A Simple and Valid Method to Determine Thermoregulatory Sweating Threshold and Sensitivity


Thermal and Mountain Medicine Division
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760-5007

Approved for public release; distribution unlimited

Sweating threshold temperature and sweating sensitivity responses are measured to evaluate thermoregulatory control. However, analytic approaches vary, and no standardized methodology has been validated. This study validated a simple and standardized method, segmented linear regression (SReg), for determination of sweating threshold temperature and sensitivity. Archived data were extracted for analysis from studies in which local arm sweat rate (m(sw)) and esophageal temperature (Tes) were measured under a variety of conditions. The relationship m(sw)/Tes from 16 experiments was analyzed by seven experienced raters (Rater), using a variety of empirical methods, and compared against SReg for the determination of sweating threshold temperature and sweating sensitivity values. Individual interrater differences (n=324 comparisons) and differences between Rater and SReg (n=110 comparisons) were evaluated within the context of biologically important limits of magnitude (LOM) via a modified Bland-Altman approach. The average Rater and SReg were not different by conventional t-test (P >0.05). SReg provides a simple, valid, and standardized way to determine sweating threshold temperature and sweating sensitivity values for thermoregulatory studies.

thermoregulation; heat stress; exercise; segmented linear regression

Unclassified

M09-27
**INSTRUCTIONS FOR COMPLETING SF 298**

1. **REPORT DATE.** Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. **REPORT TYPE.** State the type of report, such as final, technical, interim, memorandum, master’s thesis, progress, quarterly, research, special, group study, etc.

3. **DATES COVERED.** Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. **TITLE.** Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. **CONTRACT NUMBER.** Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. **GRANT NUMBER.** Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. **PROGRAM ELEMENT NUMBER.** Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. **PROJECT NUMBER.** Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. **TASK NUMBER.** Enter all task numbers as they appear in the report, e.g. 05; RFO330201; T4112.

5f. **WORK UNIT NUMBER.** Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. **AUTHOR(S).** Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. **PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES).** Self-explanatory.

8. **PERFORMING ORGANIZATION REPORT NUMBER.** Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. **SPONSOR/MONITOR’S ACRONYM(S).** Enter, if available, e.g. BRL, ARDEC, NADC.

10. **SPONSOR/MONITOR’S REPORT NUMBER(S).** Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

11. **SUPPLEMENTARY NOTES.** Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

12. **ABSTRACT.** A brief (approximately 200 words) factual summary of the most significant information.

13. **SUBJECT TERMS.** Key words or phrases identifying major concepts in the report.

14. **SECURITY CLASSIFICATION.** Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

15. **LIMITATION OF ABSTRACT.** This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

**Standard Form 298 Back (Rev. 8/98)**
A simple and valid method to determine thermoregulatory sweating threshold and sensitivity

Samuel N. Cheuvront,1 Shawn E. Bearden,2 Robert W. Kenefick,1 Brett R. Ely,1 David W. DeGroot,1 Michael N. Sawka,1 and Scott J. Montain1

1United States Army Research Institute of Environmental Medicine, Natick, Massachusetts; and 2Department of Biological Sciences, Idaho State University, Pocatello, Idaho

Submitted 9 March 2009; accepted in final form 4 May 2009

Cheuvront SN, Bearden SE, Kenefick RW, Ely BR, DeGroot DW, Sawka MN, Montain SJ. A simple and valid method to determine thermoregulatory sweating threshold and sensitivity. J Appl Physiol 107: 69–75, 2009. First published May 7, 2009; doi:10.1152/japplphysiol.00250.2009.—Sweating threshold temperature and sweating sensitivity responses are measured to evaluate thermoregulatory control. However, analytic approaches vary, and no standardized methodology has been validated. This study validated a simple and standardized method, segmented linear regression (SReg), for determination of sweating threshold temperature and sensitivity. Archived data were extracted for analysis from studies in which local arm sweat rate (m˙sw; ventilated dew-point temperature sensor) and esophageal temperature (Tes) were measured under a variety of conditions. The relationship m˙sw/Tes from 16 experiments was analyzed by seven experienced raters (Rater), using a variety of empirical methods, and compared against SReg for the determination of sweating threshold temperature and sweating sensitivity values. Individual interrater differences (n = 324 comparisons) and differences between Rater and SReg (n = 110 comparisons) were evaluated within the context of biologically important limits of magnitude (LOM) via a modified Bland-Altman approach. The average Rater and SReg outputs for threshold temperature and sensitivity were compared (n = 16) using inferential statistics. Rater employed a very diverse set of criteria to determine the sweating threshold temperature and sweating sensitivity for the 16 data sets, but interrater differences were within the LOM for 95% (threshold) and 73% (sensitivity) of observations, respectively. Differences between mean Rater and SReg were within the LOM 90% (threshold) and 83% (sensitivity) of the time, respectively. Rater and SReg were not different by conventional t-test (P > 0.05). SReg provides a simple, valid, and standardized way to determine sweating threshold temperature and sweating sensitivity values for thermoregulatory studies.

