



OPEN

## Elevated body temperature is associated with depressive symptoms: results from the TemPredict Study

Ashley E. Mason<sup>1✉</sup>, Patrick Kasl<sup>2</sup>, Severine Soltani<sup>2</sup>, Abigail Green<sup>3</sup>, Wendy Hartogensis<sup>1</sup>, Stephan Dilchert<sup>4</sup>, Anoushka Chowdhary<sup>5</sup>, Leena S. Pandya<sup>1</sup>, Chelsea J. Siwik<sup>6</sup>, Simmie L. Foster<sup>7,8</sup>, Maren Nyer<sup>7,8</sup>, Christopher A. Lowry<sup>9</sup>, Charles L. Raison<sup>10</sup>, Frederick M. Hecht<sup>1,12</sup> & Benjamin L. Smarr<sup>2,11,12</sup>

Correlations between altered body temperature and depression have been reported in small samples; greater confidence in these associations would provide a rationale for further examining potential mechanisms of depression related to body temperature regulation. We sought to test the hypotheses that greater depression symptom severity is associated with (1) higher body temperature, (2) smaller differences between body temperature when awake versus asleep, and (3) lower diurnal body temperature amplitude. Data collected included both self-reported body temperature (using standard thermometers), wearable sensor-assessed distal body temperature (using an off-the-shelf wearable sensor that collected minute-level physiological data), and self-reported depressive symptoms from > 20,000 participants over the course of ~ 7 months as part of the TemPredict Study. Higher self-reported and wearable sensor-assessed body temperatures when awake were associated with greater depression symptom severity. Lower diurnal body temperature amplitude, computed using wearable sensor-assessed distal body temperature data, tended to be associated with greater depression symptom severity, though this association did not achieve statistical significance. These findings, drawn from a large sample, replicate and expand upon prior data pointing to body temperature alterations as potentially relevant factors in depression etiology and may hold implications for development of novel approaches to the treatment of major depressive disorder.

Depression has become a health crisis of epidemic proportions<sup>1</sup>. Globally, the prevalence of major depressive disorder (MDD) has risen over the last several generations in countries across the world<sup>2</sup>. The last decade has seen a particularly significant increase in depression in the United States, with prevalence rates increasing by 33% between 2013 and 2016, with the largest increase among youth and young adults<sup>3,4</sup>. This is particularly concerning as the disease course is most likely to be malignant, and the costs of depression in terms of lost opportunities across a lifetime are likely to be highest in youth and young adulthood<sup>4</sup>. Rates of antidepressant use have increased substantially in most Western countries during the same time period<sup>5</sup>, and currently available pharmacologic agents have significant limitations in efficacy<sup>6–8</sup>. Taken together, these patterns highlight the urgent need to identify and implement new treatments for depression.

<sup>1</sup>Osher Center for Integrative Health, University of California San Francisco, San Francisco, CA, USA. <sup>2</sup>Shu Chien-Gene Lay Department of Bioengineering, University of California San Diego, San Diego, CA, USA. <sup>3</sup>Neurosciences Graduate Program, University of California San Diego, San Diego, CA, USA. <sup>4</sup>Department of Management, Zicklin School of Business, Baruch College, The City University of New York, New York, NY, USA. <sup>5</sup>Department of Psychology, The University of Arizona, Tucson, AZ, USA. <sup>6</sup>Department of Wellness and Preventative Medicine, Cleveland Clinic, Cleveland, OH, USA. <sup>7</sup>Depression Clinical and Research Program, Massachusetts General Hospital, Boston, MA, USA. <sup>8</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA. <sup>9</sup>Department of Integrative Physiology, University of Colorado Boulder, Boulder, CO, USA. <sup>10</sup>Department of Psychiatry, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA. <sup>11</sup>Halicioğlu Data Science Institute, University of California San Diego, San Diego, CA, USA. <sup>12</sup>These authors contributed equally: Frederick M. Hecht and Benjamin L. Smarr. ✉email: ashley.mason@ucsf.edu

To develop novel treatments for depression it is important to identify mechanisms that contribute to the development and/or maintenance of depressive symptoms and that may be amenable to intervention. Although depression is both biologically and behaviorally heterogeneous<sup>9,10</sup>, an important first step in treatment development is often to identify physiologic signatures among individuals with MDD that are not present among those without MDD. Although no single biological or behavioral abnormality will characterize all individuals with MDD, the identification of an abnormality associated with MDD may open the door to identifying a relatively biologically homogeneous subgroup that demonstrates a larger treatment response to interventions that target the specific abnormality<sup>11</sup>.

One physiologic characteristic that may hold potential as a therapeutic target is thermoregulatory dysregulation<sup>12</sup>, which is among the most widely reported circadian biological abnormalities in affective disorders, including MDD. This has been observed in the form of elevated body temperature, particularly at night<sup>13–16</sup>, when thermoregulatory cooling responses are critical for sleep onset and quality<sup>17,18</sup>. Such temperature elevations have also been reported during the day<sup>19,20</sup>. Notably, some data suggest that these aberrant temperature elevations improve upon clinical recovery in MDD<sup>21,22</sup>. Data have shown that circadian amplitude is blunted in depression<sup>23</sup> and that smaller differences between average body temperature during time awake (typically during the daytime) and during time asleep (typically during the nighttime) are associated with greater depressive symptoms<sup>21</sup>. Additionally, data have shown that individuals with depression have lower circadian body temperature amplitudes that increase upon clinical recovery<sup>13,21</sup>. These alterations are notable in light of the narrow range of human body temperature<sup>24,25</sup>.

Although suggestive, these results derive from small-scale studies conducted with limited sample sizes (< 300) and in controlled settings<sup>13–15,19,20</sup>. Thus, an important step toward understanding the association between thermoregulatory dysfunction and depressive symptoms is to establish this association outside of clinical laboratory settings and in larger samples. Establishing a more definitive association between measures of body temperature and depression would set the stage for further explorations of depression treatment modalities that target the thermoregulatory system. To accomplish this, the current analyses tested associations between monthly depression symptom scores and (1) self-collected body temperature submitted to the research team via online surveys and (2) wearable sensor device-collected distal body temperature automatically transmitted to the research team over the internet, in a large international sample ( $N > 20,000$ ) of adults who participated in an online study focused on COVID-19 detection beginning in March of 2020. Initial analyses used self-reported body temperature, which participants collected using personal thermometers (e.g., oral temperature) and reported via daily surveys. Further analyses used minute-level distal body temperature data collected using an off-the-shelf wearable sensor device. In total, using wearable sensor-assessed distal body temperature, we analyzed four body temperature metrics examined in prior research<sup>13–16,19–21</sup>: (1) distal body temperature while awake; (2) distal body temperature while asleep; (3) the difference between the average distal body temperatures while asleep and awake; and (4) the diurnal distal body temperature amplitude. Based on prior studies with small samples and our prior theoretical work<sup>12–16,19–21</sup>, we hypothesized that higher self-reported body temperature and wearable sensor-assessed distal body temperatures, lower diurnal distal body temperature amplitude, and smaller differences between awake and asleep distal body temperature, would be associated with higher levels of depressive symptoms.

## Results

### Study overview

The TemPredict Study, initiated in March of 2020, sought to assess whether off-the-shelf wearable sensor devices collect data that could be used to screen large numbers of individuals for the early stages of SARS-CoV-2 infection<sup>26,27</sup>. The primary findings included that physiological data allowed for prediction of a COVID-19 infection 2.75 days prior to diagnosis. All participants wore a commercially available off-the-shelf wearable sensor device, the Oura Ring (Oura Health, Oulu, Finland) that collected distal body temperature (one value per minute while worn on the finger) and paired with a smartphone app. Participants also completed daily surveys that asked for a self-collected body temperature measurement (assessed with a handheld thermometer; not assessed using the wearable device), a baseline survey that collected demographic data, and monthly surveys that included mental health assessments.

The self-reported body temperature analytic sample included 20,880 individuals (Supplementary Fig. S1). The mean age (standard deviation [SD]) was 46.9 (12.6) years; 53% male; 47% female. For adjusted analyses, we excluded participants who reported their biological sex as ‘other’ and participants whose biological sex was missing ( $n = 17$ ). We first computed each participant’s average T-score across a maximum of seven monthly PROMIS depression assessments. We then computed the average of all participants’ mean T-scores (SD); this value was 51.49 (7.37), which is within normal limits (WNL). On average, participants completed 3.6 of 7 possible PROMIS depression assessments and self-reported a total of 559,664 body temperature assessments (mean of 27 daily temperature reports per participant).

The wearable sensor-assessed body temperature analytic sample included 21,064 individuals (Supplementary Fig. S2). The mean age (SD) was 46.5 (12.1) years; 56% male; 44% female. The average T-score across a maximum of seven monthly PROMIS depression assessments in this sample was 50.94 (7.26), which is also WNL. On average, participants completed 4.0 of 7 possible PROMIS depression assessments; we used 2 weeks of wearable sensor device distal body temperature data prior to each completed PROMIS assessment for analyses. Participants had an average of 35.5 days of distal body temperature data with at least (1) 4 h of distal body temperature data during the asleep period and (2) at least 4 h of distal body temperature data during the awake period for each 24-h period available within our timeframe of analyses. Both analytic samples (which we drew from the same participant pool) included a geographically diverse set of individuals from 106 different countries.

