angular frequency (\( \omega = 2\pi/T \)), \( T \) is the period (i.e. 24h and corresponding to 24h) and \( \phi \) is the acrophase of the curve. After verification of the presence of a significant circadian rhythm by testing the zero-amplitude assumption (no rhythm) using \( F \)-statistics, mesor, amplitude and acrophase as well as goodness-of-fit indices (residual summed squares (SSRES) and total variance accounted for by the fitted cosine function \( R^2 \)) were obtained for each subject and each temperature variable and summarized as group means and standard deviations (i.e. median for SSRES and \( R^2 \)) for further analyses. Wilcoxon signed rank tests were employed to test fitting parameters and goodness-of-fit indices (mesor, acrophase, amplitude, SSRES, and \( R^2 \)) between the two temperature recording techniques for significant differences. Finally, bias and limits of agreement were reported for the fitted cosine curve parameters to compare the random error component between methods.

3. Results

Subject characteristics \((n = 7)\) are given in Table 1. Age ranged between 21 years and 42.6 years, with four subjects being older than 30 years. Based on their body mass index (BMI), two of the subjects were considered as slightly overweight (BMI: 27.0 and 28.3 kg/m²), while two subjects were rather lean (BMI: 22.6 and 23.1 kg/m²). Due to high compliance both temperature recording techniques, the rectal probe as well as the Double Sensor, were generally well accepted by the subjects throughout the study. A total of 322 measurements for rectal and Double Sensor temperature recordings, respectively, were included in final analyses.

Temperature recordings from the Double Sensor were moderately related to rectal temperature measurements \((r = 0.704, P < 0.001)\). The corresponding scatter plot is given in Fig. 2A. A number of data points \((n = 19)\) exceeding the lower 95% confidence interval deserved closer inspection (grey-shaded data points). It was found that this scatter was predominantly caused by single subject during the interval between 23:30 and 09:00h. Given the better association for the subject’s remaining data, it can be speculated that the rectal probe might not have been correctly fixed in its original position. Deleting the data cloud increased the correlation coefficient to \( r = 0.803; P < 0.001 \). Similar findings were observed for CCC, which increased from its initial level of 0.678–0.773.

Agreement between methods is indicated in Fig. 2B. There was a slight \((0.08 ± 0.32 °C)\), but significant difference between \(T_{REC}\) and \(T_{DS}\) \((P < 0.001)\). 95% of the differences were located between −0.72 and +0.55 °C. Furthermore, as shown in the folded cumulative distribution plot (Fig. 3) 85% of the differences varied between the \(a\) priori defined cut-off values of ±0.5 °C, whereas about 69% lay between ±0.25 °C. Omission of the apparent extreme values indicated that the rectal probe might not have been correctly fixed in its original position. Deleting the data cloud increased the correlation coefficient to \( r = 0.803; P < 0.001 \). Similar findings were observed for CCC, which increased from its initial level of 0.678–0.773.

\[ \text{Fig. 2. (A) Scatter plot of rectal and Double Sensor temperature recordings. } T_{REC}, \text{ rectal temperature. } T_{DS}, \text{ Double Sensor temperature. (B) Bland–Altman plot of the difference between the methods against the mean of methods. } T_{REC}, \text{ rectal temperature. } T_{DS}, \text{ Double Sensor temperature. The solid line indicates the bias between methods and dashed lines are 95% limits of agreement (±1.96 S.D.). Proportional error is indicated by a significant relationship between differences and means of methods. Extreme values are grey-shaded. Reported statistics are based on all data shown.} \]

### Table 1

Descriptive subject characteristics (means ± S.D.).

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Height (m)</th>
<th>Body mass (kg)</th>
<th>BMI (kg/m²)</th>
<th>(T_{REC}) (°C)</th>
<th>(T_{DS}) (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.9 ± 8.0</td>
<td>1.79 ± 0.04</td>
<td>80.8 ± 5.6</td>
<td>25.3 ± 2.0</td>
<td>36.82 ± 0.37</td>
<td>36.91 ± 0.45</td>
</tr>
</tbody>
</table>

\(T_{REC}\), rectal temperature; \(T_{DS}\), Double Sensor temperature.

* Significantly different from \(T_{REC}\) \((P < 0.001)\).
Folded cumulative distribution plot for differences between rectal and Double Sensor temperature recordings. Percentiles were computed for each ranked difference, whereas for all percentiles exceeding 50, percentiles were ‘folded’, i.e. transformed as 100—percentile and then plotted against the differences.

cated above, 90% of the data were within limits of ±0.5 °C, and 73% within limits of ±0.25 °C.