thermoregulation; heat stress; exercise; segmented linear regression

BODY TEMPERATURE IS BELIEVED to be regulated by a proportionate control system (12, 14), defined as the graded response of a controlled variable (e.g., sweating) to the displacement of the regulated variable (e.g., body temperature). Both peripheral and central thermal receptors provide afferent input to the hypothalamus from which the resultant effector signal is initiated (16). The fundamental elements of this model include the core temperature (predominant influence) at which sweating increases above some baseline (threshold), and the change in gain or slope of the sweating response once initiated (sensitivity) (3, 12). Changes in either or both threshold and sensitivity of sweating can be observed in response to dehydration (10, 11, 23, 31), heat acclimatization (1, 28, 29, 34), exposure to altitude (19), physical training (29), age (1), exercise intensity (23), sleep loss (30), sex hormones (8, 20, 35), circadian (34, 36, 39), as well as several nonthermal factors (17, 18, 21), to name a few (33). Yet, despite extensive study of factors that influence thermoregulatory sweating, no standardized methods have been employed or validated to determine sweating threshold temperature and sensitivity. Potential differences among the many apparent nonstandardized methods have also not been evaluated.

Quantitative methods to determine thermoregulatory sweating variables currently include combinations of visual estimation, least squares regression, and empiricism. To determine threshold temperature, a decision is made regarding the onset of external thermoregulatory sweating. This can be a local sweating rate (m˙sw) that exceeds nonsweating transepidermal water loss (7, 22), a purely statistical deviation of successive points beyond random mathematical variation (4), or principled extrapolation to zero central drive (27), which accounts for the latency period from sweat gland stimulation to sweat emergence (7, 27). In addition, any m˙sw value in excess of the detection limit of a dew point capsule may also be used (13, 19, 20, 23, 31), as this will be slightly above levels of resting transepidermal water loss in temperate environments (5, 22). Very often, however, the methodologies employed are not provided in detail. Sweating sensitivity is more consistently calculated from the slope of the linear regression line, which defines the relationship between local m˙sw and esophageal temperature (Tes); however, it is rarely reported which portion of the gain response (early, late, combined) (21) is used for slope calculations.

Any empirical method for determining the thermoregulatory sweating threshold temperature or sweating sensitivity should be reproducible when applied to relatively unchanged circumstances (e.g., Ref. 4). But once experimental effects are imposed for a comparison of at least two trials, methodological differences among empirical procedures can potentially contribute to measurement error (noise). This is important, because differences in the interpretation of an experimental effect (signal) depend largely on the measurement variability (signal-to-noise ratio), particularly where purely statistical conclusions are drawn. If the level of agreement is poor or inconsistent among the different procedures used by investigators all trying to interpret the same experimental effect, it is possible that different conclusions may be drawn owing solely to methodology. Therefore, an important information void can be filled by quantifying the magnitude of differences among empirical procedures. This information also provides the first step for
establishing the data quality against which a more standardized procedure can be compared.

