

<https://doi.org/10.1038/s44294-024-00037-9>

# Understanding wrist skin temperature changes to hormone variations across the menstrual cycle

Check for updates

Georgianna Lin<sup>1</sup> ✉, Jin Yi Li<sup>1</sup>, Ken Christofferson<sup>1</sup>, Shwetak N. Patel<sup>2</sup>, Khai N. Truong<sup>1</sup> & Alex Mariakakis<sup>1</sup>

Consumer devices are increasingly used to monitor peripheral body temperature (PBT) for menstrual cycle tracking, but the link between PBT and hormone variations remains underexplored. This study examines the relationship between these variables with a focus on nightly wrist skin temperature (WST). Fifty participants provided physiological and self-reported data, including WST, daily step counts, glucose levels, hormone levels (E3G, LH), and diary entries. Results show a negative correlation between WST and hormone levels when E3G and LH are below average, and this trend was robust to demographics and self-reported stress. Increased variance between mid-cycle hormonal peaks and WST fluctuations may stem from differences between basal body temperature (BBT) and WST. This research suggests that algorithms reliant on body temperature for tracking hormonal changes or other aspects of the menstrual cycle may need to account for increased variance in WST trends if they are meant to be deployed on wearable devices.

The menstrual cycle is a complex physiological process involving hormonal changes and interactions across multiple body systems<sup>1</sup>. One notable pattern that has been reported in clinical literature is the biphasic fluctuation of resting-state body temperature, with temperatures higher post-ovulation than pre-ovulation<sup>2</sup>. This fluctuation is influenced by the interplay of hormones such as estrogen, which lowers temperature, and progesterone, which has a thermogenic effect throughout the menstrual cycle. These hormonal dynamics regulate the release of luteinizing hormone (LH). Estrogen provides positive feedback in the late follicular phase to trigger the mid-cycle LH peaks, and progesterone inhibits LH release during the luteal phase while counteracting the positive feedback effect of elevated estrogen on LH.

This biphasic fluctuation pattern has led to the development of algorithms that rely on periodic changes in body temperature for purposes such as menstrual cycle tracking, pregnancy preparation, and contraception<sup>2</sup>. Previous research has typically relied on resting-state temperature measured at internal body sites like the esophagus or rectum because these sites closely reflect the thermal characteristics of central arterial or mixed venous blood<sup>2</sup>. However, measuring temperature at these sites can be inconvenient due to their invasiveness. Recent work has been done using peripheral body temperature (PBT) obtained from self-monitoring consumer devices like smartwatches and rings<sup>3–8</sup>, yet little has been done to explore either the relationship between PBT and hormone variations or the effectiveness of PBT in capturing anticipated patterns across the menstrual cycle<sup>2</sup>.

Research efforts examining the relationship between hormones and temperature have often centered around studies in which hormones are administered to rats and individuals with anovulatory cycles<sup>9–11</sup> or collected during select parts of the menstrual cycle<sup>2,12</sup>. During a study in which hormones were administered to four patients with anovulatory cycles, Davis et al.<sup>9</sup> observed a rise in temperature after the patients received progesterone and a subsequent decline following the discontinuation of estradiol and progesterone. Another study by Morris et al.<sup>13</sup> involved 27 participants collecting hormones over 10–14 consecutive days mid-cycle to measure the relationship between basal body temperature (BBT) and the LH surge. Their findings revealed that the LH surge occurred within 1 day of the BBT nadir for 81% of subjects who had such a distinct fiducial point in their hormone data. Other studies have utilized urinary LH tests to compare body temperature across the follicular and luteal phases, but these studies often do not capture fluctuations across the entire menstrual cycle<sup>6,7,14,15</sup>. Little to no work has examined how hormone fluctuations align with PBT throughout the entire menstrual cycle, perhaps due to the practical challenges of recording daily hormone and PBT data.

We investigate the relationship between resting PBT during sleep and daily hormone variation. In light of recent innovations in at-home hormone testing and continuous passive sensing, we compare WST recorded by smartwatches with levels of LH and estrone-3-glucuronide (E3G), one of the urinary metabolites of estradiol, using urinary hormone tests. We also

<sup>1</sup>University of Toronto, Computer Science, Toronto, ON, Canada. <sup>2</sup>University of Washington, Computer Science & Engineering, Seattle, WA, USA.

✉ e-mail: [blue.lin@mail.utoronto.ca](mailto:blue.lin@mail.utoronto.ca)

examine the influence of potential confounding factors like activity level and stress on this relationship across the entire menstrual cycle.

## Results

### Participant and cycle characteristics

A summary of participants' data is provided in Table 1, both collectively across the entire menstrual cycle and split across menstrual cycle phases. All physiological signals, as well as self-reported sleep issues, exhibited statistically significant variations across menstrual cycle phases according to  $\chi^2$  and repeated measures ANOVA tests.

Figure 1 shows the trends of nightly WST, daily E3G, and daily LH levels across the entire menstrual cycle. Participants' WST measurements peaked during the luteal phase ( $34.0 \pm 1.0$  °C) and were lowest during the late follicular phase ( $33.7 \pm 0.8$  °C). Furthermore, levels of E3G and LH were found to be highest during the ovulation phase ( $179.0 \pm 141.7$  ng/mL,  $10.3 \pm 16.4$  mIU/mL), with the lowest E3G values being recorded during the menstrual phase ( $97.7 \pm 73.3$  ng/mL) and the lowest LH values being recorded during the luteal phase ( $4.8 \pm 14.1$  mIU/mL).