## Self-reported body temperature

### Depression symptom frequencies

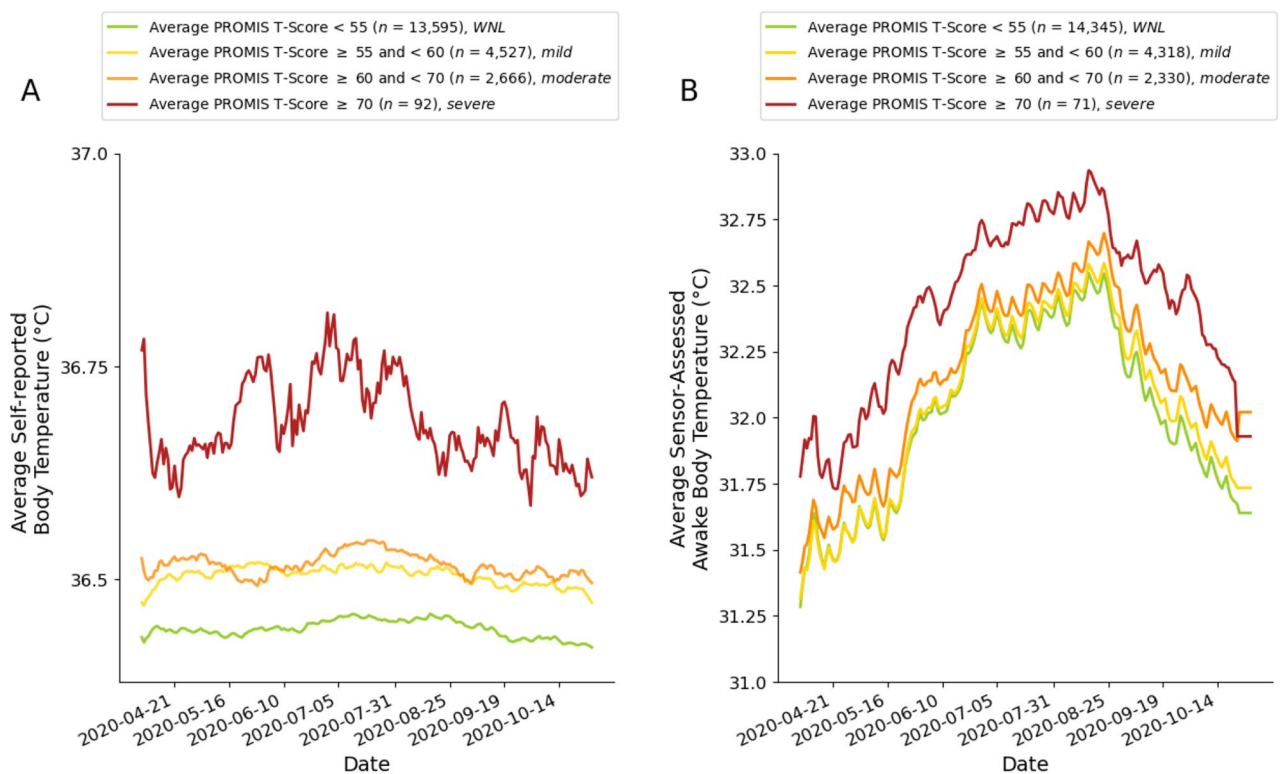
Of the 20,880 participants with self-reported temperature data, 13,595 had an average depression symptom severity WNL (T-score < 55), 4527 had an average depression symptom severity within the mild range (T-score  $\geq$  55 and < 60), 2666 had an average depression symptom severity within the moderate range (T-score  $\geq$  60 and < 70), and 92 had an average depression symptom severity within the severe range (T-score  $\geq$  70) of the PROMIS depression assessment<sup>28</sup>. The average depression symptom severity score across participants' average scores was 51.5 (the average within-person SD = 2.5), and the average range between participants' lowest to highest scores was 5.2 (SD = 5.5), and the median was 3.9 (interquartile range [IQR] = 0–8.3).

### Graphical depictions of self-report temperature data

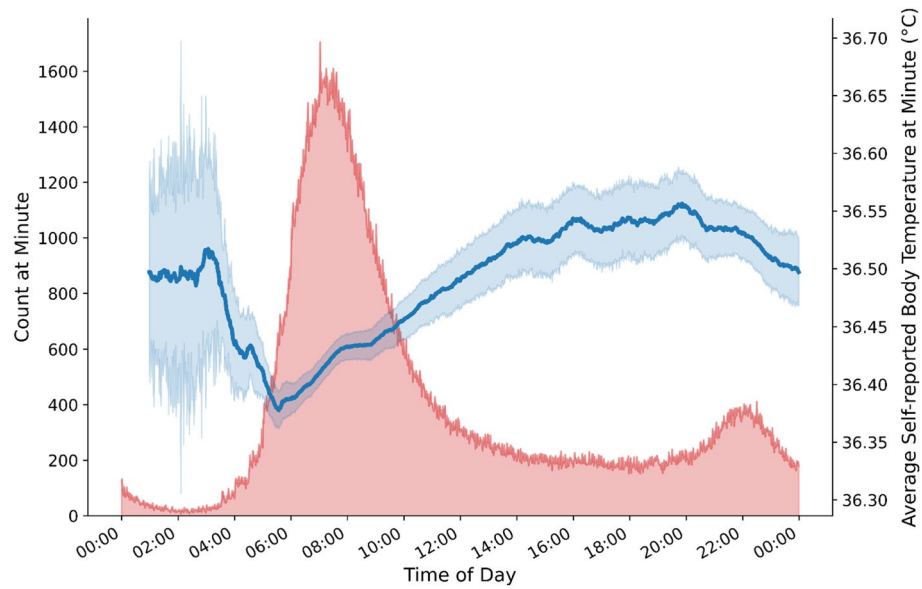
Figure 1 illustrates the differences in average self-reported body temperature across participants with average PROMIS depression symptom T-scores WNL, and in the mild, moderate, and severe ranges. The self-reported, time-stamped body temperature survey data (adjusted for local time zone) revealed that participants' self-reported temperatures followed expected diurnal patterns (Fig. 2), with lower body temperatures in the early morning and rising body temperatures during daytime hours that fell in the evening hours. This is consistent with the interpretation that participants self-collected their body temperature around the same time that they reported their body temperature to the study. As shown, participants tended to self-report their body temperature more often in the morning than in the evening.

### Linear regression models

We used linear regression models to assess whether the average daily self-reported body temperature recorded over seven months was correlated with the average PROMIS depression T-score across seven monthly assessments



**Figure 1.** Average self-reported body temperature (A) and average wearable sensor-assessed distal body temperature (B) plotted by PROMIS depression symptom T-score categories. Figure panels show that individuals with PROMIS depression symptom T-scores within normal limits (WNL; green) have the lowest average self-reported and wearable sensor-assessed distal body temperatures, with increasing average self-reported and wearable sensor-assessed distal body temperatures among individuals in the mild (yellow), moderate (orange), and severe (red) PROMIS depression symptom T-score categories. *Note.* We took the self-reported body temperature data from each calendar day (Panel A; most frequently reported by participants in the morning) and the average of any available wearable sensor-assessed distal body temperature data from each calendar day (during awake time; Panel B) and plotted those average values for the different depression groups. We smoothed average self-reported and wearable sensor-assessed body temperatures (°C) using an exponentially weighted moving average with a 7-day window before and after each timepoint as a function of the average NIH PROMIS Adult Health Profile instrument for depression (Form 4a) score. See Supplementary Fig. S3 for unsmoothed figure panels.



**Figure 2.** Average self-reported body temperature by time-of-day. Figure depicts expected diurnal pattern of lowest self-reported body temperatures reported in the early morning hours and higher self-reported body temperatures during daytime hours. *Note.* Blue line depicts average self-reported body temperature (right Y axis) by time of day; blue shading indicates standard error of the mean. Red shading indicates number of responses (left Y axis) provided at each minute (X axis).

from the same period. We found that these were positively correlated (greater body temperature associated with greater depression T-score) in an unadjusted model ( $b = 1.6661$ ; 95% CI [1.4128, 1.9195];  $p = 7.1 \times 10^{-38}$ ) and in a model adjusted for age, biological sex, and self-reported body temperature survey time stamp ( $b = 0.8595$ ; 95% CI [0.6118, 1.1071];  $p = 1.1 \times 10^{-11}$ ), Table 1). We calculated E-values as sensitivity analyses. E-values for these linear models ranged from 1.38 to 1.60, and in all cases were larger than observed effects of body temperature, age, or biological sex on depression (Table 1), meaning that it is unlikely that unaddressed or unmeasured confounders had effects that better explained the reported (observed) effects. Standardized adjusted regression analyses showed that body temperature accounted for unique variance in PROMIS depression T-scores beyond the known variance accounted for by biological sex and age (Supplementary Table S1).

Model	Model parameter	<i>b</i>	SE	<i>p</i>	95% CI (LB, UB)	E-value for point estimate (E-value for 95% CI)
Unadjusted ( <i>n</i> = 20,880) $r^2 = 0.008$ $p = 7.1 \times 10^{-38}$	Intercept	51.488	0.051		[51.388, 51.588]	1.60 (1.53)
	Self-report body temperature	1.666	0.129	$7.1 \times 10^{-38}$	[1.413, 1.920]	
Adjusted ( <i>n</i> = 20,863) $r^2 = 0.078$ $p = 2.3 \times 10^{-128}$	Intercept	52.940	0.0764		[52.790, 53.089]	1.38 (1.30)
	Self-report body temperature	0.859	0.126	$1.1 \times 10^{-11}$	[0.612, 1.107]	
	Age	- 0.103	0.004	$9.3 \times 10^{-146}$	[- 0.111, - 0.096]	
	Biological sex	- 2.741	0.099	$1.8 \times 10^{-165}$	[- 2.935, - 2.547]	
	Time of day (B1)	- 0.526	0.117	$7.5 \times 10^{-6}$	[- 0.756, - 0.296]	
	Time of day (B2)	- 0.774	0.095	$4.6 \times 10^{-16}$	[- 0.961, - 0.588]	

**Table 1.** Unadjusted and adjusted linear models regressing average PROMIS depression symptom T-scores onto average self-reported body temperature. *Note.* Models predict depression symptom T-scores assessed using the Patient-Reported Outcomes Measurement Information System (PROMIS), Adult Health instrument for depression (Form 4a), modified to reflect a 1-month timeframe<sup>63</sup>. We used survey timestamps to compute the cosine (Time of Day B1) and sine (Time of Day B2) of  $2\pi \times t$  for each self-reported body temperature measurement, where “t” was the decimal proportion of the day during which the daily survey was completed (see “Methods”). CI, confidence interval; LB, lower bound; UB, upper bound. Mean self-reported body temperatures were centered around the grand mean of 36.54, such that the intercept represents the mean PROMIS depression T-score for participants with average self-reported body temperatures at the analytic sample mean. Age was centered around the overall analytic sample mean of 46.94.

### Logistic regression models

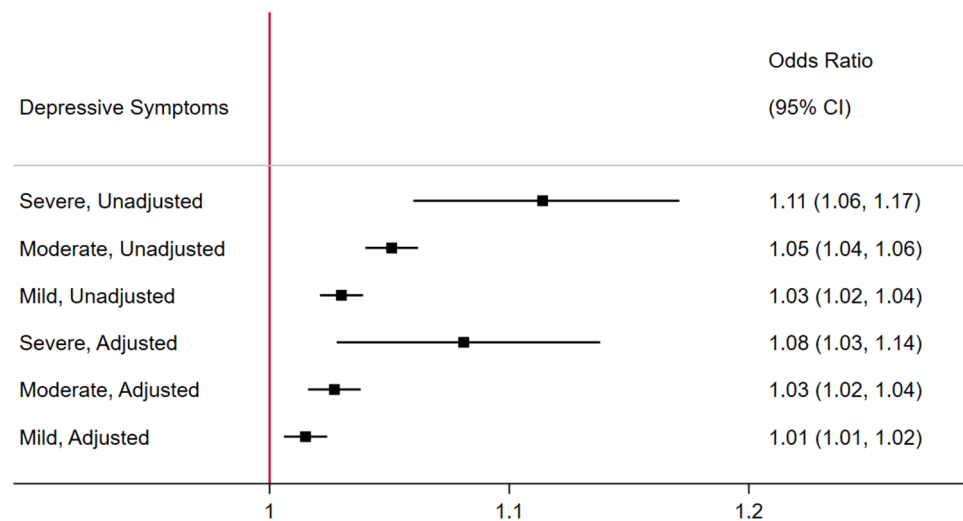
We next tested whether greater self-reported body temperature was associated with increased odds of having depression using separate logistic regression models for each category of depression (mild, moderate, and severe). In unadjusted analyses (Table 2), the odds ratio for having average PROMIS depression T-scores within the mild range vs. WNL were significantly increased with each 0.1 °C increase in average body temperature (OR: 1.030; 95% CI [1.021, 1.039];  $p = 1.4 \times 10^{-11}$ ). Similarly, there were increased (but progressively greater) odds of having (with each 0.1 °C increase in body temperature) PROMIS depression T-scores within the moderate range (OR = 1.051 [95% CI, 1.040, 1.062],  $p = 2.3 \times 10^{-20}$ ) and of having PROMIS depression T-scores within the severe range (OR: 1.114; 95% CI [1.060, 1.171],  $p = 2.1 \times 10^{-5}$ ) relative to PROMIS depression T-scores WNL (Fig. 3). Similar regression analyses adjusting for covariates attenuated these differences somewhat (mild vs. WNL; OR: 1.015; 95% CI [1.006, 1.024],  $p = 8.7 \times 10^{-4}$ ; moderate vs. WNL OR: 1.027 95% CI [1.016, 1.038],  $p = 1.8 \times 10^{-6}$ ; severe vs. WNL: OR = 1.081; 95% CI [1.028, 1.138],  $p = 0.0026$ ; Table 3). E-values for these logistic models ranged from 1.09 to 1.30, and in all cases were larger than observed effects of self-reported body temperature, age, or biological sex, on depression (Tables 2 and 3).