The identification of a standardized and valid method for determining sweating threshold temperature and sensitivity holds great potential for advancing the objective assessment of when an intervention or stressor has modified control. Thermoregulatory sweating, like many other biological effector systems (37), can be viewed as two line segments with a breakpoint. Baseline sweating has an approximate slope of zero, which changes in abrupt monotonic fashion in response to heat stress. The “hockey stick” shape of the relationship $\dot{m}_{sw}/T_{es}$ makes it possible to identify a breakpoint in the two lines using segmented regression (SReg) (15). A similar approach (piecewise linear regression) was advocated 20 years ago for fitting similar biological responses (37), but this and the variety of other available regression-based algorithms vary in the constraints, assumptions, and level of difficulty associated with their use. Given the basic nature of thermoregulatory sweating and its measurement (7, 13, 22, 26), it seems probable that use of simplified SReg will provide a standardized solution to this measurement dilemma.

The first purpose of this study was to evaluate between-investigator differences in empirically measured sweating threshold and sensitivity values. To accomplish this, several trained and experienced researchers (Rater) determined the sweating threshold temperature and sensitivity values from a preselected set of data. The second purpose of this study was to compare Rater-derived outcomes to those obtained using SReg on the same data. The third purpose was to qualitatively compare these findings to meaningful biological limits of magnitude (LOM) to appreciate the potential importance of the observed differences. We hypothesized that SReg would provide a standardized and valid way to determine sweating threshold temperature and sensitivity values in thermoregulatory studies.

![Fig. 1. Sixteen separate observations of local sweating rate ($\dot{m}_{sw}$) vs. esophageal temperature ($T_{es}$). Panels are truncated to 1 mg·cm$^{-2}$·min$^{-1}$ on the y-axis to optimize simultaneous display of all graphs on the same scale. All x-axis data contained at least 4 min of a discriminate baseline $\dot{m}_{sw}$, but may not be apparent (i.e., panels 12 and 14) due to the narrow $T_{es}$ ranges in those panels (~0.30°C) plotted on a requisite wider scale (1.50°C). All rater evaluations were performed with complete data sets to afford graphing and analysis on any desired scale.](https://www.jap.org/article.figures)
METHODS

Subjects and data. The volunteers for this study were 9 men (mean ± SD; age 23 ± 2 yr, height 175 ± 2 cm, weight 77.3 ± 4.5 kg, maximum oxygen uptake 52 ± 7 ml·kg⁻¹·min⁻¹), who gave voluntary, informed, written consent to participate in the original studies, where measurements of sweating threshold temperature and sensitivity were made. Research was conducted under the provisions outlined in Army Regulation 70–25 and US Army Medical Research and Materiel Command Regulation 70–25 and was approved by the US Army Research Institute of Environmental Medicine Human Use Review Committee.

Data sets (Fig. 1) from volunteers performing past experiments were selected from US Army Research Institute of Environmental Medicine study archives. Table 1 describes and contrasts the experimental conditions that make up the 16 data sets selected for analysis. Data selection was predicated on representing a variety of \( \dot{m}_{sw}/T_{es} \) responses under conditions known to influence baseline transepidermal diffusion rates, sweating threshold temperature, sensitivity, and strong peripheral modifiers of these responses (5, 6, 22, 23, 27, 28).

Measurements. \( T_{es} \) was obtained from a thermistor placed within the esophagus at a depth equal to ~25% of standing height (38). \( T_{es} \) measurements were employed because their response is more rapid and, therefore, preferred for sweating threshold temperature and sensitivity analysis (32). Volunteers spit into a cup during data collection to avoid spurious lowering of \( T_{es} \) from swallowing. If inadvertent swallowing occurred, the artifact was removed. \( T_{es} \) was measured continuously and recorded at 1-min intervals. The \( \dot{m}_{sw} \) of the upper arm was determined from a continuously ventilated dew point temperature sensor within a 15.9-cm² capsule (13). The flow rate through the capsule was modified (based on pilot testing) between different flow rates to optimize the dew point temperature rise within the capsule. Dew point temperature was interfaced to a data-acquisition system for continuous measurement. The \( \dot{m}_{sw} \) was analyzed from 1-min sampling intervals, since this interval is well within the latency period that extends between sweat gland stimulation and sweat emergence (7). It was also selected for consistency, since more frequent samples were unavailable from all of the archived data sets. Although SReg results were very similar between 20-s and 1-min samples, where this comparison was possible (unpublished observations), 1-min samples displayed less scatter and afforded more analytic certainty related to the assumptions of ordinary regression (24).