### The effect of hormone levels on nightly WST

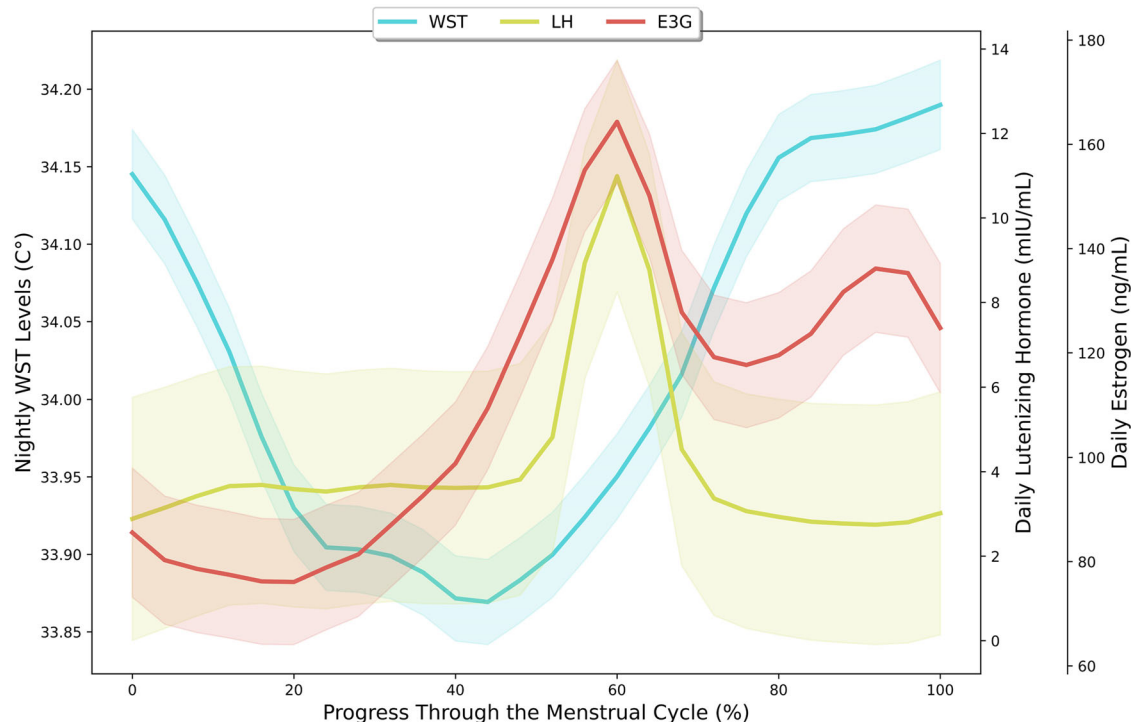
As shown in Table 2, four out of the ten different models yielded statistically significant relationships between hormone levels and nightly WST. In the linear model, higher LH levels were associated with statistically significant decreases in nightly WST throughout the menstrual cycle ( $\beta = -0.002$ ,  $p < 0.001$ , 95% confidence intervals (CI)  $[-0.004, -0.0009]$ ). In the quadratic model, statistically significant linear ( $\beta = -0.006$ ,  $p < 0.001$ , 95% CI  $[-0.009, -0.003]$ ) and quadratic terms ( $\beta = 0.00004$ ,  $p < 0.001$ , 95% CI  $[0.00001, 0.00006]$ ) indicated that higher levels of LH predicted lower nightly WST. However, this effect was only present for below-average levels of LH as seen in Supplementary Fig. 1. The interaction between E3G and LH was also statistically significant ( $\beta = -0.00001$ ,  $p < 0.001$ , 95% CI  $[-0.00002, -0.00005]$ ), suggesting that LH has significant relationships with WST despite being strongly correlated to E3G. For both the ovulation and the luteal phases, the inclusion of quadratic terms for the hormone level predictors revealed significant relationships between hormone and nightly WST levels. During the luteal phase, both the linear and quadratic terms for LH levels were significant, where higher LH levels were associated with

**Table 1 | Participant demographics and raw physiological signals summarized**

	All phases	Menstrual	Late-follicular	Ovulation	Luteal
Physiological signals (mean ± standard deviation)					
WST (°C)***	33.9 ± 0.9	33.9 ± 0.9	33.7 ± 0.8	33.8 ± 1.0	34.0 ± 1.0
LH (mIU/mL)***	6.4 ± 15.3	5.7 ± 17.1	5.4 ± 10.7	10.3 ± 16.4	4.8 ± 14.1
E3G (ng/mL)***	135.5 ± 110.8	97.7 ± 73.3	104.3 ± 80.0	179.0 ± 141.7	150.3 ± 110.7
Blood glucose (mmol/L)***	6.17 ± 0.76	6.19 ± 0.78	6.04 ± 0.71	6.15 ± 0.74	6.26 ± 0.78
Step count	7307.4 ± 5272.5	7362.0 ± 5436.0	7667.5 ± 5383.4	7015.5 ± 5133.9	7206.7 ± 5181.7
Demographics (mean ± standard deviation)					
Age (yrs)	20.7 ± 2.7				
Height (cm)	162.7 ± 5.9				
Weight (kg)	58.7 ± 10.7				
BMI (kg/m <sup>2</sup> )	22.6 ± 4.3				
Daily self-reported experiences (total count of daily entries, % within menstrual phase subset)					
Physical activity level					
Very low	675 (21%)	132 (22%)	171 (21%)	154 (21%)	218 (20%)
Low	1183 (36%)	234 (38%)	271 (34%)	265 (37%)	412 (37%)
Moderate	1006 (31%)	189 (31%)	252 (31%)	219 (30%)	346 (31%)
High	333 (10%)	41 (7%)	92 (11%)	72 (10%)	128 (11%)
Very high	56 (2%)	16 (3%)	18 (2%)	9 (1%)	13 (1%)
Sleep Issues**					
Did not experience	623 (19%)	87 (14%)	156 (19%)	155 (21%)	224 (20%)
Very low	573 (18%)	115 (19%)	168 (21%)	102 (14%)	188 (17%)
Low	788 (24%)	161 (26%)	191 (24%)	165 (23%)	271 (24%)
Moderate	716 (22%)	148 (24%)	153 (19%)	171 (24%)	244 (22%)
High	358 (11%)	65 (11%)	91 (11%)	75 (10%)	127 (11%)
Very high	203 (6%)	36 (6%)	46 (6%)	57 (8%)	64 (6%)
Stress					
Did not experience	428 (13%)	63 (10%)	103 (13%)	112 (15%)	150 (13%)
Very low	287 (9%)	54 (9%)	72 (9%)	52 (7%)	109 (10%)
Low	570 (17%)	117 (19%)	141 (18%)	122 (17%)	190 (17%)
Moderate	1061 (33%)	199 (33%)	265 (33%)	225 (31%)	371 (33%)
High	614 (19%)	125 (20%)	146 (18%)	141 (19%)	202 (18%)
Very high	301 (9%)	54 (9%)	78 (10%)	73 (10%)	96 (9%)

Participant demographics and raw physiological signals reported in aggregate and separated according to menstrual phase.