### Receiver operating characteristics (ROC) curve analyses

ROC curve analyses for each logistic regression model (Fig. 4) showed better discernment based on the adjusted rather than unadjusted models comparing PROMIS depression T-scores between the severe and WNL ranges (AUC = 0.762 vs. unadjusted AUC = 0.635), between the moderate and WNL ranges (AUC = 0.672 vs. unadjusted AUC = 0.557), and between the mild and WNL ranges (AUC = 0.612 vs. unadjusted AUC = 0.537). Using Youden's Index, which locates the threshold value that maximizes the distance between the ROC curve and the line of chance, to identify optimally performing threshold values from each ROC curve resulted in 85.87% sensitivity

Model	Model Parameter	OR	SE	$p$	95% CI [LB, UB]	E-Value of point estimate (E-Value of 95% CI)
Severe Depressive Symptoms vs. WNL 13,687 observations Pseudo $r^2 = 0.0157$	Intercept	0.006	0.001		[0.005, 0.008]	1.30 (1.20)
	Self-report body temperature	1.114	0.028	$2.1 \times 10^{-5}$	[1.060, 1.171]	
Moderate Depressive Symptoms vs. WNL 16,261 observations Pseudo $r^2 = 0.0059$	Intercept	0.194	0.004		[0.186, 0.203]	1.19 (1.16)
	Self-report body temperature	1.051	0.006	$2.3 \times 10^{-20}$	[1.040, 1.062]	
Mild Depressive Symptoms vs. WNL 18,122 observations Pseudo $r^2 = 0.0022$	Intercept	0.333	0.006		[0.322, 0.344]	1.14 (1.11)
	Self-report body temperature	1.030	0.005	$1.4 \times 10^{-11}$	[1.021, 1.039]	

**Table 2.** Unadjusted logistic models regressing PROMIS depression symptom T-score categories (mild, moderate, severe, with T-scores within normal limits [WNL] as the reference category) onto average self-reported body temperature (scaled per 0.1 °C). *Note.* CI, confidence interval; LB, lower bound; UB, upper bound.



**Figure 3.** Forest plot depicting odds ratios from logistic regression models. Adjusted and unadjusted models predict PROMIS depression T-score categories (mild, moderate, severe) vs. depressive symptoms within normal limits (WNL) from self-reported body temperature (scaled per 0.1 °C).

Model	Model parameter	OR	SE	<i>p</i>	95% CI [LB, UB]	E-Value of point estimate (E-Value of 95% CI)
Severe Depressive Symptoms vs. WNL 13,676 observations Pseudo $r^2 = 0.0742$	Intercept	0.007	0.001		[0.005, 0.010]	1.24 (1.13)
	Self-report body temperature	1.081	0.028	0.0026	[1.028, 1.138]	
	Age	0.941	0.009	$8.1 \times 10^{-10}$	[0.922, 0.959]	
	Biological sex	0.580	0.124	0.011	[0.381, 0.883]	
	Time of day (B1)	0.894	0.208	0.63	[0.567, 1.410]	
	Time of day (B2)	0.557	0.113	0.0039	[0.375, 0.829]	
Moderate Depressive Symptoms vs. WNL 16,246 observations Pseudo $r^2 = 0.0582$	Intercept	0.261	0.008		[0.245, 0.278]	1.13 (1.10)
	Self-report body temperature	1.027	0.006	$1.8 \times 10^{-06}$	[1.016, 1.038]	
	Age	0.965	0.002	$3.2 \times 10^{-81}$	[0.961, 0.968]	
	Biological sex	0.477	0.021	$3.2 \times 10^{-62}$	[0.437, 0.520]	
	Time of day (B1)	0.786	0.040	$2.7 \times 10^{-6}$	[0.710, 0.869]	
	Time of day (B2)	0.762	0.032	$1.1 \times 10^{-10}$	[0.701, 0.827]	
Mild Depressive Symptoms vs. WNL 18,109 observations Pseudo $r^2 = 0.0257$	Intercept	0.431	0.011		[0.410, 0.454]	1.09 (1.06)
	Self-report body temperature	1.015	0.005	$8.7 \times 10^{-4}$	[1.006, 1.024]	
	Age	0.979	0.001	$2.9 \times 10^{-47}$	[0.976, 0.982]	
	Biological sex	0.591	0.021	$1.6 \times 10^{-50}$	[0.552, 0.633]	
	Time of day (B1)	0.869	0.036	$7.0 \times 10^{-4}$	[0.801, 0.942]	
	Time of day (B2)	0.855	0.029	$3.2 \times 10^{-6}$	[0.800, 0.913]	

**Table 3.** Adjusted logistic models regressing PROMIS depression symptom T-score categories (mild, moderate, severe, with T-scores within normal limits [WNL] as the reference category) onto average self-reported body temperature (scaled per 0.1 °C). *Note.* See Table 2 note.

to detect PROMIS depression T-scores within the severe range based on the adjusted model, but with specificity of 34.05%; the best performance was 96.91% sensitivity to detect PROMIS depression T-scores within the moderate range (with 63.40% specificity) from the unadjusted model. Sensitivity was lowest (42.72%) for detection of PROMIS depression T-scores within the mild range based on an unadjusted analysis. Specificity was lowest (31.55%) for detection of PROMIS depression T-scores within the moderate range based on the adjusted model (Table 4).

### Wearable sensor-assessed temperature

#### *Depression symptom frequencies*

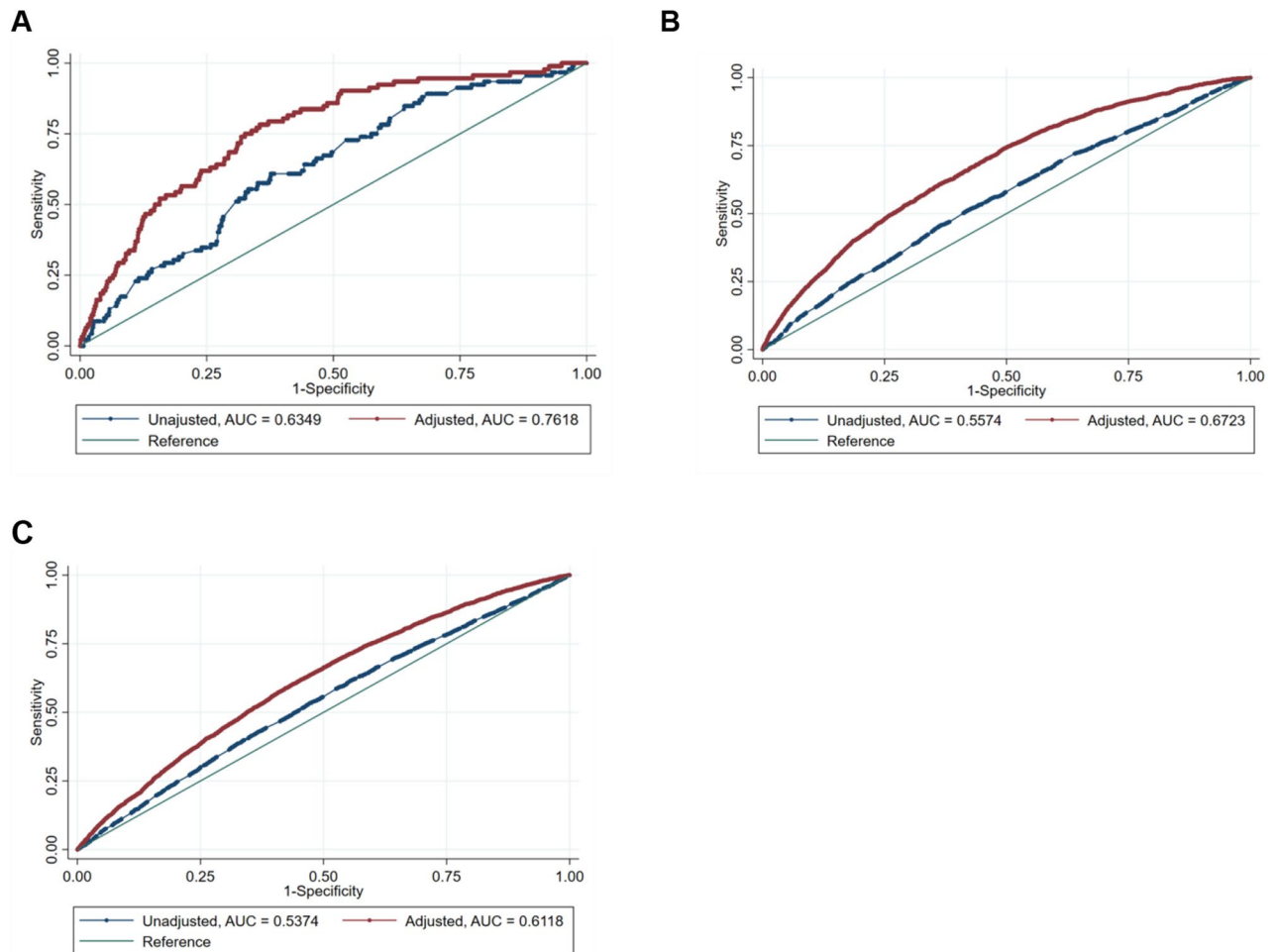
Of the 21,064 participants with wearable sensor-assessed distal body temperature data, 14,345 had an average depression symptom severity WNL, 4318 had an average depression symptom severity within the mild range, 2330 had an average depression symptom severity within the moderate range, and 71 had an average depression symptom severity within the severe range of the PROMIS depression assessment<sup>28</sup>. The average depression symptom severity score across participants' mean scores was 50.9 (SD = 2.2), and the mean range between participants' lowest to highest scores was 5.5 (SD = 5.6), and the median was 4.9 (IQR = 0–8.7).