Table 1. Descriptive conditions that produced the panels represented in Figure 1

<table>
<thead>
<tr>
<th>Panel No.</th>
<th>Intensity, %VO₂max</th>
<th>Mode (TM/CE)</th>
<th>Dehydration, %body mass</th>
<th>Acclimatization (Y/N)</th>
<th>Environment (T°, %RH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>TM</td>
<td>0</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>TM</td>
<td>3</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
<td>TM</td>
<td>0</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>TM</td>
<td>3</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>7</td>
<td>45</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>8</td>
<td>45</td>
<td>TM</td>
<td>0</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>CE</td>
<td>0</td>
<td>N</td>
<td>20, 50</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>CE</td>
<td>0</td>
<td>N</td>
<td>40, 20</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>CE</td>
<td>0</td>
<td>N</td>
<td>20, 50</td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td>CE</td>
<td>0</td>
<td>N</td>
<td>40, 20</td>
</tr>
<tr>
<td>15</td>
<td>25</td>
<td>TM</td>
<td>0</td>
<td>Y</td>
<td>30, 50</td>
</tr>
<tr>
<td>16</td>
<td>25</td>
<td>TM</td>
<td>5</td>
<td>Y</td>
<td>30, 50</td>
</tr>
</tbody>
</table>

\( \dot{m}_{sw} \), maximum \( \dot{O}_2 \) uptake; TM, treadmill; CE, cycle ergometer; Y/N, Yes/No; \( T_{db} \), dry bulb temperature; RH, relative humidity.

RImax measures were truncated to remove obvious plateau data, but no consideration was given to distinguishing between early or late portions of the linear phase of the sweating transient (21).

Raters. Seven investigators experienced with local sweating measurements and skilled in biological threshold detection agreed to independently estimate the \( T_{es} \) sweating threshold temperature and calculate sweating sensitivity (slope) using 16 data sets that included only \( T_{es} \) (°C), \( \dot{m}_{sw} \) (mg·cm⁻²·min⁻¹), and time (min). These raters (Rater) were permitted to use any empirical method they wished and provided documentation describing “how” they performed their analyses. None of the investigators was informed of the study purpose until after data analysis was completed.

SReg. SReg is a method of regression analysis that identifies the intersection of two line segments formed when an abrupt change, or breakpoint, occurs in the dependent variable (y-axis) as a function of the independent variable (x-axis). The breakpoint is the intersection of the two lines with the smallest residual sums of squares. The “user-defined” equation in the nonlinear regression tab of GraphPad Prism (version 4.0) was used for this purpose, although the same is possible with other statistical software packages. The equation used to fit two intersecting lines is (25):

\[
\begin{align*}
\ y_1 &= a_1 + b_1 \times x_1 \\
\ y_2 &= a_2 + b_2 \times (x - x_0) \\
\ y &= IF \ (x < x_0, y_1, y_2)
\end{align*}
\]

where \( y_1 \) defines the intercept and slope of the first line segment, \( y \) at \( x_0 \) is the \( y \) value of the first line segment when \( x = x_0 \), and \( y_2 \) computes the second regression segment from the \( x_0 \) breakpoint. The final line defines \( y \) for all values of \( x \), where, if \( x < x_0 \), then \( y = y_1 \); otherwise, \( y = y_2 \) (25). Initial estimate values for \( a_1 \) (0.05), \( b_1 \) (0.00), \( x_0 \) (37.00), and \( b_2 \) (1.00) were used to provide a reasonable starting point for the actual mathematical solutions to \( x_0 \) (temperature threshold) and \( b_2 \) (sensitivity or slope). These inputs were held constant for all 16 data sets regressing \( \dot{m}_{sw}/T_{es} \).