Significant results (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ) of  $\chi^2$  and repeated measures ANOVA tests of symptom variance across cycle phases are also indicated.



**Fig. 1 | Plot of nightly WST, LH, and E3G levels throughout the menstrual cycle.** A plot of nightly wrist skin temperature (WST), luteinizing hormone (LH), and estrogen (E3G) levels throughout all participants’ menstrual cycles starting from menstruation with LOESS smoothing. The shaded regions indicate 95% confidence intervals.

**Table 2 | Model results for the effect of hormone levels on nightly WST**

	E3G (ng/mL) β-coefficient [95% CI]	LH (mIU/mL) β-coefficient [95% CI]	E3G × LH β-coefficient [95% CI]
<b>Linear model</b>			
Entire cycle	0.00004 [−0.0001, 0.0002]	−0.002 [−0.004, −0.0009]***	−0.00001 [−0.00002, −0.000005]***
Menstrual	−0.00001 [−0.0005, 0.0005]	−0.004 [−0.01, 0.004]	−0.00003 [−0.00009, 0.00003]
Follicular	−0.0001 [−0.0006, 0.0003]	0.002 [−0.003, 0.007]	0.00001 [−0.00002, 0.00005]
Ovulation	0.00002 [−0.0003, 0.0003]	0.0003 [−0.002, 0.002]	−0.000002 [−0.00001, 0.000008]
Luteal	0.0001 [−0.0002, 0.0004]	−0.002 [−0.006, 0.001]	−0.00002 [−0.00004, 0.000008]
<b>Quadratic model</b>			
<b>Entire cycle</b>			
Linear term	0.00008 [−0.0002, 0.0003]	−0.006 [−0.009, −0.003]***	−0.000002 [−0.00001, 0.000009]
Quadratic term	−0.00000006 [−0.000001, 0.000001]	0.00004 [0.00001, 0.00006]***	−
<b>Menstrual</b>			
Linear term	−0.00004 [−0.0006, 0.0005]	−0.004 [−0.01, 0.005]	−0.00004 [−0.0001, 0.00003]
Quadratic term	0.0000003 [−0.000002, 0.000003]	−0.00001 [−0.00006, 0.00004]	−
<b>Follicular</b>			
Linear term	−0.0002 [−0.0007, 0.0004]	0.005 [−0.003, 0.01]	−0.00001 [−0.00007, 0.00005]
Quadratic term	−0.0000006 [−0.000003, 0.000002]	−0.00004 [−0.0001, 0.00005]	−
<b>Ovulation</b>			
Linear term	−0.0004 [−0.0009, 0.00007]	0.001 [−0.003, 0.006]	−0.000008 [−0.00002, 0.000006]
Quadratic term	0.000002 [0.0000002, 0.000003]*	−0.00001 [−0.00005, 0.00002]	−
<b>Luteal</b>			
Linear term	0.0004 [−0.00003, 0.0008]	−0.008 [−0.01, −0.002]**	0.000008 [−0.00003, 0.000004]
Quadratic term	−0.000001 [−0.000003, 0.0000008]	0.00006 [0.00002, 0.0001]**	−

The results of linear mixed-effects models that were generated to estimate associations between hormones and nightly wrist skin temperature (WST). Significant model results are indicated as such: \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001.

lower WST for below-average levels of LH as seen in Supplementary Fig. 2 (linear term:  $\beta = -0.008, p < 0.01, 95\% \text{ CI } [-0.01, -0.002]$ ; quadratic term:  $\beta = 0.00006, p < 0.01, 95\% \text{ CI } [0.00002, 0.0001]$ ). There was also a statistically significant relationship between E3G and WST in the ovulation phase for the quadratic term, with higher estrogen levels leading to lower WST ( $\beta = 0.000002, p < 0.05, 95\% \text{ CI } [0.0000002, 0.000003]$ ). As shown in Supplementary Fig. 3, this relationship was only present for below-average E3G values.

**Univariate associations between WST, hormones, and confounds**

As reported in Table 3, we examined confounding variables that could affect nightly WST levels by including them in the four univariate models with statistically significant results. No significant associations were found between WST and demographic characteristics, step count, and stress levels. However, a relationship was found between glucose and WST for the two entire-cycle models ( $\beta = 0.04, p < 0.001, 95\% \text{ CI } [0.02, 0.06]$ ). High physical activity was correlated with lower WST ( $\beta = -0.05, p < 0.05, 95\% \text{ CI } [-0.11, -0.002]$ ) in the quadratic model for the entire cycle. Moreover, high levels of sleep issues were associated with statistically significant increases in WST throughout the entire cycle (linear model:  $\beta = 0.07, p < 0.05, 95\% \text{ CI } [0.002,$

$0.13]$ ; quadratic model:  $\beta = 0.06, p < 0.05, 95\% \text{ CI } [0.001, 0.12]$ ), and such effects increased in magnitude during the ovulation phase in particular ("high":  $\beta = 0.14, p < 0.05, 95\% \text{ CI } [0.03, 0.26]$ ; "very high":  $\beta = 0.15, p < 0.05, 95\% \text{ CI } [0.02, 0.28]$ ).

**Multivariate associations between WST, hormones, and confounds**

To better assess whether the significance of the identified relationships between nightly WST and hormone levels occurred merely due to confounding variables, we created a multivariate linear mixed-effects model with all the statistically significant confounds and the hormone variables. Further details of these models' results can be found in Table 4. Besides self-reported physical activity levels, all statistically significant relationships reported above persisted in both direction and significance. Thus, the relationships between WST, E3G, and LH levels were still notable in the presence of confounding variables.