#### *Distributions of wearable sensor-assessed body temperature metrics by depression symptom severity*

Probability density plots illustrate differences in the distribution of 4 wearable sensor-assessed distal body temperature metrics by severity of depressive symptoms (Fig. 5A–D). The awake distal body temperature distributions shift slightly higher from WNL to mild, and from WNL to moderate, with the most pronounced shift from WNL to severe depressive symptoms (see Kolmogorov–Smirnov D-statistics, below). In contrast, the asleep–awake distal body temperature difference and the diurnal distal body temperature amplitude plots demonstrate similar distribution shifts in the other direction; the visible separation of curves largely shows that differences in the mean values for these two variables decrease as depression symptom severity (in comparison to depression symptom scores WNL) increases. The plot for asleep distal body temperature shows more overlap in distributions but some separation of the curves is evident. For diurnal distal body temperatures, the mean within-individual standard deviation was 0.96 for the WNL category, 0.95 for the mild category, 0.91 for the moderate category, and 0.88 for the severe category.

#### *Magnitude of associations between wearable sensor-assessed body temperature metrics and depression symptom severity*

Kolmogorov–Smirnov D-statistics, which can be considered a measure of effect size, ranged from 0.094 to 0.225 for comparisons of severe symptoms versus WNL for all 4 distal body temperature metrics, while they ranged from 0.014 to 0.056 for comparisons between moderate symptoms versus WNL, and mild symptoms versus WNL (Table 5). Within each of the 4 distal body temperature metrics, the D-statistic was largest for the comparisons of severe symptoms versus WNL and smallest for comparisons of mild symptoms versus WNL. Relative to individuals WNL, individuals with severe depressive symptoms had distal body temperature distributions that were higher during the awake and asleep periods and had lower diurnal amplitudes and smaller asleep–awake

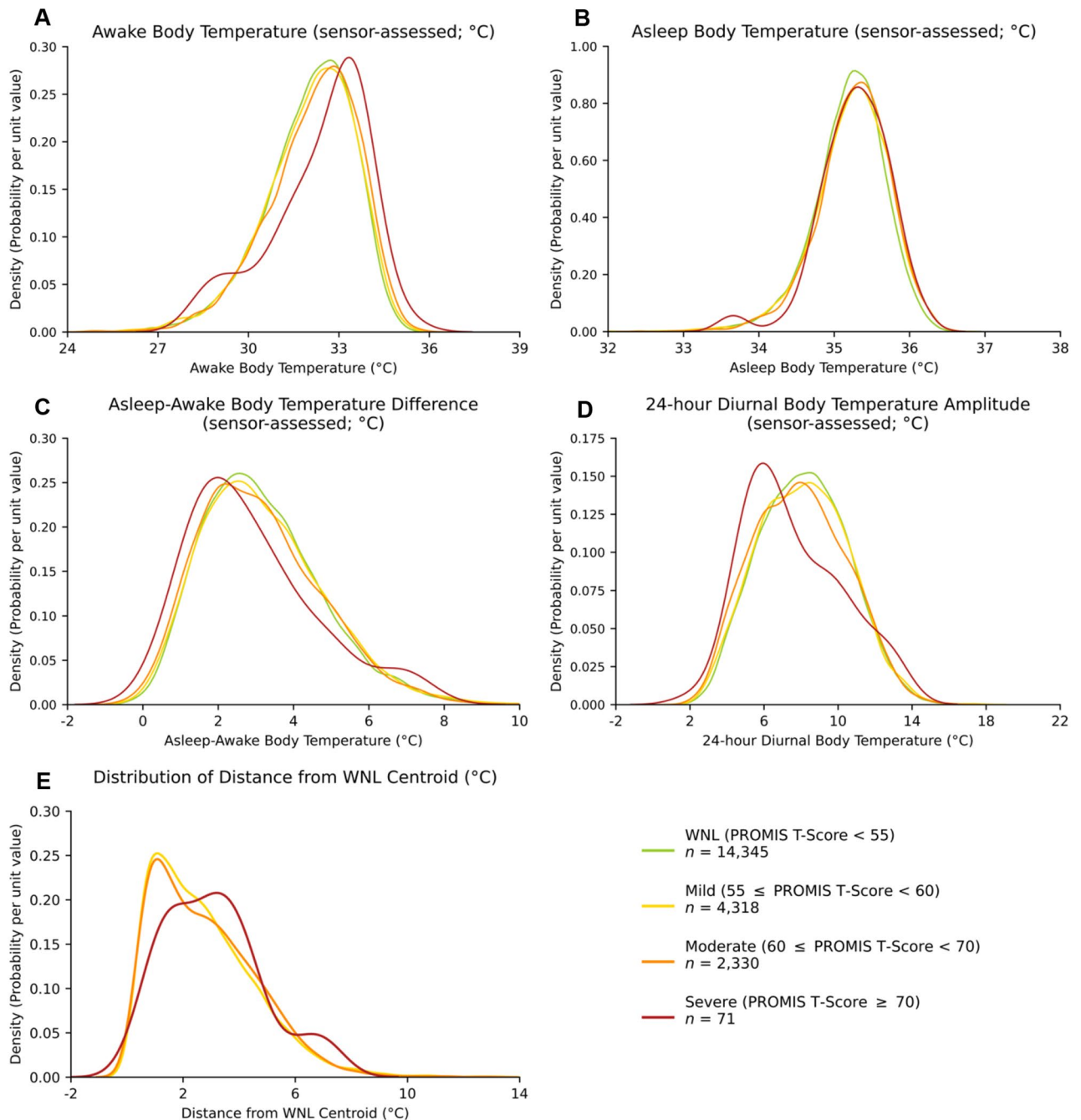


**Figure 4.** Receiver operating characteristic (ROC) analyses. Figure panel depicts ROC curves based on unadjusted and adjusted logistic regression models for PROMIS depression T-scores in the severe range (Panel A), the moderate range (Panel B), and the mild range (Panel C), versus PROMIS depression symptom T-scores within normal limits (WNL).

Model	Outcome	Temperature Cut point (°C)	Sensitivity	Specificity	ROC AUC
Unadjusted	Severe vs WNL	36.613	60.87%	62.34%	0.635
Adjusted	Severe vs WNL	36.375	85.87%	34.05%	0.762
Unadjusted	Moderate vs WNL	36.624	96.91%	63.40%	0.557
Adjusted	Moderate vs WNL	36.350	75.17%	31.55%	0.672
Unadjusted	Mild vs WNL	36.624	42.72%	63.46%	0.537
Adjusted	Mild vs WNL	36.500	58.54%	47.47%	0.612

**Table 4.** Receiver operating characteristic (ROC) curve analyses for each logistic regression model predicting PROMIS depression symptom T-score categories (mild, moderate, and severe) versus PROMIS depression symptom T-scores within normal limits (WNL) from self-reported body temperature. *Note.* AUC = Area under the curve; Adjusted = models adjusted for age, biological sex, and self-reported body temperature assessment time of day (see Table 1 note and “Methods”).

temperature differences than individuals with moderate or mild depressive symptoms. The associated statistical tests revealed statistically significant differences between distribution of awake distal body temperature, the asleep-awake difference in distal body temperature, and the diurnal distal body temperature amplitude, but not for asleep distal body temperature, when comparing distributions of these metrics among participants with severe symptoms versus WNL. The distributions of all 4 metrics in participants with moderate symptoms were statistically significantly different from the distributions among participants with symptoms WNL. Only the distribution of asleep distal body temperature in participants with mild symptoms had statistically significant differences compared to participants WNL; comparisons of the distributions of the other 3 metrics between mild symptoms and WNL were not statistically significant.



**Figure 5.** Figure panel depicting wearable sensor-assessed body temperature metrics. Probability Density Plots showing distributions of: Awake distal body temperature (Panel A); asleep distal body temperature (Panel B); asleep-awake distal body temperature difference (Panel C); diurnal distal body temperature amplitude separately for individuals with PROMIS depression T-scores within normal limits (WNL; green), and within the mild (yellow), moderate (orange), and severe (red) PROMIS depression symptom T-score categories (Panel D); distribution of the Euclidean Distance (within mild, moderate, and severe depression subgroups) of three distal body temperature metrics (awake temperature, asleep-awake temperature difference, and diurnal distal body temperature amplitude) from the centroid of those metrics among individuals with depression scores WNL (Panel E).

The Kolmogorov-Smirnov D-statistics shown in Table 5 are a measure of the maximum vertical distance between WNL and all other depression symptom categories on the empirical Cumulative Distribution Function (eCDF) plots (Supplementary Fig. S4). The eCDF plots, as with the density plots in Fig. 5A–D, are another illustration of the shifts in distal body temperature metric distributions with depressive symptom severity. These analyses show that the distal body temperature distributions differed the most between individuals with depressive symptoms in the severe range and individuals with depression symptoms WNL.



Symptom severity (vs. WNL)	Variable	RBC [95% CI: LB, UB]	D-Statistic	p
Severe	Awake distal body temperature	0.173 [0.026, 0.320]	0.225	0.0012
	Asleep distal body temperature	0.088 [-0.044, 0.219]	0.094	0.52
	Asleep–awake distal body temperature difference	-0.129 [-0.270, 0.012]	0.161	0.045
	Diurnal distal body temperature amplitude	-0.158 [-0.303, -0.013]	0.201	0.0053
Moderate	Awake distal body temperature	0.055 [0.029, 0.081]	0.049	$1.2 \times 10^{-4}$
	Asleep distal body temperature	0.072 [0.046, 0.097]	0.056	$7.8 \times 10^{-6}$
	Asleep–awake distal body temperature difference	-0.028 [-0.053, -0.002]	0.038	0.0067
	Diurnal distal body temperature amplitude	-0.051 [-0.077, -0.025]	0.054	$1.6 \times 10^{-5}$
Mild	Awake distal body temperature	0.008 [-0.012, 0.028]	0.014	0.53
	Asleep distal body temperature	0.050 [0.030, 0.070]	0.051	$6.7 \times 10^{-8}$
	Asleep–awake distal body temperature difference	0.005 [-0.015, 0.025]	0.017	0.29
	Diurnal distal body temperature amplitude	-0.008 [-0.028, 0.012]	0.020	0.15

**Table 5.** Kolmogorov–Smirnov tests comparing wearable sensor-assessed distal body temperature metrics across PROMIS depression symptom T-score categories (severe, moderate, mild) versus PROMIS depression symptom T-scores within normal limits (WNL), with Rank Biserical Correlation (RBC) and its 95% confidence intervals (lower bound, upper bound) for each comparison. *Note.* See Table 2 note.

The rank biserial correlation coefficients (RBC) comparing each distal body temperature metric for individuals with severe, moderate, or mild depressive symptoms (relative to individuals WNL) showed that the RBC increased as depressive symptoms increased for awake distal body temperature and asleep distal body temperature, consistent with the D-statistics and the increasing mean distal body temperatures with increasing depression symptom severity for these metrics (Table 6). The asleep–awake difference in distal body temperature and diurnal distal body temperature amplitude generally both decreased as symptom severity increased, and correspondingly, the RBC decreased as symptom severity increased.