Qualitative LOM. The importance of measurement differences among procedures depends, in part, on the magnitude of the expected change in the variable. Published reproducibility (biological noise) of the temperature threshold for sweating is 0.1°C when measured serially over a 20-min interval (4). This can be considered the smallest difference worth detecting. No reproducibility data were available for sweating sensitivity, defined by the slope of \( \dot{m}_{sw}/T_{es} \), due to the sweat gland priming artifact that occurs over such a short interval (4, 26). It is well documented (33) that a variety of factors modify the control of thermoregulatory sweating. The largest range of effects in threshold temperature (limit 1, 0.50°C; limit 2, −0.25°C) and sensitivity (slope) (limit 1, 0.15 mg·cm⁻²·min⁻¹·°C⁻¹; limit 2, −0.30 mg·cm⁻²·min⁻¹·°C⁻¹) were identified among the many studies referenced in the Introduction (1, 8, 10, 11, 17, 18, 19–21, 23, 28–31, 36, 39) and used as liberal LOM for quantitative data assessment. Both the overall differences between procedural methods (SReg vs. Rater) and differences among raters within a given data set (orthogonal contrasts) were interpreted in relation to the LOM. A within-subjects experiment with one volunteer was also evaluated similarly to directly examine how the variety of Rater techniques might affect interpretation of the same treatment effect.

Statistical analysis. The analytic performances of Rater and SReg were determined using sweating threshold temperature and sensitivity as outcome measures. From among the 16 data sets, the number of possible interrater differences within Rater was determined to be 21 in each set (7(7 – 1)/2 = 21) and 336 in all (21 × 16). Data were compiled using tabled, pairwise orthogonal contrasts. All 336 individual and 16 mean differences were plotted using a modified Bland–Altman approach (2). Briefly, the practical importance of the individual and mean differences within (Rater) were examined by plotting
and interpreting the percentage of differences within LOM for biological noise (4) and reported experimental effects on control of sweating (limits I and 2) (1, 8, 10, 11, 17–21, 23, 28–31, 34–36, 39). This approach is simple, easy to interpret (2), and allows data evaluation against an evidentiary standard other than zero, similar to equivalence testing (9). A liberal range was selected for the LOM, so that differences outside the LOM would be unequivocally important. Because interrater differences have not been previously evaluated, this approach enabled a descriptive assessment of Rater performance against which the merit of using SReg could be determined.

Quantitative analysis began with a simple Student t-test comparing the mean absolute Rater response for each of the 16 data sets against the associated SReg value (n = 16). A total of 112 possible individual differences were calculated between Rater and SReg (7 * 16 = 112). The individual differences (112 vs. 336) were likewise compared quantitatively using a Student t-test, without concern for unequal sample sizes or inequality of variances (40). Bland-Altman analysis was performed as described above for qualitative analysis of the SReg mean (n = 16) and individual (n = 112) differences. Last, the strength of the relationship between the mean SReg and Rater values was examined using Pearson product moment correlation. All data are presented as means ± SD, unless otherwise indicated. Statistical significance was accepted at P < 0.05. The qualitative importance of any difference was considered marginal, independent of P value, if it was smaller than the liberal LOM. Taken together, quantitative and qualitative comparisons were used to determine the validity of SReg with respect to Rater.

RESULTS

Descriptive. Baseline m˙sw was measured before the start of exercise. Under the variety of conditions described in Table 1, values ranged from 0.008 to 0.220 mg·cm⁻²·min⁻¹ (Fig. 1), which is consistent with resting transepidermal water losses across temperate to hot environments (5, 6, 22).

The variety of analytic approaches employed by the seven independent raters involved combinations of visual estimation, least squares regression, and empiricism. Rater 5 identified the threshold as the Tₚₑ when m˙sw exceeded 0.06 mg·cm⁻²·min⁻¹. Raters 4 and 7 identified threshold as the intercept of the regression equation of m˙sw/Tₚₑ after visually identifying an approximate breakpoint for slope calculation. This is also the approach of rater 5 when sweating baselines exceeded 0.06 mg·cm⁻²·min⁻¹. Rater 3 used the first data point to exceed >1 SD from the variation in successive data points on the baseline sweating segment of the zero slope response. Raters 1, 2, and 6 identified the threshold visually, using descriptive terms like “first obvious inflection point” when graphing m˙sw/Tₚₑ or m˙sw/time. Rater 1 identified 14 of 16 data sets (sets 1–13, 16) as having two distinct slopes beyond threshold. Rater 6 identified 10 of 16 data sets as having two distinct slopes (sets 1–6, 8, 13, 15, and 16). Rater 2 identified two distinct slopes for data sets 15 and 16 only. In each case of biphasic slope identification, only the early-phase sweating transient was analyzed. Rater 7 felt that data sets 14 and 15 could not be properly analyzed and left these untouched, so those data sets are not reported for rater 7. Raters 3, 4, and 5 made no distinctions between early- or late-phase sweating transients for any data set.