**The temporal difference between WST and hormone variations**

Figure 2 shows the temporal alignment between mid-cycle hormone peaks<sup>13</sup>, WST peaks, and WST nadirs. E3G and LH peaks were observed around the day of ovulation with small differences of

**Table 3 | Univariate model results for associations between hormone levels, nightly WST, and confounding variables**

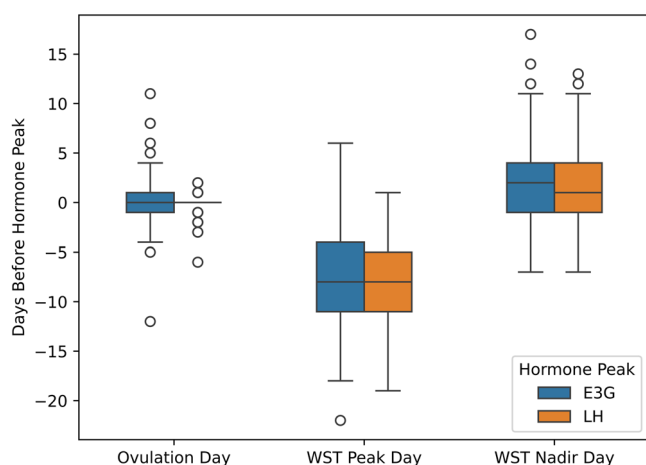
Covariate	Linear entire cycle $\beta$ -coefficient [95% CI]	Quadratic entire cycle $\beta$ -coefficient [95% CI]	Quadratic ovulation $\beta$ -coefficient [95% CI]	Quadratic luteal $\beta$ -coefficient [95% CI]
<b>Physiological signals</b>				
Blood glucose (mmol/L)	0.04 [0.02, 0.06]***	0.04 [0.02, 0.06]***	0.006 [-0.03, 0.05]	-0.02 [-0.05, 0.01]
Step count	0.0000001 [-0.000003, 0.000003]	0.0000002 [-0.000003, 0.000003]	-0.000005 [-0.00001, 0.000004]	-0.0000006 [-0.000005, 0.000004]
<b>Demographics</b>				
Age (yrs)	0.004 [-0.08, 0.08]	0.004 [-0.08, 0.09]	-0.009 [-0.09, 0.07]	0.003 [-0.06, 0.04]
Height (cm)	-0.01 [-0.06, 0.04]	-0.01 [-0.06, 0.04]	-0.008 [-0.06, 0.04]	-0.007 [-0.06, 0.04]
Weight (kg)	-0.02 [-0.04, 0.006]	-0.02 [-0.04, 0.006]	-0.02 [-0.04, 0.002]	-0.01 [-0.03, 0.01]
BMI (kg/m <sup>2</sup> )	-426.71 [-1103.79, 257.91]	-421.91 [-1103.52, 267.22]	-466.70 [-1107.03, 172.28]	-253.88 [-995.95, 487.24]
<b>Daily self-reported experiences</b>				
<b>Physical activity level</b>				
Very low	Reference	Reference	Reference	Reference
Low	-0.02 [-0.06, 0.01]	-0.03 [-0.06, 0.01]	0.02 [-0.05, 0.09]	-0.04 [-0.10, 0.01]
Moderate	-0.02 [-0.06, 0.02]	-0.02 [-0.06, 0.02]	0.03 [-0.04, 0.11]	-0.04 [-0.10, 0.02]
High	-0.05 [-0.11, 0.0006]	-0.05 [-0.11, -0.002]*	-0.09 [-0.18, 0.008]	-0.01 [-0.09, 0.06]
Very high	-0.05 [-0.16, 0.05]	-0.06 [-0.16, 0.04]	0.006 [-0.20, 0.21]	-0.03 [-0.19, 0.13]
<b>Sleep issues</b>				
Did not experience	Reference	Reference	Reference	Reference
Very low	-0.06 [-0.11, 0.003]	-0.06 [-0.11, 0.002]	0.12 [-0.005, 0.24]	-0.06 [-0.14, 0.03]
Low	0.01 [-0.04, 0.07]	0.01 [-0.04, 0.06]	0.06 [-0.04, 0.16]	-0.02 [-0.10, 0.06]
Moderate	-0.009 [-0.06, 0.05]	-0.01 [-0.07, 0.05]	0.06 [-0.04, 0.17]	-0.02 [-0.11, 0.06]
High	0.07 [0.002, 0.13]*	0.06 [0.001, 0.12]*	0.14 [0.03, 0.26]*	0.005 [-0.09, 0.10]
Very high	-0.03 [-0.10, 0.05]	-0.03 [-0.10, 0.04]	0.15 [0.02, 0.28]*	0.02 [-0.09, 0.13]
<b>Stress</b>				
Did not experience	Reference	Reference	Reference	Reference
Very low	-0.03 [-0.10, 0.05]	-0.04 [-0.11, 0.04]	-0.07 [-0.22, 0.08]	-0.008 [-0.12, 0.10]
Low	-0.05 [-0.11, 0.02]	-0.05 [-0.12, 0.01]	-0.04 [-0.16, 0.08]	-0.03 [-0.13, 0.07]
Moderate	-0.03 [-0.09, 0.03]	-0.03 [-0.10, 0.03]	0.001 [-0.12, 0.12]	-0.009 [-0.11, 0.09]
High	-0.03 [-0.09, 0.04]	-0.03 [-0.09, 0.04]	0.04 [-0.08, 0.16]	-0.06 [-0.16, 0.04]
Very high	-0.06 [-0.13, 0.02]	-0.06 [-0.13, 0.01]	0.08 [-0.05, 0.22]	0.05 [-0.05, 0.16]

The results of the univariate linear mixed-effects models that were generated to estimate relationships between hormone levels, nightly wrist skin temperature (WST), and confounds. Significant model results are indicated as such: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