#### *Standardized (to WNL) comparisons of body temperature metrics between mild, moderate, and severe depression symptom severities*

A probability density plot showing the distribution of the Euclidean Distance of awake distal body temperature, the asleep–awake difference in distal body temperature, and diurnal distal body temperature amplitude to the centroid (mean of these measures among participants WNL), appears in Fig. 5E. This plot shows that all levels of depression symptom severity (mild, moderate, severe) shifted away from the WNL centroid but with much overlap between curves for mild and moderate depressive symptoms, and less overlap with the curve for severe depressive symptoms. This indicates that people with severe depressive symptoms differed more from the WNL values than did people with mild or moderate depressive symptoms. That is, the severe depressive symptoms showed a more pronounced shift away from WNL values compared to the WNL centroid, though this was not statistically significant (Kruskal–Wallis test,  $p = 0.065$ ).

## Discussion

To our knowledge, this is the largest study to date to examine the association between body temperature, assessed using both self-report methods and wearable sensors, and depressive symptoms in a geographically broad sample. In these analyses, higher levels of depressive symptoms were associated with higher body temperatures during time awake. We observed this finding using body temperatures assessed (1) at most once per day via self-collection using handheld thermometers, and (2) at most once per minute via an unobtrusive wearable sensor device (worn on the finger). We found that distal body temperatures collected by the wearable sensor device during sleep were fairly similar across depression categories and were higher than awake distal body temperatures;

Variable	PROMIS Depression Symptom T-Score Category Mean 95% CI: [LB, UB]			
	WNL	Mild	Moderate	Severe
Awake distal body temperature	32.038 [32.015, 32.062]	32.051 [32.006, 32.096]	32.166 [32.105, 32.226]	32.366 [31.983, 32.750]
Asleep distal body temperature	35.201 [35.193, 35.209]	35.237 [35.222, 35.253]	35.260 [35.240, 35.279]	35.266 [35.156, 35.377]
Asleep–awake distal body temperature difference	3.163 [3.137, 3.189]	3.186 [3.138, 3.235]	3.094 [3.029, 3.160]	2.900 [2.502, 3.298]
Diurnal distal body temperature amplitude	8.149 [8.110, 8.187]	8.119 [8.046, 8.191]	7.942 [7.842, 8.043]	7.575 [6.958, 8.191]

**Table 6.** Mean values and 95% confidence interval (lower bound, upper bound) of wearable sensor-assessed distal body temperature metrics for each PROMIS depression symptom T-score category. *Note.* See Table 2 note.

this resulted in smaller asleep–awake distal body temperature differences with increasing depressive symptom severity. We also found that the association between self-reported body temperature and depressive symptoms was robust to adjusting for time of day at which body temperature was assessed (as done in prior work with smaller samples)<sup>19</sup>. We also observed these associations between self-reported body temperature and depressive symptoms in models that statistically accounted for demographic factors that can affect body temperature<sup>29</sup>.

These findings confirm associations between depressive symptoms and body temperature reported in smaller studies (< 300 participants)<sup>13–16,19–21</sup>. Specifically, these analyses replicated prior results showing that daytime self-reported body temperature was associated with greater depressive symptoms<sup>19,20</sup> and build on one prior study showing that the asleep–awake body temperature difference was more than twice as large among controls relative to individuals with depression<sup>21</sup>. In contrast to prior work<sup>13,21</sup>; however, we did not observe an association between increased body temperature during sleep time among individuals with greater depressive symptoms. This may be due in part to key differences in measurement; prior studies<sup>13,21</sup> monitored rectal (core) temperature during the night, whereas the wearable sensor in these analyses was collected from the skin of the finger (distal). As core temperature typically decreases whereas peripheral temperature increases during sleep<sup>30</sup>, it is possible that body temperature increases we observed during sleep are specific to distal temperature. Additionally, these analyses went beyond single self-reported daily body temperature assessments by including minute-level wearable sensor-assessed distal body temperature data over the course of several months, which prior studies of depressive symptoms and body temperature have not done. Importantly, we observed these associations outside of the controlled laboratory setting, which lends further external validity to these findings.

It is uncertain whether the elevated body temperature observed in depression reflects increased metabolic heat production, decreased ability to induce thermoregulatory cooling, or a combination of both. Body temperature reflects a balance between metabolic heat generation and thermoregulatory heat loss, with these processes under tight control by an integrated neural and immune-based feedback system that involves both bodily and central nervous system processes<sup>12</sup>. Available data suggest that the inadequate ability to activate thermoregulatory cooling mechanisms, as indexed by a reduced ability to sweat, may play an important role in the body temperature alterations observed in depression<sup>15,31</sup>. Based on these and other findings, we have previously proposed that MDD may be associated with dysregulation of afferent warm signals from distal cutaneous sensors to the central nervous system, i.e., involving heat-defense mechanisms traveling along the spinoparabrachial pathway<sup>32</sup>, leading secondarily to decreases in sweating as assessed by measurement of skin conductance levels<sup>12,33–36</sup>. Indeed, a recent systematic review and narrative synthesis concluded that lower skin conductance level in individuals with depression versus healthy controls has been a consistent finding<sup>37</sup>.

Evidence suggests that individuals with MDD may have altered electrodermal activity (EDA), which is now the preferred term that encompasses historical terms related to electrical characteristics of the skin, such as electrodermal level, electrodermal response, galvanic skin response, psychogalvanic reflex, skin conductance, skin conductance level, skin conductance response, and sympathetic skin response. EDA is typically assessed as changes in the amount of sweat secreted by eccrine sweat glands in the hypodermis of the palmar region of the hand and plantar region of the foot. EDA has a tonic and a phasic component, with the tonic component being related to skin conductance level, and with the phasic component being related to faster-changing elements of the signal that can be associated with an acute stimulus (skin conductance response) or “spontaneous” or “nonspecific” (nonspecific skin conductance response)<sup>38,39</sup>. Reduced EDA was first associated with depression in 1890<sup>40</sup>, an observation repeatedly observed in depression since then<sup>37,41</sup>. The most consistent findings in individuals with MDD (relative to healthy controls), have been lower skin conductance level, increased skin conductance response latency, and lower skin conductance response amplitude, suggesting alterations of multiple elements of EDA in persons with depression<sup>37</sup>.

Correlations between body temperature metrics and depressive symptoms suggest potential common underlying pathophysiological mechanisms. For example, chronic stressors that contribute to risk for depression<sup>42</sup> may also impact thermoregulation dysregulation of the hypothalamic pituitary adrenal axis<sup>43–45</sup>. In addition, abnormal glutamate/gamma-aminobutyric acid signaling has been observed in MDD post-mortem brains<sup>46</sup>, and this altered excitatory/inhibitory balance could also contribute to dysregulated body temperature, as shown in models of hot flashes<sup>47</sup>. Finally, low-grade (micro- or para-) inflammation<sup>48</sup> could lead to both elevated body temperature<sup>49</sup> and depressive symptoms<sup>50,51</sup>. Interestingly, temperature sensitive channels have been implicated in MDD and bipolar disorder and have also been shown to regulate inflammatory responses<sup>52</sup> and body temperature<sup>53,54</sup>. Thus, temperature-sensitive immune-regulatory channels are potential targets for development of MDD treatments.

Associations between body temperature and depression might be relegated to the realm of academic interest were it not for data showing that interventions directly targeting thermoregulatory systems have yielded antidepressant effects. Although it may seem counterintuitive that interventions that temporarily raise body temperature could benefit a condition characterized by increased body temperature, acute exposure to high heat induces counter-regulatory thermoregulatory cooling processes that produce longer-term and sustained reductions in body temperature<sup>55</sup>. In the context of MDD, whole-body hyperthermia has been reported to elicit a rapid and sustained reduction in depressive symptoms following a single whole-body hyperthermia treatment designed to raise core body temperature to 38.5 °C<sup>56,57</sup>. Notably, in one of these trials, participants with depression and higher body temperatures prior to WBH tended to experience larger antidepressant responses<sup>57</sup>. Antidepressant effects have also been observed for other heat-based interventions, including hot yoga<sup>58</sup>, hyperthermic baths<sup>59</sup>, and infrared sauna lamps<sup>56,57</sup>. Though these initial studies suggest that alterations in abilities to regulate body temperature may be associated with at least some cases of depression, clarifying the biological pathways through which body temperature is altered in some individuals with depression may reveal more specific pathogenic mechanisms amenable to targeted treatment for individuals with depression and elevated body temperature.

Several limitations of this study warrant discussion and have implications for further investigation. Though most self-reported body temperatures were likely collected using oral assessment methods, participants used

personal thermometers to collect their body temperature, and did not report the body site from which they collected their temperature. Unless the site from which participants measured their temperature varied systematically by levels of depressive symptoms, this measurement issue is likely to have made it more difficult to observe an association between body temperature and mood because of increased noise in temperature measurement (thus attenuating observed effects and plausibly rendering them conservative estimates of the true association). Thus, that the reported effects emerged despite this methodological noise is noteworthy. Additionally, we report here the most conservative statistical analysis of difference between groups by aggregating across time to determine average temperature parameters per person. As Fig. 1 makes clear, there are substantial changes for all groups across the year, and we do not attempt to use these changes to reduce the unexplained variance in our comparisons. Given the consistency of the differences visible across the months of data collection, we expect that future classification analyses could generate substantially larger effect sizes than we report here by statistically accounting for time of day, week, month, and year in order to reduce the contribution of these non-random sources of variance from the statistical comparisons.

Although adjusted analyses accounted for the time of day at which participants collected their body temperature, standardizing the time of day at which participants collected their body temperature would have reduced measurement variability. Future research using self-reported body temperature should standardize the thermometer device used, the site from which participants collect their body temperature, and the time of day temperature measurements are collected. Wearable sensor-assessed body temperature may have some additional benefits, such as the provision of continuous assessments that can provide significantly more analytic power. Additionally, wearable sensors may reduce problems associated with human errors in self-collecting body temperature (e.g., improper thermometer placement) and self-reporting (e.g., forgetting to report a value, incorrectly typing in a value).

The wearable sensor-based temperature measurements we used in this study supported the main findings of associations between depressive symptoms and self-reported temperature. Further, they extended these findings, for example, by allowing us to assess associations between depressive symptoms and the difference between awake and asleep body temperature. The wearable sensor devices we used have their own limitations, however. The wearable sensor devices measured dermal temperature on the finger, which can differ from core body temperature. Self-reported body temperatures were likely oral body temperature, which is a metric of core body temperature. The wearable sensor may have artifacts due to measurement errors, for example, due to removing the ring from a finger. To limit measurement artifacts from the wearable sensor, we took steps such as removing the upper and lower 5% body temperature assessments.

We used average depressive symptom scores and body temperature values in these analyses. An important remaining question is whether changes in depressive symptoms are associated with changes in body temperature, and over what timescale. We were unable to assess this well, as depressive symptoms were relatively stable in most of our cohort, with the median range between the highest and lowest depression T-score being 3.9, which may constitute a clinically meaningful difference but relative to categorical depression levels represents a relatively small difference. Although limited prior data suggest that body temperature may decrease when depressive symptoms decrease<sup>21,57</sup>, we were not able to address whether altering body temperature may improve depressive symptoms. As noted, small pilot studies suggest this might be the case<sup>57</sup>, but this is a key area for more rigorous, larger studies in the future. The current analyses reinforce the rationale for such intervention research.