In all, there were 110 individual Rater sweating threshold temperature and sweating sensitivity calculations and 324 interrater contrasts for analysis. Of the sensitivity calculations, 84 were calculated without, and 26 with, distinction for early- and late-phase slopes. Of the 16 possible Rater averages for each data set, 14 were calculated using n = 7 raters and 2 (data sets 14 and 15) were calculated from n = 6 raters.

Analytic. Figure 2, A (threshold) and B (sensitivity), shows the mean (triangles) and individual (circles) interrater differences for each of the 16 data sets using modified Bland-Altman plots. The shaded band represents variability from biological noise. For mean values, 100% of threshold and sensitivity differences fell within the LOM. For individual values, 307/324 (95%) threshold differences and 235/324 (73%) sensitivity differences fell within limits I and 2, respectively. Any pattern for the threshold values out of agreement was not obvious, but approximately one-half of values out of agreement for sensitivity differences were due to raters 1 and 6. The data quality of Rater was, therefore, established for comparison to SReg.

Table 2 presents descriptive and inferential analytic data comparing mean threshold and sensitivity values for the SReg and Rater methods. In only one case for threshold (data set 5) and in no case for sensitivity was SReg outside the range of values reported for Rater. No significant differences were
observed between SReg and the mean Rater methods for either threshold or sensitivity calculations. The 110 individual Rater vs. SReg differences were also not different \( (P > 0.05) \) compared with the 324 interrater differences within Rater for either threshold or slope. The correlation coefficients between SReg and mean Rater (Table 2 values) indicated a strong, positive relationship for threshold \( (r = 0.97) \) and sensitivity \( (0.99) \).

Figure 3, A (threshold) and B (sensitivity), shows the mean (triangles) and individual (circles) differences between SReg and Rater values using modified Bland-Altman plots. For mean threshold values, 8 of 16 (50%) mean differences fell within the shaded band of biological noise, whereas all 16 (100%) fell within limits 1 and 2. For individual threshold values, 99/110 (90%) differences fell within limits 1 and 2. For sensitivity calculations, all 16 (100%) mean differences and 91/110 (83%) individual differences fell within limits 1 and 2, respectively. Thus the average differences between SReg and Rater were smaller than the largest experimental effects reported in the literature for sweating threshold temperature and sensitivity. In general, SReg was in agreement with individual Rater calculations between 80 and 90% of the time. Seven of the eleven threshold values out of agreement were for data sets 1 and 5. Fourteen of the nineteen sensitivity calculations out of agreement were due to raters 1 and 6, both of whom identified two distinct slopes for the majority of data sets.

To illustrate the potential for Rater techniques to influence interpretation of the same treatment effects, a within-subjects design was performed. Table 3 shows the changes in threshold and \( m_{\text{sw}} \) sensitivity for a single volunteer tested once when euhydrated (Panel 5) and once when dehydrated (Panel 6) by 3% of body mass. Exercise intensity and environmental conditions were the same in both trials (Table 1). The threshold temperature obtained using SReg was slightly higher than the range reported using Rater, while sensitivity was within the Rater range. However, the Rater range for the observed change in \( T_{\text{es}} \) threshold (0.12–0.41) and \( m_{\text{sw}} \) sensitivity (0.07–0.26) was large and underscores the potential for very different interpretations of the same experimental effect when applying a variety of Rater techniques.

**DISCUSSION**

Sweating threshold temperature and sweating sensitivity values are frequently reported to evaluate changes in thermoregulatory control; however, no standardized methods have been validated. This was the first study to characterize the measurement variability for numerous empirical approaches employed by experienced investigators and to assess the magnitude of interrater differences. It was then possible to ascertain the validity of a simple, standardized methodology for determining sweating threshold temperature and sweating sensitivity values by comparing SReg to the variety of Rater approaches. Data selected for this study included diverse conditions (Table 1), which produced a heterogeneous sample of sweating responses for analysis (Fig. 1).