Visual inspection revealed that the Double Sensor underestimates rectal temperature at lower temperatures and overestimates rectal temperature at higher temperatures. This proportional error was also confirmed by linear regression analysis ($r = -0.248, P < 0.001$). Recalculation of the limits of agreement after data deletion of the apparent artefacts (grey-shaded data points) according to Fig. 2A reduced the upper bound to +0.49 °C and the lower bound to −0.57 °C. Graphical inspection of time series data for each individual subject revealed distinctive circadian patterns. This rhythm was confirmed by cosinor analysis as amplitudes were significantly different from 0 for all subjects ($P < 0.001$). Fig. 4 shows cosinor analysis for an individual subject for both rectal and Double Sensor recordings.

The high degree of visual agreement between measurements indicated in Fig. 4 was also confirmed by parameters of the fitted cosine curves ($y(t) = 36.63 + 0.53 \cos(t + 20.68); R = 0.979$ and $y(t) = 36.63 + 0.50 \cos(t + 20.46); R = 0.962$ for rectal and Double Sensor data, respectively). This finding was also observed in the total group. Scatter plots and group cosinor analyses for rectal and Double Sensor temperature recordings are given in Fig. 5A and B.

Comparison of curve characteristics between rectal and Double Sensor recordings is given in Table 2. When circadian temperature profiles were quantified by mesor, acrophase, and amplitude, no significant difference were found between rectal and Double Sensor temperature recordings ($P = 0.310–0.866$).

Further information about the degree of agreement between methods is provided by the limits of agreement. Of particular interest is the random error component for acrophase and amplitude. As indicated in Table 2, individual differences for amplitude range between −0.21 and 0.19 °C. More importantly, differences in acrophase were as low as about 1 h (1.18–1.49 h). Finally, it should be noted though that Double Sensor recordings were characterized by higher variation compared to rectal temperature measurements (Fig. 5), yielding a poorer model fit. This was also statistically confirmed by a significantly greater amount of total variance accounted for by the rectal cosine curve model (0.93 vs. 0.79, $P < 0.05$) and significantly lower residual summed squares (0.48 vs. 1.69, $P < 0.05$) compared to the Double Sensor fitted cosine curve.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Rectal (°C)</th>
<th>Double Sensor (°C)</th>
<th>Bias (°C)</th>
<th>LoA* (°C)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesor (°C)</td>
<td>36.79 ± 0.13</td>
<td>36.90 ± 0.24</td>
<td>−0.11</td>
<td>−0.67 to 0.45</td>
<td>0.310</td>
</tr>
<tr>
<td>Acrophase (rad)</td>
<td>5.55 ± 0.21</td>
<td>5.51 ± 0.22</td>
<td>0.04</td>
<td>−0.31 to 0.39</td>
<td>0.735</td>
</tr>
<tr>
<td>Amplitude (°C)</td>
<td>0.48 ± 0.07</td>
<td>0.49 ± 0.08</td>
<td>−0.01</td>
<td>−0.21 to 0.19</td>
<td>0.866</td>
</tr>
</tbody>
</table>

* LoA, limits of agreement (bias ± 1.96 S.D.).
4. Discussion

Monitoring circadian rhythms is not only important for understanding thermoregulatory and cardiovascular adaptation processes in space but as well for exploring physiological problems associated with microgravity conditions such as body weight loss and de-synchronisation of internal rhythms, which might affect the overall health and performance of astronauts/cosmonauts during long-term space missions.