Although we were unable to assess many potential confounders, sensitivity analysis using E-values showed that unmeasured confounders would need to be larger than observed effects of several known relevant variables to explain away all observed effects in these models. Although there are several known risk factors for depression that were not included in this analysis, to function as a confounder, these factors would also need to be associated with elevated body temperature. Genetics play a role in depression, and plausibly may be related to body temperature; however, it seems possible that genetic links to depression and temperature may relate to genetic mediators of depression through a thermoregulatory pathway, rather than constituting unmeasured confounders in observed associations. Illness can play a role in depression and may also influence body temperature, and we were not able to control for this; however, the magnitude of unmeasured potential confounding effects of chronic illnesses causing both depression and elevated body temperature would need to be substantially larger than our observed effects, given the magnitude of the E-values we observed. In these data, ROC curve analysis illustrated the strength of the association between body temperature and depression, with moderate-to-high AUC values across all models, and the highest being for severe depression.

These data confirm an association between body temperature and depressive symptoms in a large sample using body temperature metrics (1) that participants self-collected and self-reported and (2) assessed by wearable sensor devices that automatically transmitted data to the research team, and replicate prior research showing that individuals with depression have smaller asleep–awake body temperature differences. Though depression is biologically, behaviorally, and psychologically heterogeneous, these findings suggest that body temperature may be a candidate biological marker of depression for some individuals with depressive symptoms. Treatments that target elevated body temperature in individuals with both elevated body temperature and depression may warrant further evaluation.

## Methods

### Study design and participants

The TemPredict Study<sup>26,27</sup> was a prospective, worldwide, cohort study that continuously collected physiological metrics (e.g., body temperature) using an off-the-shelf wearable device (Oura Ring) with the primary aim of developing an algorithm to identify the onset of COVID-19. In addition to data collected by the Oura Ring, participants self-reported their body temperature daily and completed monthly surveys. Average study engagement

was 127.1 days for the self-reported body temperature analyses and 87.7 days for the sensor-assessed distal body temperature analyses. Eligible participants were at least 18 years of age, possessed a smartphone that could pair with the Oura Ring, and could communicate in English. Eligible participants either already possessed an Oura Ring that they used for participation or were provided with an Oura Ring for participation (the study provided Oura Rings to frontline healthcare workers at several participating healthcare institutions). We conducted participant recruitment through invitations delivered within the Oura smartphone app; these invitations contained embedded weblinks for prospective participants to review study details and to begin the study enrollment process<sup>60</sup>. All participants provided electronic informed consent. Participants provided consent to use data from their personally owned Oura Ring prior to their enrollment; we therefore analyzed data collected by the wearable devices beginning in April 2020. We did not pay participants for participation. The University of California, San Francisco, Institutional Review Board (IRB; IRB# 20-30408) and U.S. Department of Defense, Human Research Protection Office (HRPO; HRPO# E01877.1a) approved of all study procedures, and all research was performed in accordance with relevant guidelines and regulations and the Declaration of Helsinki. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines<sup>61</sup>.

## Measures

### *Depression severity*

Participants received a monthly survey (Qualtrics Survey Software) via email that included the Patient Reported Outcomes Medical Information System (PROMIS) Adult Health Profile instrument for depression (v1.0 Form 4a) with the recall period modified (as described elsewhere<sup>62</sup>) to reflect the past month rather than the past 7 days<sup>63</sup>.

### *Self-reported body temperature*

Participants self-reported their demographic information, including age and biological sex, in a baseline survey. Participants self-reported their self-assessed body temperature once per day by accessing a daily self-reported survey (Qualtrics Survey Software) that was linked within the Oura App. The daily survey asked participants to report the highest body temperature value they collected within the last day (in °F or °C, per participant preference). For these daily self-reported temperature assessments, participants used their own thermometers (which were not provided by the study, and which may have been of any brand), and we did not ask participants to report the type of thermometer used or the method of body temperature assessment. Qualtrics generated a time stamp for each completed survey.

### *Wearable sensor-assessed body temperature*

All participants wore the Oura Ring Gen2 (ouraring.com), a commercially available wearable sensor device (Oura Health, Oulu, Finland), on a finger of their choosing, as previously described<sup>27</sup>. The Oura Ring connects to the Oura App (available in Google Play Store and Apple App Stores) via Bluetooth. Users can wear the ring continuously during daily activities and sleep in wet and dry environments. The Oura Ring assesses distal body temperature using a negative temperature coefficient (NTC) thermistor (resolution of 0.07 °C) on the internal surface of the ring, which is in relatively consistent contact with the skin. The wearable sensor registers distal (dermal, peripheral) body temperature readings from the palm side of the finger base every minute. As described elsewhere<sup>27,64</sup>, the Oura Ring also assesses other physiological metrics, including sleep parameters, heart rate, heart rate variability, respiratory rate, and activity metrics.

## Data analyses

We performed statistical analyses using Stata v16 and Python v3.8.10 with the *statsmodels* (v0.13.2), *scipy* (v1.8.0), and *scikit\_posthocs* (v0.8.0) packages.

### *Variable preparation*

**Depression severity.** Each item on the PROMIS depression Form 4a is scored from 1 (*never*) to 5 (*always*), and these item scores are summed to create a total raw summary score. We converted raw PROMIS depression summary scores to T-scores using the conversion tables in the PROMIS scoring manual (healthmeasures.net<sup>65</sup>). PROMIS T-scores are referenced to a general adult population with a mean score of 50 and standard deviation of 10, with higher T-scores indicating greater depression symptomatology<sup>63,66</sup>.

**Self-report body temperature.** We excluded self-reported body temperature data points that were above 38.0 °C, as these are not representative of baseline temperature and meet criteria for clinical fever. We excluded these periods of potential illness because the focus of this analysis was body temperature under usual conditions of daily life, and so we excluded known periods of aberrant physiology. We used the Qualtrics auto-generated timestamp to create self-reported body temperature “time of day variables” as done in prior work examining associations between oral temperature and self-reported depressive symptoms<sup>19</sup>. Specifically, we calculated the cosine (B1) and sine (B2) of  $2\pi t$  for each self-reported body temperature value, where “t” was the decimal proportion of the day during which the daily survey was completed.

**Wearable sensor-assessed distal body temperature.** For each participant, we used the two weeks of distal body temperature data prior to each completed PROMIS assessment. For each 24-h period, we discarded wearable sensor-assessed distal body temperature values below the 5th percentile value and above the 95th percentile. This served to exclude periods of non-wear and periods of clinical fever and to minimize influence from non-representative extreme values (e.g., outliers), as done in prior research with wearable sensor-assessed metrics<sup>67–69</sup>. We

used hypnogram data to determine periods of asleep versus awake. As done in prior studies of body temperature and depression<sup>19,21</sup>, we computed metrics for the daily average distal body temperature during minutes awake (“awake distal body temperature”), the daily average during sleeping minutes (“asleep distal body temperature”), and the difference between those averages (“asleep–awake distal body temperature difference”). To compute the diurnal distal body temperature amplitude, we calculated the daily maximum distal body temperature as the highest daily value and the daily minimum distal body temperature as the lowest daily value for each person and calculated the difference between these two values.

#### *Statistical approach*

**Self-report body temperature.** Analyses of self-reported body temperature data included participants who completed one or more PROMIS depression assessments and who provided at least one self-reported temperature assessment, resulting in a sample of 20,880 participants.

We computed an unadjusted linear regression model with participants’ average self-reported body temperature as the predictor variable and average PROMIS depression T-score as the outcome variable. We also computed a multiple linear regression model that adjusted for age and biological sex, both of which can affect body temperature<sup>29</sup>, as well as body temperature survey time stamp (as done in prior analyses of depression and body temperature<sup>19</sup>). We included each participant’s body temperature time-of-day variables (average B1 and B2) in adjusted models to hold constant the average effect of time of day for each participant’s average body temperature. We plotted body temperatures for each depression T-score category.

We then conducted unadjusted and adjusted logistic regression models with categorical depression level variables (mild, moderate, and severe, each with WNL as the reference category) as outcomes. We used established thresholds for PROMIS depression T-scores, with separate models for mild (T-score 55–59.99) vs. WNL (T-score below 55); moderate (T-score 60–69.99) vs. WNL; and severe (T-score 70 or greater) vs. WNL, as the outcome variable, and self-reported body temperature as the predictor, scaled per 0.1 °C<sup>63</sup>. We report odds ratios with 95% confidence intervals (CIs) for the odds for each depression level with a 0.1 °C increase in average self-reported body temperature and visualized this in a forest plot. For each linear and logistic regression model, we performed sensitivity analyses by calculating E-values<sup>70,71</sup>, which provide estimates of the magnitude of unmeasured confounding that could explain the observed effects and have been recommended as a type of sensitivity analysis<sup>70,72</sup>. We conducted receiver operating characteristics (ROC) curve analyses for each logistic regression model and calculated the area under the curve (AUC) to evaluate how well self-reported body temperature could identify depressive symptoms in three depression comparisons (WNL vs. mild; WNL vs. moderate, and WNL vs. severe), and used Youden’s Index to identify body temperature thresholds to evaluate sensitivity and specificity based on those threshold values<sup>73</sup>. Youden’s index can be calculated as (sensitivity + specificity – 1) and ranges between 1 and 0, with 1 indicating perfect sensitivity and specificity and 0 indicating results no better than chance<sup>73</sup>. Maximizing Youden’s index in ROC analysis identifies the value of a continuous predictor that best distinguishes cases and non-cases, and is equivalent to locating the point on the curve that is furthest from the line of chance<sup>74</sup>. To control for multiple comparisons, we set the significance level at  $p < 0.01$  for all statistical tests (all two-sided).

**Wearable sensor-assessed body temperature.** Analyses involving wearable sensor-assessed distal body temperature data included participants who completed one or more PROMIS depression assessments and wore an Oura Ring for at least seven 24-h periods with at least 4 h of asleep and awake distal body temperature data recorded during each 24-h period, resulting in a sample of 21,064 participants. We included distal body temperature data collected in the 2 weeks prior to each completed PROMIS depression assessment.

We grouped participants into WNL, mild, moderate, or severe depression PROMIS depression symptom categories<sup>75,76</sup>. We plotted probability density plots to illustrate differences in the distributions of the four distal body temperature metrics by severity of depressive symptoms. We then calculated Kolmogorov–Smirnov Distance (D-statistic) and *p*-values, together with rank biserial correlations and their corresponding 95% confidence intervals, to quantify the magnitude of the association between distal body temperature metrics and depression symptom severity. We also created empirical Cumulative Distribution Function (eCDF) plots. Briefly, the Kolmogorov–Smirnov D-statistic is proportional to the maximum vertical distance between two eCDF plots, and ranges from 0 (distributions entirely overlapping) to 1 (distributions completely separated), functioning as a relative effect size metric for the comparison between two empirical distributions<sup>77,78</sup>. Together, the above analyses allowed us to compare the distributions of distal body temperature metrics of the three depression symptom categories (mild, moderate, severe) relative to WNL. We also calculated Rank Biserial Correlations<sup>79,80</sup> with 95% confidence intervals<sup>81</sup>. To further explore distal body temperature differences between depression symptom categories, we computed the standardized distal body temperature metrics within each level of depression symptom severity (mild, moderate, and severe) against those from the WNL centroid (the Euclidean Distance). We calculated the WNL centroid as mean of awake distal body temperature, asleep–awake temperature difference in distal body temperature, and diurnal distal body temperature amplitude; we did not include asleep distal body temperature within this centroid calculation as asleep distal body temperature did not diverge based on depression symptom category. We calculated the Euclidean Distance as the distance between the WNL centroid and each participants’ values on each of three distal body temperature metrics (awake distal body temperature, asleep–awake temperature difference in distal body temperature, and diurnal distal body temperature amplitude). We then created density plots for the distribution of the Euclidean Distance; we did this within each depression category (mild, moderate, severe) and computed a Kruskal–Wallis test.