**Table 4 | Multivariate model results for associations between hormone levels, nightly WST, and confounding variables**

Covariate	Linear entire cycle β-coefficient [95% CI]	Quadratic entire cycle β-coefficient [95% CI]	Quadratic ovulation β-coefficient [95% CI]	Quadratic luteal β-coefficient [95% CI]
Physiological signals				
E3G (ng/mL)	0.00005 [−0.0001, 0.0002]	0.0001 [−0.00008, 0.0003]	−0.0005 [−0.0008, −0.00008]*	0.0004 [0.00007, 0.0007]*
LH (mIU/mL)	−0.002 [−0.003, −0.001]***	−0.007 [−0.01, −0.005]***	0.001 [−0.002, 0.005]	−0.008 [−0.01, −0.003]***
E3G × LH	−0.00001 [−0.00002, −0.000006]***	0.0000009 [−0.000008, 0.00001]	−0.000008 [−0.00004, 0.00001]	0.000008 [−0.00002, 0.00003]
E3G quadratic	–	−0.0000001 [−0.000001, 0.0000008]	0.000002 [0.0000007, 0.000003]**	−0.000001 [−0.000003, 0.0000003]
LH quadratic	–	0.00005 [0.00003, 0.00007]***	−0.00002 [−0.00002, 0.000003]	0.00006 [0.00003, 0.00009]***
Glucose (mmol/L)	0.04 [0.01, 0.06]**	0.04 [0.02, 0.06]***	–	–
Daily self-reported experiences				
Physical activity level				
Very low	Reference	Reference	Reference	Reference
Low	–	−0.02 [−0.06, 0.02]	–	–
Moderate	–	−0.02 [−0.06, 0.02]	–	–
High	–	−0.05 [−0.10, 0.01]	–	–
Very high	–	−0.07 [−0.18, 0.04]	–	–
Sleep Issues				
Did not experience	Reference	Reference	Reference	Reference
Very low	−0.03 [−0.09, 0.04]	−0.04 [−0.10, 0.03]	0.12 [−0.006, 0.24]	–
Low	0.03 [−0.03, 0.09]	0.02 [−0.04, 0.08]	0.06 [−0.04, 0.17]	–
Moderate	0.008 [−0.05, 0.07]	−0.0004 [−0.06, 0.06]	0.06 [−0.05, 0.16]	–
High	0.09 [0.02, 0.16]**	0.08 [0.01, 0.15]*	0.13 [0.02, 0.25]*	–
Very high	−0.03 [−0.10, 0.05]	−0.03 [−0.11, 0.05]	0.14 [0.01, 0.27]*	–

The results of the multivariate linear mixed-effects models that were generated to estimate changes in nightly wrist skin temperature (WST). Significant model results are indicated as such: \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001.



**Fig. 2 | Boxplot of day differences between peaks in hormone levels and cycle events.** A boxplot displaying the time difference (in days) between peaks in luteinizing hormone (LH), and estrogen (E3G) levels and the following events: day of ovulation, peaks in nightly wrist skin temperature (WST) across the cycle, and nadir in WST across the cycle. The center line represents the median, bounds of box represents the lower (Q1) and upper (Q3) quartiles, and the whiskers extend to show the remaining distribution, with outliers represented with circles.

approximately  $0.2 \pm 2.5$  days and  $-0.2 \pm 2.5$  days, respectively. Furthermore, hormonal peaks were typically followed shortly after by days of WST nadirs, but with a slightly wider range. WST nadirs occurred  $2.3 \pm 4.5$  days before the day of peak E3G and  $1.8 \pm 4.0$  days before the day of peak LH. Meanwhile, WST peaks occurred  $8.2 \pm$

$4.2$  days after the day of peak E3G and  $7.6 \pm 4.8$  days after the day of peak LH.

### Discussion

Previous research suggests a strong connection between body temperature variations and menstrual cycle phases, potentially resulting from the secretion of hormones during this time<sup>2,16</sup>. To analyze these relationships across the menstrual cycle, we investigated the correlation between E3G and LH levels obtained from daily hormone tests in conjunction with nightly WST readings collected through a smartwatch. We did this by building a hierarchy of models to understand the effects of hormones, behaviors, and other physiological parameters on WST. We then identified a negative linear relationship between LH and WST, while the relationship between E3G and WST was only evident when considering non-linear associations and stratifying by menstrual cycle phase. Specifically, higher E3G levels during the ovulation phase were linked to lower WST, whereas higher E3G levels during the luteal phase corresponded to elevated WST. These relationships remained robust even after accounting for confounding factors such as glucose levels, physical activity levels, stress, and sleep quality.

Previous studies have found that E3G contributes to a cooling effect days within ovulation, while progesterone plays a key role in the elevation of body temperature in the luteal phase<sup>2</sup>. Consistent with previous research findings<sup>7,8,17</sup>, we observed a biphasic pattern of WST with the lowest values coinciding with peak levels of E3G and LH and the highest values occurring during the luteal phase after a dramatic decrease in LH. Since progesterone acts as a primary inhibitor of LH, the negative correlation observed between LH and body temperature is likely a result of its interaction with estrogen and progesterone throughout the menstrual cycle<sup>18</sup>. Additionally, the positive quadratic relationship between E3G and nightly WST is likely associated with the

interaction between LH and progesterone during the luteal phase, wherein peak progesterone levels result in elevated body temperature.

WST troughed around two days near the mid-cycle hormone peaks, and WST peaked approximately 7 days after. However, our study revealed a higher degree of variation compared to previous works that reported 0–2 days as the time between hormone peaks and resting-state body temperature nadirs<sup>13,19</sup> and 1–4 days as the time between hormone peaks and resting-state body temperature peaks<sup>20,21</sup>. This discrepancy may be attributed to differences in detecting ovulation based on temperature measured at peripheral versus internal body sites<sup>17,19</sup>. Zhu et al.<sup>17</sup> achieved higher true-positive and false-positive rates in ovulation detection when using WST instead of BBT, but we did not account for such discrepancies in our data analysis plan. Another explanation for the discrepancy between our work and prior literature is variation in hormonal concentrations between and within individuals at different time points across their menstrual cycles<sup>2</sup>. Prior studies have not collected longitudinal hormonal data because of the burden it imposes on study participants<sup>12,15,22–24</sup>, resulting in a limited scope of how hormone variation at different time points influence cycle phase timing and thermoregulation.