The application of SReg consistently produced sweating threshold and sensitivity values within the range of Rater estimates (Table 2). The differences between SReg and mean Rater estimates were always well within qualitative LOM, while the differences between SReg and individual rater estimates \( (n = 110) \) were also smaller than the same limits >80% of the time (Fig. 3). Despite differences in methods used by our raters, the average differences among raters with respect to SReg \( (n = 16) \) were well within qualitative LOM. Interrater differences \( (n = 324) \) were also within these limits for 95% (threshold) and 73% (sensitivity) of the comparisons. It is important to view these findings within the context of the desired experimental signal-to-noise ratio (effect size). The qualitative LOM, which define performance acceptability herein, are the largest range of treatment effects reported in the literature. Much smaller changes (signal) in the sweating threshold temperature would be expected, for example, when studying a short circadian window or a modest level of dehydration (23, 36). Conversely, more noise in the denominator might also be characteristic of certain situations or populations.

---

**Table 2. Comparison of \( T_{\text{es}} \) threshold and sweating sensitivity calculations using SReg vs. Rater methods**

<table>
<thead>
<tr>
<th>Panel No.</th>
<th>( T_{\text{es}} ) Threshold, °C</th>
<th>( m_{\text{sw}} ) Sensitivity, mg·cm(^{-2} )·min(^{-1} )·°C(^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SReg</td>
<td>Rater</td>
</tr>
<tr>
<td>1</td>
<td>37.05</td>
<td>37.19 (37.01–37.31)</td>
</tr>
<tr>
<td>2</td>
<td>37.13</td>
<td>37.23 (37.08–37.40)</td>
</tr>
<tr>
<td>3</td>
<td>37.62</td>
<td>37.61 (37.50–37.72)</td>
</tr>
<tr>
<td>4</td>
<td>36.90</td>
<td>36.86 (36.79–36.92)</td>
</tr>
<tr>
<td>5</td>
<td>36.35</td>
<td>36.55 (36.44–36.63)</td>
</tr>
<tr>
<td>6</td>
<td>36.78</td>
<td>36.87 (36.71–37.02)</td>
</tr>
<tr>
<td>7</td>
<td>37.07</td>
<td>37.03 (36.84–37.14)</td>
</tr>
<tr>
<td>8</td>
<td>36.83</td>
<td>36.93 (36.74–36.92)</td>
</tr>
<tr>
<td>9</td>
<td>36.88</td>
<td>36.72 (36.56–37.07)</td>
</tr>
<tr>
<td>10</td>
<td>36.78</td>
<td>36.91 (36.55–37.08)</td>
</tr>
<tr>
<td>11</td>
<td>37.24</td>
<td>37.31 (37.15–37.40)</td>
</tr>
<tr>
<td>12</td>
<td>36.74</td>
<td>36.82 (36.60–36.90)</td>
</tr>
<tr>
<td>13</td>
<td>36.90</td>
<td>37.00 (36.69–37.09)</td>
</tr>
<tr>
<td>14</td>
<td>37.41</td>
<td>37.43 (37.35–37.46)</td>
</tr>
<tr>
<td>15</td>
<td>37.01</td>
<td>37.00 (36.77–37.19)</td>
</tr>
<tr>
<td>16</td>
<td>36.87</td>
<td>37.01 (36.70–37.17)</td>
</tr>
<tr>
<td>Mean</td>
<td>36.97</td>
<td>37.04</td>
</tr>
<tr>
<td>SD</td>
<td>0.30</td>
<td>0.26</td>
</tr>
</tbody>
</table>

\( T_{\text{es}} \), esophageal temperature; \( m_{\text{sw}} \), sweating rate; SReg, segmented regression; Rater, analysis by seven experienced raters. Values in parentheses represent Rater range.
similar to piecewise linear regression (37), to objectively compare early and late sweating phases, if desired. This technique can also be used to objectively remove one phase or the other entirely from the analysis. Our a priori decision was to make no distinction between phases, because this seemed the most common approach in the literature. The uncoached choices made by our raters also support this decision.