Given the number of drawbacks of present technological advances for monitoring core body temperature – under rest and exercise – there is a great demand for developing an easy-to-operate and non-invasive technology to measure core body temperature in humans. In the present study we evaluated a new skin temperature and heat flux measurement device, called Double Sensor, to monitor core body temperature changes during 24 h 6-h head-down tilt bed-rest. Though 95% of individual differences between the methods ranged between −0.72 and +0.55 °C, it should be noted that 85% (i.e. 90% after omission of apparent extreme values) of the differences were located within the a priori defined limits of ±0.5 °C. In this regard it seems noteworthy that rectal temperature recordings themselves are not without error. In contrast, oesophageal temperature can be considered as ‘Gold Standard’ for monitoring core body temperature (Imrie and Hall, 1990). Given that data from patients after postoperative rewarming shows that 95% of individual differences between oesophageal and rectal temperature can be as high as ±0.82 °C (Braeuer et al., 1997), the predefined limits of clinical acceptance of ±0.5 °C should be considered as rather conservative and the observation that in the present study 85% of the differences were located within this interval seems to be very promising. Furthermore, it was found that circadian core temperature profiles could be well approximated by the Double Sensor. There were no significant mean differences for cosine curve fitted parameters (P = 0.310–0.866) and differences in amplitude and acrophase for an individual were as low as about 0.2 °C and 1 h, respectively. It should be noted that Double Sensor recordings were characterized by higher variability compared to rectal measurements (Fig. 5) though. Given that increased variability will reduce the sensitivity to detect changes in body temperature, the higher variability observed for the Double Sensor compared to rectal temperature recordings deserves closer inspection. Though standard deviations for mesor, acrophase, and amplitude did not significantly differ between rectal and Double Sensor methodology (P = 0.160, P = 0.913, and P = 0.754 for mesor, acrophase, and amplitude, respectively), power analysis was performed to quantify the impact of the higher variability of the Double Sensor methodology on outcomes of prospective studies. Based on a two-sided paired t-test, a level of significance of 0.05, a power of 0.80, and an observed difference of 0.5 °C for mesor, a sample size of n = 5 and n = 3 for Double Sensor and rectal temperature methodology, respectively, would be needed to obtain a significant difference. Similarly, to detect a significant difference of 0.25 °C in amplitude and a significant difference of 0.5 rad (corresponding to approximately 2 h) in acrophase, sample sizes of n = 3 for both rectal and Double Sensor methodology, and n = 5 and n = 4 for Double Sensor and rectal methodology would be required, respectively. Thus, even if the available pool of subjects is small, we are confident that the Double Sensor technology is accurate enough for detecting relatively small meaningful differences in circadian rhythm profiles compared to rectal temperature recordings. Thus, in spite of relative marked variability for single ‘spot checks’, we therefore conclude that the Double Sensor placed at the forehead seems to be a valuable approach for monitoring 24-h core body temperature changes (circadian rhythm) during long-term 6-h head-down tilt bed-rest.

These results confirm the findings from our previous study where we validated the Double Sensor during treadmill exercise (Gunga et al., 2008) with limits of agreement ranging between −0.90 and +1.6 °C. It should be noted though that poor subject compliance might have affected the accuracy of both rectal and Double Sensor recordings and generated additional artefacts. For instance, as indicated in Fig. 2A, a number of extreme data points could be identified that pertained to a specific time interval of an individual subject. Deletion of this subject increased the correlation between rectal and Double Sensor measurements to r = 0.803 (CCC = 0.773) and decreased the limits of agreement for single measurements to +0.49 °C and the lower bound to −0.57 °C. Thus, individual differences between the two methods might have been inflated by poor subject compliance. Furthermore, with regard to the Double Sensor it seems noteworthy that some subjects showed slight skin irritation after long-term application of the Double Sensor, which might have further jeopardized agreement between methods. These side-effects are, however, relatively mild compared to the limitations associated with other approaches outlined above. Finally, as indicated by the significant correlation between the differences of methods and the means of methods (r = −0.248, P < 0.001), there was significant proportional error. With increasing temperature, Double Sensor recordings tended to overestimate core body temperature, indicating the need to specifically verify the accuracy of the Double Sensor technology at higher core body temperatures. Presently, however, we cannot verify any of these assumptions and further studies are necessary to assess the effects of improved sensor fixation and careful assessment of subject compliance.

In conclusion, the knowledge and technological development gained from the study could promote applications of continuous core body temperature measurements in clinical, occupational, sports and environmental medicine on earth and in space. The Double Sensor device might be an effective, non-invasive and easy-to-operate technology to gain fundamental insights into cardio-circulatory regulation, thermoregulation, and circadian rhythms in humans and we suggest further studies to fully reveal the underlying potential of the technique by improving its technology and testing it in various subject cohorts and diverse settings under different conditions.

Acknowledgements

We wish to express our thanks to the whole Berlin BedRest Study team, as well as Mrs. Himmelsbach-Wegner and Mrs. Eveline Hofmann for their help to prepare the manuscript. Furthermore, we would like to acknowledge outstanding support and cooperation from all subjects and thank them for dedicating their time and effort towards the successful completion of the study. Daniel L. Belavý was supported by a post-doctoral fellowship from the Alexander von Humboldt Foundation. The 2nd Berlin Bed-Rest Study (BBR2-2) was supported by grant 14431/02/NL/SH2 from the European Space Agency and grant 50WB0720 from the German Aerospace Center (DLR). The 2nd Berlin Bed-Rest Study was also sponsored by Novotec Medical, Charité Universitätsmedizin Berlin, Siemens, Osteomedical Group, Wyeth Pharma, Servier Deutschland, P&G, Kubitvent, Seca, Astra-Zeneka and General Electric. In recent years Prof. Gunga has also been serving as a consultant to the Draegerwerk AG & Co. KGaA. Dr. Sattler and Dr. Koch are employees of the Draegerwerk AG & Co. KGaA. The other coauthors do not have any financial interest in any company related to this manuscript.

References