Our findings underscore the limitation of algorithms that assume a fixed offset between hormonal effects and temperature, as they fail to account for physiological diversity within and across individuals. However, the notable alignment between WST and hormones suggests that wearable devices could serve as accessible tools for understanding hormonal patterns. This is particularly significant given the current inaccessibility and high cost of hormone testing. Wearables present a potentially ubiquitous and less expensive alternative, prompting the need for further research into the feasibility and usability of such devices for hormonal insights. Clinicians may also contemplate the development of algorithms to predict hormonal health using signals collected by wearables, while carefully considering the privacy, security, and legal implications<sup>25</sup>.

We acknowledge several limitations in our work. Our study cohort consisted mostly of young adults, which may not fully represent all individuals who menstruate. Future work is needed on how our results generalize to other relevant demographics, including those who are perimenopausal or undergoing hormonal intervention. To enable daily hormone testing and support long-term study adherence<sup>26</sup>, we opted for commodity devices suitable for at-home usage. The hormone analyzer<sup>27–30</sup>, glucose monitor<sup>31</sup>, and smartwatch<sup>32</sup> we chose have all shown strong correlations with tests conducted in clinical settings, but they may still have marginally inferior reliability and precision during continuous data collection. Devices like clinically validated actigraphs or sensors designed for other peripheral sites may also yield different results compared to those used in our study<sup>33,34</sup>. Therefore, future studies should consider conducting similar analyses in clinical settings or using other data collection methods to support our observations.

Another limitation is that we might not have collected all potentially relevant hormones. Mira fertility testers only collected granular E3G and LH data at the time of this study, omitting progesterone and other relevant menstrual hormones such as follicle-stimulating hormone (FSH). Therefore, we encourage future studies to incorporate a more comprehensive range of hormonal data. We also recognize that there may be other confounding factors that influence temperature and were not considered in our study, such as alcohol consumption<sup>35</sup>, travel<sup>36</sup>, and illness<sup>8</sup>. Although we asked participants to self-report their sleep quality, sleep issues can be induced by various factors including light exposure, the lunar cycle, and restlessness due to high motor activity; each of these factors may themselves influence blood glucose metabolism and nightly WST<sup>34,37–40</sup>. We did not record all of these variables to avoid overburdening participants with their electronic diary entries, but future studies could examine these relationships more deliberately to investigate their interplay with hormonal variations and individual differences in body temperature.

## Methods

### Subjects

Fifty Canadian volunteers were recruited via social media groups and workspaces operated by women's health advocacy organizations in the Greater Toronto Area. Recruitment was limited to menstruating individuals who were between 18 and 30 years old and had not used hormonal therapy or hormonal contraception for at least three months prior to joining the study. Participants were informed that they could request to withdraw from the study at any point and that those who become pregnant and/or fail to adhere to the study protocol would be de-enrolled.

### Study design and data collection

The study protocol received approval from the Research Ethics Board at the University of Toronto (Protocol #41568), which evaluates research according to institutional, provincial, and federal practices along with the Declaration of Helsinki to ensure best practices in the protection of human subjects. Each participant electronically provided informed consent through e-signatures. For three consecutive months, each participant collected a variety of physiological signals associated with menstrual health via commercial wearable devices. Additionally, they self-reported aspects of their daily living in electronic diaries completed each morning. These diaries asked participants about their physical activity levels, assessed using a 5-point scale ranging from "very low" to "very high," alongside their sleep issues and stress levels, evaluated on a 6-point scale from "did not experience" to "very high".

Participants were given a Fitbit Sense smartwatch<sup>41</sup> to measure their nightly WST and a Dexcom G6 continuous glucose monitor (CGM)<sup>42</sup> to measure blood glucose. Participants were instructed to wear both devices at all times except during necessary maintenance activities like charging or sensor replacement. To collect WST data, users must be in an extended period of sleep of at least three hours while wearing their watch. Whenever this requirement is satisfied, Fitbit reports an average nightly temperature value based on the data collected during that period of sleep. Beyond these continuous data streams, participants were also asked to measure their LH and E3G hormone levels using the Mira Plus Starter Kit<sup>43</sup> daily. Measuring hormones with this device requires completing a disposable urine test wand and then waiting roughly 16 min for the Mira analyzer to finish its processing. Since the analyzer requires users to refrain from liquid consumption for at least two hours prior to collecting a urine sample, participants were advised to complete this hormone test every morning shortly after waking up.

Other researchers have assessed the accuracy and reliability of these devices. For example, Fitbit's PBT estimates yield an estimation error of  $0.23 \pm 0.08$  °C when compared against temperatures taken using the ThermoScan PRO 4000 ear thermometer<sup>44,45</sup>, while other studies have shown minimal differences between Fitbit's activity measurements (e.g., step count, heart rate) and corollaries collected by manual methods<sup>32</sup>. Meanwhile, David et al.<sup>31</sup> measured 98.7% matching between the Dexcom G6 CGM and point-of-care testing according to a Clarke Error Grid analysis. Finally, Nakhuda et al.<sup>27</sup> confirmed a strong correlation between serum E2 concentrations measured by immunoassay analyzers and urine E3G concentrations quantified by Mira devices. Their findings corroborate Mira's performance according to other studies that have observed 99% accuracy in ovulation prediction and 20% in coefficient of variation<sup>28–30</sup>.

### Data processing

Mira estimates the start and end dates of each menstrual cycle phase, along with the day of ovulation, via rises and falls in the collected E3G and LH levels. With this information, we generated four phase labels: follicular (the day after the last day of menstrual flow to the first day of the fertile window), ovulation (the fertile window), luteal (the last day of the fertile window to the first day of menstrual flow), and menstruation (days with menstrual flow). Cycles that had more than 4 consecutive days of E3G and LH data missing, more than 40% of E3G and LH data missing, and/or more than 20% of nightly WST data missing were dropped from the dataset. Based on these

criteria, 49 out of the 172 total cycles (28.4%) were eliminated, yielding a dataset with 123 cycles from 45 participants. To consolidate continuous physiological data into daily values, we calculated the total step count and median glucose levels for each day (12:00 AM–11:59 PM). We ensured the representativeness of summary statistics by removing days with less than 18 hours of continuous physiological data. For the cycles that remained, Fitbit, Dexcom CGM, and diary data were available for averages of  $98.3 \pm 0.9\%$ ,  $84.5 \pm 7.8\%$ , and  $89.7 \pm 2.4\%$  respectively within a given cycle.