Although the interrater differences for sweating threshold temperature and sweating sensitivity values were quantitatively and qualitatively acceptable, the potential exists for any two laboratories using Rater to arrive at very different conclusions about the same data set (Fig. 2 and Table 3). In Table 3, for example, rater 5 would have reported that dehydration produced a small change in the $T_{es}$ threshold (0.12°C), which is near the level of unimportant biological noise (0.10°C) (4). In contrast, rater 6 would have reported a change nearly four times larger (0.41°C) and more consistent with the literature. There was, however, very close agreement among three raters (0.39–0.41°C) and SReg (0.43°C). Similar findings were apparent for sensitivity calculations (Table 3). Greater methodological details and more precise identification of analysis parameters within manuscripts should reduce interrater measurement variability. While the degree to which SReg or any other method can complement intuition and improve computations is ultimately determined by the care of the end user, evidence for adopting a standardized method of doing sweating threshold temperature and sensitivity analyses is now available. A standardized procedure would logically reduce interrater measurement noise and focus interpretation of sweating control data around common factors of influence (33). A prospective study comparing the use of SReg by different investigators analyzing the same data would provide definitive supporting evidence.

Conclusions. In conclusion, SReg offers a simple, standardized, and valid method of determining sweating threshold temperatures and sweating sensitivity values. Compared with values obtained from a range of empirical approaches used by experienced investigators, the majority of differences outside the LOM for both threshold temperature and sweating sensitivity were explainable by differences in the perception and analysis of distinct sweating phases. It is, therefore, recommended that the phase(s) of the sweating response targeted for analysis be clearly identified to further improve SReg for sweating threshold temperature and sweating sensitivity anal-

![Fig. 3. Bland-Altman plots of the differences between segmented regression (SReg) and Rater determinations of sweating threshold (A) and sensitivity (B). Circles represent individual differences ($n = 110$), and triangles represent mean differences ($n = 16$). The shaded band and LOM are the same as described in Fig. 2 legend.](image)

Table 3. Within-subject changes in $T_{es}$ threshold and $m_{sw}$ sensitivity between euhydration (panel 5) and dehydration (panel 6) conditions using Rater or SReg methods

<table>
<thead>
<tr>
<th>Rater No.</th>
<th>$\Delta T_{es}$ Threshold, °C</th>
<th>$\Delta m_{sw}$ Sensitivity, mg cm$^{-2}$ min$^{-1}$ °C$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.39</td>
<td>−0.60</td>
</tr>
<tr>
<td>2</td>
<td>0.39</td>
<td>−0.07</td>
</tr>
<tr>
<td>3</td>
<td>0.26</td>
<td>−0.14</td>
</tr>
<tr>
<td>4</td>
<td>0.23</td>
<td>−0.12</td>
</tr>
<tr>
<td>5</td>
<td>0.12</td>
<td>−0.14</td>
</tr>
<tr>
<td>6</td>
<td>0.41</td>
<td>−0.26</td>
</tr>
<tr>
<td>7</td>
<td>0.40</td>
<td>−0.15</td>
</tr>
<tr>
<td>Mean</td>
<td>0.31</td>
<td>−0.21</td>
</tr>
<tr>
<td>SD</td>
<td>0.11</td>
<td>−0.18</td>
</tr>
<tr>
<td>SReg</td>
<td>0.43</td>
<td>−0.14</td>
</tr>
</tbody>
</table>

$\Delta$, Change.
yses. The application of these findings will improve comparisons and communications within and between independent laboratories vested in the study of factors that influence control of thermoregulatory sweating.

ACKNOWLEDGMENTS

The authors thank Dr. Christopher H. Schmid, Director of the Tufts University Biostatistics Research Center, and Dr. Warren W. Tryon (Fordham University) for expert statistical assistance. We also thank Drs. Lacy Alexander Holowatz, Nisha Charkoudian, and David A. Low.

DISCLAIMER

The opinions or assertions contained herein are the private views of the authors and should not be construed as official or reflecting the views of the Army of the Department of Defense. Approved for public release: distribution unlimited.

REFERENCES