## Data availability

Oura's data use policy does not permit us to make wearable sensor-assessed (via the Oura Ring) data available to third parties. Self-report data (e.g., self-reported body temperature, self-reported depressive symptoms) can be made available; those seeking to reproduce our findings should contact Ashley Mason, PhD, and Benjamin Smarr, PhD for an online application to access the study data portal. This application process will require requesters to make a written commitment expressing agreements to not duplicate data, to not share data with third parties, and/or other confidentiality precautions.

Received: 26 September 2023; Accepted: 6 January 2024

Published online: 05 February 2024

## References

- World Health Organization. Health statistics and information systems: Depression (2021, accessed 29 august 2022). <https://www.who.int/news-room/fact-sheets/detail/depression>.
- The changing rate of major depression. Cross-national comparisons. Cross-National Collaborative Group. *JAMA* **268**(21), 3098–3105. <https://doi.org/10.1001/jama.1992.03490210080039> (1992).
- Major Depression. National Institute of Mental Health (NIMH) (2022, accessed 6 Dec 2022). <https://www.nimh.nih.gov/health/statistics/major-depression>.
- Zisook, S. *et al.* Effect of age at onset on the course of major depressive disorder. *Am. J. Psychiatry* **164**(10), 1539–1546. <https://doi.org/10.1176/appi.ajp.2007.06101757> (2007).
- Luo, Y., Kataoka, Y., Ostinelli, E. G., Cipriani, A. & Furukawa, T. A. Prescription patterns of antidepressants in the treatment of adults with major depression in the us between 1996 and 2015: A population representative survey based analysis. *Front. Psychiatry* **2020**, 11. <https://doi.org/10.3389/fpsy.2020.00035> (2020).
- Trivedi, M. H. *et al.* Evaluation of outcomes with citalopram for depression using measurement-based care in STAR\*D: Implications for clinical practice. *AJP* **163**(1), 28–40. <https://doi.org/10.1176/appi.ajp.163.1.28> (2006).
- Rush, A. J. *et al.* Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR\* D report. *Am. J. Psychiatry* **163**(11), 1905–1917. <https://doi.org/10.1176/appi.ajp.163.11.28> (2006).
- Cipriani, A. *et al.* Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *FOC*. **16**(4), 420–429. <https://doi.org/10.1176/appi.focus.16407> (2018).
- Strawbridge, R., Young, A. H. & Cleare, A. J. Biomarkers for depression: Recent insights, current challenges and future prospects. *Neuropsychiatr. Dis. Treatment* **2017**, 13. <https://doi.org/10.2147/NDT.S114542> (2017).
- Østergaard, S. D., Jensen, S. O. W. & Bech, P. The heterogeneity of the depressive syndrome: When numbers get serious. *Acta Psychiatr. Scand.* **124**, 495–496. <https://doi.org/10.1111/j.1600-0447.2011.01744.x> (2011).
- Miller, A. H. & Raison, C. L. Burning down the house: Reinventing drug discovery in psychiatry for the development of targeted therapies. *Mol. Psychiatry* **2022**, 2. <https://doi.org/10.1038/s41380-022-01887-y> (2022).
- Raison, C. L., Hale, M. W., Williams, L., Wager, T. D. & Lowry, C. A. Somatic influences on subjective well-being and affective disorders: The convergence of therosensory and central serotonergic systems. *Front. Psychol.* **5**(1580), 1–22. <https://doi.org/10.3389/fpsyg.2014.01580> (2015).
- Avery, D. H., Wildschiodtz, G. & Rafaelsen, O. J. Nocturnal temperature in affective disorder. *J. Affect. Disord.* **4**(1), 61–71. [https://doi.org/10.1016/0165-0327\(82\)90020-9](https://doi.org/10.1016/0165-0327(82)90020-9) (1982).
- Souetre, E. *et al.* Twenty-four-hour profiles of body temperature and plasma TSH in bipolar patients during depression and during remission and in normal control subjects. *Am. J. Psychiatry* **145**(9), 1133. <https://doi.org/10.1176/ajp.145.9.1133> (1988).
- Avery, D. H., Shah, S. H., Eder, D. N. & Wildschindtz, G. Nocturnal sweating and temperature in depression. *Acta Psychiatr. Scand.* **100**(4), 295–301. <https://doi.org/10.1111/j.1600-0447.1999.tb10864.x> (1999).
- Duncan, W. C. Jr. Circadian rhythms and the pharmacology of affective illness. *Pharmacol. Therapeut.* **71**(3), 253–312. [https://doi.org/10.1016/S0163-7258\(96\)00092-7](https://doi.org/10.1016/S0163-7258(96)00092-7) (1996).
- Harding, E. C., Franks, N. P. & Wisden, W. The temperature dependence of sleep. *Front. Neurosci.* **2019**, 13. <https://doi.org/10.3389/fnins.2019.00336> (2019).
- Bach, V. *et al.* Chapter 14—Thermoregulation in Wakefulness and Sleep in Humans. In *Handbook of Clinical Neurology, Vol 98. Sleep Disorders Part I* (eds. Montagna, P. & Chokroverty, S.) 215–227 (Elsevier, 2011). <https://doi.org/10.1016/B978-0-444-52006-7.00014-9>.
- Rausch, J. L. *et al.* Depressed patients have higher body temperature: 5-HT transporter long promoter region effects. *Neuropsychobiology* **47**(3), 120–127. <https://doi.org/10.1159/000070579> (2003).
- Song, X. *et al.* Predictive markers of depression in hypertension. *Medicine* **97**(32), 1–8. <https://doi.org/10.1097/MD.00000000000011768> (2018).
- Szuba, M. P., Guze, B. H. & Baxter, L. R. Electroconvulsive therapy increases circadian amplitude and lowers core body temperature in depressed subjects. *Biol. Psychiatry* **42**(12), 1130–1137. [https://doi.org/10.1016/S0006-3223\(97\)00046-2](https://doi.org/10.1016/S0006-3223(97)00046-2) (1997).
- Avery, D. H., Wildschiodtz, G. & Rafaelsen, O. J. REM latency and temperature in affective disorder before and after treatment. *Biol. Psychiatry* **17**(4), 463–470 (1982).
- Sou ete, E. *et al.* Circadian rhythms in depression and recovery: Evidence for blunted amplitude as the main chronobiological abnormality. *Psychiatry Res.* **28**(3), 263–278. [https://doi.org/10.1016/0165-1781\(89\)90207-2](https://doi.org/10.1016/0165-1781(89)90207-2) (1989).
- Geneva, I. I., Cuzzo, B., Fazili, T. & Javaid, W. Normal body temperature: A systematic review. *Open Forum Infect. Dis.* **6**(4), 032. <https://doi.org/10.1093/ofid/ofz032> (2019).
- Taylor, N. A. S., Tipton, M. J. & Kenny, G. P. Considerations for the measurement of core, skin and mean body temperatures. *J. Therm. Biol.* **46**, 72–101. <https://doi.org/10.1016/j.jtherbio.2014.10.006> (2014).
- Smarr, B. L. *et al.* Feasibility of continuous fever monitoring using wearable devices. *Sci. Rep.* **10**(1), 1–11. <https://doi.org/10.1038/s41598-020-78355-6> (2020).
- Mason, A. E. *et al.* Detection of COVID-19 using multimodal data from a wearable device: Results from the first TemPredict Study. *Sci. Rep.* **12**(1), 1–15. <https://doi.org/10.1038/s41598-022-07314-0> (2022).
- Kroenke, K. *et al.* Minimally important differences and severity thresholds are estimated for the PROMIS depression scales from three randomized clinical trials. *J. Affect. Disord.* **266**, 100–108. <https://doi.org/10.1016/j.jad.2020.01.101> (2020).
- Waalens, J. & Buxbaum, J. N. Is older colder or colder older? The association of age with body temperature in 18,630 individuals. *J. Gerontol.: Ser. A.* **66A**(5), 487–492. <https://doi.org/10.1093/gerona/66A5> (2011).
- Szymusiak, R. Chapter 20—body temperature and sleep. In *Handbook of Clinical Neurology, Vol 156. Thermoregulation: From Basic Neuroscience to Clinical Neurology Part I* (ed. Romanovsky, A. A.):341–351 (Elsevier, 2018). <https://doi.org/10.1016/B978-0-444-63912-7.00020-5>.