For each menstrual cycle, we performed a series of preprocessing steps to account for noise in our sensor data. We used a process similar to the one used by Majjala et al.<sup>46</sup> to prepare the WST data; after applying linear interpolation to standardize the temporal spacing between data points and account for missing data, we applied a 5<sup>th</sup>-order Butterworth low-pass filter<sup>47</sup> with a cutoff at 10% of the target sampling frequency (1 sample per day) to attenuate noise in the temperature signal. We then identified the peaks and nadirs of nightly WST within each cycle (from the end of one menstrual phase to the start of the next) by pinpointing the lowest temperature before menstrual bleeding, followed by a minimum of three consecutive days of rising temperature during which the highest local maximum was recorded. This process aligns closely with the “three-over-six rule”<sup>48</sup>, which states that a notable peak in WST occurs when it is elevated by at least 0.2 °C for at least five of the preceding six days and it remains elevated for a minimum of three consecutive days<sup>8</sup>. We also identified peaks in E3G and LH levels by pinpointing the highest local maximum in every cycle. On average, the E3G peak was typically  $1.3 \pm 0.5$  times higher than the next highest E3G local maximum within a given cycle, while the LH peak was typically  $2.2 \pm 5.0$  times higher than the next highest LH local maximum within a given cycle.

## Statistics

To assess variations across the menstrual cycle and its phases, we present descriptive statistics for participant demographics and physiological signals in Table 1. We present these statistics both in aggregate and separated by menstrual cycle phase with  $\chi^2$  and repeated measures ANOVA test results to report notable differences. After visualizing nightly WST, E3G, and LH levels across temporally normalized cycles with LOESS smoothing, we then assess the impact of our collected data on nightly WST levels using linear mixed-effects models. We consider both linear and curvilinear relationships<sup>49</sup> across the entire menstrual cycle and stratified by menstrual cycle phase. We developed univariate models with potential confounders for every significant relationship identified, after which we constructed a multivariate model including all confounders with a statistically significant association with nightly WST in the univariate models. In all models, hormone levels were centered according to each person’s average levels, and we incorporated an interaction between E3G and LH to assess their dependency. Random slopes and intercepts were utilized in all models to account for WST variance between and within participants. The results of our models were reported according to  $\beta$ -coefficients, 95% CI, and the significance of each relationship within models, considering a statistical significance threshold of  $p < 0.05$ . We applied Holm–Bonferroni adjustments to correct for multiple comparisons. We conclude our analysis by presenting box plots showing the timing differences between WST peaks and nadirs and mid-cycle hormone peaks.

## Data availability

On reasonable request, the dataset generated during and/or analyzed in this paper is available from the corresponding author.

## Code availability

Data was collated in SQL databases, pre-processed with Python 3.10, and analyzed using R-3.5.1. The code used for the analysis of the dataset is available upon reasonable request from the corresponding author.

Received: 25 May 2024; Accepted: 19 September 2024;  
Published online: 04 October 2024

## References

- Messinis, I. E., Messini, C. I. & Dafopoulos, K. Novel aspects of the endocrinology of the menstrual cycle. *Reprod. Biomed. Online* **28**, 714–722 (2014).
- Baker, F. C., Sibozza, F. & Fuller, A. Temperature regulation in women: Effects of the menstrual cycle. *Temperature* **7**, 226–262 (2020).
- Chen, W., Kitazawa, M. & Togawa, T. Estimation of the biphasic property in a female’s menstrual cycle from cutaneous temperature measured during sleep. *Ann. Biomed. Eng.* **37**, 1827–1838 (2009).
- Murayama, Y. et al. Determination of biphasic menstrual cycle based on the fluctuation of abdominal skin temperature during sleep. *Adv. Biomed. Eng.* **12**, 28–36 (2023).
- Yu, J.-L. et al. Tracking of menstrual cycles and prediction of the fertile window via measurements of basal body temperature and heart rate as well as machine-learning algorithms. *Reprod. Biol. Endocrinol.* **20**, 118 (2022).
- Luo, L. et al. Detection and prediction of ovulation from body temperature measured by an in-ear wearable thermometer. *IEEE Trans. Bio Med. Eng.* **67**, 512–522 (2020).
- Goodale, B. M. et al. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: observational study. *J. Med. Internet Res.* **21**, e13404 (2019).
- Shilahi, M. et al. Modern fertility awareness methods: wrist wearables capture the changes in temperature associated with the menstrual cycle. *Biosci. Rep.* **38**, BSR20171279 (2018).
- Davis, M. E. & Fugo, N. W. The cause of physiologic basal temperature changes in women. *J. Clin. Endocrinol.* **8**, 550–563 (1948).
- Buxton, C. L. & Atkinson, W. B. Hormonal factors involved in the regulation of basal body temperature during the menstrual cycle and pregnancy. *J. Clin. Endocrinol.* **8**, 544–549 (1948).
- Charkoudian, N. & Stachenfeld, N. Sex hormone effects on autonomic mechanisms of thermoregulation in humans. *Autonomic Neurosci.* **196**, 75–80 (2016).
- Stephenson, L. A. & Kolka, M. A. Esophageal temperature threshold for sweating decreases before ovulation in premenopausal women. *J. Appl. Physiol.* **86**, 22–28 (1999).
- Morris, N. M., Underwood, L. E. & Easterling Jr, W. Temporal relationship between basal body temperature nadir and luteinizing hormone surge in normal women. *Fertil. Steril.* **27**, 780–783 (1976).
- BS, H., RC, M. & TG, K. et al. Novel technique for confirmation of the day of ovulation and prediction of ovulation in subsequent cycles using a skin-worn sensor in a population with ovulatory dysfunction: a side-by-side comparison with existing basal body temperature algorithm and vaginal core body temperature algorithm. *Front. Bioeng. Biotechnol.* **10**, 807139 (2022).
- Coyne, M. D., Kesick, C. M., Doherty, T. J., Kolka, M. A. & Stephenson, L. A. Circadian rhythm changes in core temperature over the menstrual cycle: method for noninvasive monitoring. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **279**, R1316–R1320 (2000).
- Schmalenberger, K. M. et al. A systematic review and meta-analysis of within-person changes in cardiac vagal activity across the menstrual cycle: Implications for female health and future studies. *J. Clin. Med.* **8**, <https://www.mdpi.com/2077-0383/8/11/1946> (2019).
- Zhu, T. Y. et al. The accuracy of wrist skin temperature in detecting ovulation compared to basal body temperature: Prospective comparative diagnostic accuracy study. *J. Med. Internet Res.* **23**, e20710 (2021).
- Skinner, D. C., Caraty, A. & Allingham, R. Unmasking the Progesterone Receptor in the Preoptic Area and Hypothalamus of the Ewe: No Colocalization with Gonadotropin-Releasing Neurons. *Endocrinology* **142**, 573–579 (2001).
- Quagliarello, J. & Arny, M. Inaccuracy of basal body temperature charts in predicting urinary luteinizing hormone surges. *Fertil. Steril.* **45**, 334–337 (1986).