31. Bagg, C. E. & Crookes, T. G. Palmar digital sweating in women suffering from depression. *Br. J. Psychiatry* **112**(493), 1251–1255. <https://doi.org/10.1192/bjp.112.493.1251> (1966).
32. Nakamura, K. & Morrison, S. F. A thermosensory pathway mediating heat-defense responses. *Proc. Natl. Acad. Sci.* **107**(19), 8848–8853 (2010).
33. Hale, M. W., Dady, K. F., Evans, A. K. & Lowry, C. A. Evidence for in vivo thermosensitivity of serotonergic neurons in the rat dorsal raphe nucleus and raphe pallidus nucleus implicated in thermoregulatory cooling. *Exp. Neurol.* **227**(2), 264–278 (2011).
34. Hale, M. W., Raison, C. L. & Lowry, C. A. Integrative physiology of depression and antidepressant drug action: Implications for serotonergic mechanisms of action and novel therapeutic strategies for treatment of depression. *Pharmacol. Therapeut.* **137**(1), 108–118. <https://doi.org/10.1016/j.pharmthera.2012.09.005> (2013).
35. Lowry, C., Flux, M. & Raison, C. Whole-body heating: An emerging therapeutic approach to treatment of major depressive disorder. *FOC* **16**(3), 259–265 (2018).
36. Lowry, C., Lightman, S. & Nutt, D. That warm fuzzy feeling: Brain serotonergic neurons and the regulation of emotion. *J. Psychopharmacol.* **23**(4), 392–400 (2009).
37. Sarchiapone, M. *et al.* The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis. *BMC Psychiatry* **18**(1), 22. <https://doi.org/10.1186/s12888-017-1551-4> (2018).
38. Boucsein, W. *Electrodermal activity*. Springer Science & Business Media (2012, accessed 20 Nov 2023). [https://books.google.com/books?hl=en&lr=&id=6N6rnOEZEEOC&oi=fnd&pg=PR3&dq=Boucsein,+W.+Electrodermal+activity+\(2nd+ed.\).+New+York:+Springer+Science+%26+Business+Media%3B+2012.&ots=B4hvoBKOLC&sig=WpEKOPA8ygzUAn1bqQS69pYMIaO](https://books.google.com/books?hl=en&lr=&id=6N6rnOEZEEOC&oi=fnd&pg=PR3&dq=Boucsein,+W.+Electrodermal+activity+(2nd+ed.).+New+York:+Springer+Science+%26+Business+Media%3B+2012.&ots=B4hvoBKOLC&sig=WpEKOPA8ygzUAn1bqQS69pYMIaO).
39. Braithwaite, J. J., Watson, D. G., Jones, R. & Rowe, M. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* **49**(1), 1017–1034 (2013).
40. Vigouroux, A. *Etude Sur La Résistance Électrique Chez Les Mélancoliques*. PhD Thesis. Faculté de médecine de Paris (1890).
41. Straub, R., Hole, G. & Wolfersdorf, M. Electrodermal hypoactivity in depression: Psychobiological marker or differential psychophysiological disposition?. *Schweizer Arch. Neurol. Psychiatr. Zurich Switzerl.* **143**(1), 41–59 (1992).
42. Colodro-Conde, L. *et al.* A direct test of the diathesis–stress model for depression. *Mol. Psychiatry* **23**(7), 1590–1596. <https://doi.org/10.1038/mp.2017.130> (2018).
43. Vinkers, C. H. *et al.* The effect of stress on core and peripheral body temperature in humans. *Stress* **16**(5), 520–530. <https://doi.org/10.3109/10253890.2013.807243> (2013).
44. Yaribeygi, H., Panahi, Y., Sahraei, H., Johnstson, T. P. & Sahebkar, A. The impact of stress on body function: A review. *EXCLI J.* **16**, 1057–1072. <https://doi.org/10.17179/excli2017-480> (2017).
45. Tan, C. L. & Knight, Z. A. Regulation of body temperature by the nervous system. *Neuron* **98**(1), 31–48. <https://doi.org/10.1016/j.neuron.2018.02.022> (2018).
46. Hu, Y. T., Tan, Z. L., Hirjak, D. & Northoff, G. Brain-wide changes in excitation-inhibition balance of major depressive disorder: A systematic review of topographic patterns of GABA- and glutamatergic alterations. *Mol. Psychiatry* **28**(8), 3257–3266. <https://doi.org/10.1038/s41380-023-02193-x> (2023).
47. Sun, Y. *et al.* Glutamatergic and GABAergic neurons in the preoptic area of the hypothalamus play key roles in menopausal hot flashes. *Front. Aging Neurosci.* **2022**, 14. <https://doi.org/10.3389/fnagi.2022.993955> (2023).
48. Medzhitov, R. Origin and physiological roles of inflammation. *Nature* **454**(7203), 428–435. <https://doi.org/10.1038/nature07201> (2008).
49. Evans, S. S., Repasky, E. A. & Fisher, D. T. Fever and the thermal regulation of immunity: The immune system feels the heat. *Nat. Rev. Immunol.* **15**(6), 335–349. <https://doi.org/10.1038/nri3843> (2015).
50. Miller, A. H. & Raison, C. L. The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* **16**(1), 22–34. <https://doi.org/10.1038/nri.2015.5> (2016).
51. Mac Giollabhui, N., Ng, T. H., Ellman, L. M. & Alloy, L. B. The longitudinal associations of inflammatory biomarkers and depression revisited: Systematic review, meta-analysis, and meta-regression. *Mol. Psychiatry* **26**(7), 3302–3314. <https://doi.org/10.1038/s41380-020-00867-4> (2021).
52. Miyanohara, J. *et al.* TRPM2 channel aggravates CNS inflammation and cognitive impairment via activation of microglia in chronic cerebral hypoperfusion. *J. Neurosci.* **38**(14), 3520–3533. <https://doi.org/10.1523/JNEUROSCI.2451-17.2018> (2018).
53. Song, K. *et al.* The TRPM2 channel is a hypothalamic heat sensor that limits fever and can drive hypothermia. *Science* **353**(6306), 1393–1398. <https://doi.org/10.1126/science.aaf7537> (2016).
54. Tan, C. L. *et al.* Warm-sensitive neurons that control body temperature. *Cell* **167**(1), 47–59.e15. <https://doi.org/10.1016/j.cell.2016.08.028> (2016).
55. Périard, J. D., Racinais, S. & Sawka, M. N. Adaptations and mechanisms of human heat acclimation: Applications for competitive athletes and sports. *Scand. J. Med. Sci. Sports* **25**(Suppl 1), 20–38. <https://doi.org/10.1111/sms.12408> (2015).
56. Janssen, C. W. *et al.* Whole-body hyperthermia for the treatment of major depressive disorder: A randomized clinical trial. *JAMA Psychiatry* **73**(8), 789–795. <https://doi.org/10.1001/jamapsychiatry.2016.1031> (2016).
57. Hanusch, K. U. *et al.* Whole-body hyperthermia for the treatment of major depression: Associations with thermoregulatory cooling. *Am. J. Psychiatry* **170**(7), 802–804. <https://doi.org/10.1176/appi.ajp.2013.12111395> (2013).
58. Nyer, M. *et al.* Community-delivered heated Hatha yoga as a treatment for depressive symptoms: An uncontrolled pilot study. *J. Altern. Complement. Med.* **25**(8), 814–823. <https://doi.org/10.1089/acm.2018.0365> (2019).
59. Naumann, J. *et al.* Effects of hyperthermic baths on depression, sleep and heart rate variability in patients with depressive disorder: A randomized clinical pilot trial. *BMC Complement. Altern. Med.* **17**(1), 1–9. <https://doi.org/10.1186/s12906-017-1676-5> (2017).
60. Qualtrics, X. M. Experience management software. Qualtrics (2022, accessed 27 Jul 2022). <https://www.qualtrics.com/>.
61. Von Elm, E. *et al.* The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Ann. Internal Med.* **147**(8), 573–577. <https://doi.org/10.7326/0003-4819-147-8-200710160-00010> (2007).
62. Batterham, P. J., Sunderland, M., Carragher, N. & Calear, A. L. Psychometric properties of 7- and 30-day versions of the PROMIS emotional distress item banks in an Australian Adult sample. *Assessment* **26**(2), 249–259. <https://doi.org/10.1177/1073191116685809> (2019).
63. Cella, D. *et al.* The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J. Clin. Epidemiol.* **63**(11), 1179–1194. <https://doi.org/10.1016/j.jclinepi.2010.04.011> (2010).
64. Grant, A. & Smarr, B. Feasibility of continuous distal body temperature for passive, early pregnancy detection. *PLOS Digital Health.* **1**(5), e0000034. <https://doi.org/10.1371/journal.pdig.0000034> (2022).
65. Healthmeasures.net. (2023, accessed 13 Jan 2023). <https://www.healthmeasures.net/>.
66. Reeve, B. B. *et al.* Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med. Care* **2007**, S22–S31 (2007).
67. Rodriguez-Morilla, B. *et al.* Application of machine learning methods to ambulatory circadian monitoring (ACM) for discriminating sleep and circadian disorders. *Front. Neurosci.* **2019**, 13. <https://doi.org/10.3389/fnins.2019.01318> (2023).
68. Nemati, S., Ghassemi, M. M. & Ambai, V. *et al.* Monitoring and detecting atrial fibrillation using wearable technology. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 3394–3397 (2016). <https://doi.org/10.1109/EMBC.2016.7591456>.

69. Shashikumar, S. P., Shah, A. J., Li, Q., Clifford, G. D. & Nemati, S. A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology. In *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* 141–144 (2017). <https://doi.org/10.1109/BHI.2017.7897225>.
70. VanderWeele, T. J. & Ding, P. Sensitivity analysis in observational research: Introducing the E-value. *Ann. Internal Med.* **167**(4), 268–274 (2017).
71. Mathur, M. B., Ding, P., Riddell, C. A. & VanderWeele, T. J. Website and R package for computing E-values. *Epidemiol. (Cambrid. Mass)* **29**(5), e45 (2018).
72. VanderWeele, T. J., Mathur, M. B. & Ding, P. Correcting Misinterpretations of the E-Value. *Ann. Intern. Med.* **170**(2), 131–132. <https://doi.org/10.7326/M18-3112> (2019).
73. Youden, W. J. Index for rating diagnostic tests. *Cancer* **3**(1), 32–35 (1950).
74. Böhning, D., Böhning, W. & Holling, H. Revisiting Youden's index as a useful measure of the misclassification error in meta-analysis of diagnostic studies. *Stat. Methods Med. Res.* **17**(6), 543–554. <https://doi.org/10.1177/0962280207081867> (2008).
75. McGraw, K. O. & Wong, S. P. A common language effect size statistic. *Psychol. Bull.* **111**, 361–365. <https://doi.org/10.1037/0033-2909.111.2.361> (1992).
76. Ruscio, J. A probability-based measure of effect size: Robustness to base rates and other factors. *Psychol. Methods* **13**(1), 19 (2008).
77. Vermeesch, P. Dissimilarity measures in detrital geochronology. *Earth-Sci. Rev.* **178**, 310–321 (2018).
78. Wilcox, R. R. Some practical reasons for reconsidering the Kolmogorov-Smirnov test. *Br. J. Math. Stat. Psychol.* **50**(1), 9–20. <https://doi.org/10.1111/j.2044-8317.1997.tb01098.x> (1997).
79. Cureton, E. E. Rank-biserial correlation. *Psychometrika* **21**(3), 287–290 (1956).
80. Kraemer, H. C. Correlation coefficients in medical research: From product moment correlation to the odds ratio. *Stat. Methods Med. Res.* **15**(6), 525–545. <https://doi.org/10.1177/0962280206070650> (2006).
81. Newson, R. Confidence intervals for rank statistics: Somers' D and extensions. *Stata J.* **6**(3), 309–334 (2006).

### Author contributions

Conception or design of the work: A.E.M., P.K., B.L.S., C.L.R.; acquisition of data: A.E.M., S.D., A.C., L.S.P., C.J.S., F.M.H.; analysis or interpretation of data: P.K., S.S., A.G., W.H., F.M.H.; creation of new software used in the work: P.K., S.S., A.G., B.L.S.; drafted or substantively revised the work: A.E.M., W.H., C.A.L., S.F., M.N., C.L.R., F.M.H. All authors reviewed the manuscript.

### Funding

This effort was funded under MTEC solicitation MTEC-20-12-Diagnostics-023 and the USAMRDC under the Department of Defense (#MTEC-20-12-COVID19-D.-023). The #StartSmall foundation (#7029991), and Oura Health Oy (#134650) also provided funding for this work. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government.

### Competing interests

CLR serves as a consultant for Usona Institute, Novartis, Emory Healthcare and Biogen/Sage. AEM and BLS have received financial remuneration from Oura Ring, Inc. for consulting. BLS has a financial interest in Oura Ring, Inc. CAL is Cofounder, Board Member, and Chief Scientific Officer of Mycobacteria Therapeutics Corporation, and is a member of the faculty of the Integrative Psychiatry Institute, Boulder, Colorado, the Institute for Brain Potential, Los Banos, California, and Intelligent Health Ltd, Reading, UK. All other authors do not have any competing interests to declare.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-51567-w>.

**Correspondence** and requests for materials should be addressed to A.E.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024