20. de Mouzon, J., Testart, J., Lefevre, B., Pouly, J.-L. & Frydman, R. Time relationships between basal body temperature and ovulation or plasma progestins. *Fertil. Steril.* **41**, 254–259 (1984).
21. Kusuda, M. & Onoue, T. A method for estimating the day of ovulation by basal body temperature chart. *Asia-Ocean. J. Obstet. Gynaecol.* **7**, 43–48 (1981).
22. Shibui, K. et al. Diurnal fluctuation of sleep propensity and hormonal secretion across the menstrual cycle. *Biol. Psychiatry* **48**, 1062–1068 (2000).
23. Grant, L. K. et al. Menstrual phase-dependent differences in neurobehavioral performance: the role of temperature and the progesterone/estradiol ratio. *Sleep* **43**, zsz227 (2020).
24. Cagnacci, A., Volpe, A., Paoletti, A. M. & Melis, G. B. Regulation of the 24-hour rhythm of body temperature in menstrual cycles with spontaneous and gonadotropin-induced ovulation. *Fertil. Steril.* **68**, 421–425 (1997).
25. Dong, Z., Wang, L., Xie, H., Xu, G. & Wang, H. Privacy analysis of period tracking mobile apps in the post-roe v. wade era. In *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*, 1–6 (IEEE, 2022).
26. Zaat, T. R. et al. Home- or hospital-based monitoring to time frozen embryo transfer in the natural cycle? Patient-reported outcomes and experiences from the Antarctica-2 randomised controlled trial. *Hum. Reprod.* **35**, 866–875 (2020).
27. Nakhuda, G. S., Li, N., Yang, Z. & Kang, S. At-home urine estrone-3-glucuronide quantification predicts oocyte retrieval outcomes comparably with serum estradiol. *FS Rep.* **4**, 43–48 (2023).
28. Zhang, J. Algorithm effectiveness report of mira. <https://www.miracare.com/assets/algorithm-effectiveness.pdf> (2022). [Online; accessed 14-February-2023].
29. Li, N. Precision performance of the mira confirm test and comparison with the semi-quantitatively product. <https://www.miracare.com/assets/precision-performance-of-the-mira-confirm-test-and-comparison-with-the-semi-quantitatively-product.pdf> (2022). [Online; accessed 14-February-2023].
30. Bouchard, T. P., Fehring, R. J. & Mu, Q. Quantitative versus qualitative estrogen and luteinizing hormone testing for personal fertility monitoring. *Expert Rev. Mol. Diagn.* **21**, 1349–1360 (2021).
31. Davis, G. M. et al. Accuracy of Dexcom G6 Continuous Glucose Monitoring in Non-Critically Ill Hospitalized Patients With Diabetes. *Diab. Care* **44**, 1641–1646 (2021).
32. Xie, J. et al. Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: Comparative study. *JMIR Mhealth Uhealth* **6**, e94 (2018).
33. Areas, R., Duarte, L. & Menna-Barreto, L. Comparative analysis of rhythmic parameters of the body temperature in humans measured with thermistors and digital thermometers. *Biol. Rhythm Res.* **37**, 419–424 (2006).
34. Sarabia, J. A., Rol, M. A., Mendiola, P. & Madrid, J. A. Circadian rhythm of wrist temperature in normal-living subjects: A candidate of new index of the circadian system. *Physiol. Behav.* **95**, 570–580 (2008).
35. Wasielewski, J. A. & Holloway, F. A. Alcohol's interactions with circadian rhythms: A focus on body temperature. *Alcohol Res. Health* **25**, 94 (2001).
36. Vila, G. et al. Real-time monitoring of passenger's psychological stress. *Future internet* **11**, 102 (2019).
37. Gubin, D. G. et al. Activity, sleep and ambient light have a different impact on circadian blood pressure, heart rate and body temperature rhythms. *Chronobiol. Int.* **34**, 632–649 (2017).
38. Gubin, D. et al. Blue light and temperature actigraphy measures predicting metabolic health are linked to melatonin receptor polymorphism. *Biology (Basel)* **13**, 22 (2023).
39. Cajochen, C. et al. Evidence that the lunar cycle influences human sleep. *Curr. Biol.* **23**, 1485–1488 (2013).
40. Komada, Y. et al. The relationship between the lunar phase, menstrual cycle onset and subjective sleep quality among women of reproductive age. *Int. J. Environ. Res. Public Health* **18**, 3245 (2021).
41. Fitbit. Advanced health smartwatch | fitbit sense. <https://www.fitbit.com/global/us/products/smartwatches/sense> (2020). [Online; accessed 14-February-2023].
42. Dexcom Inc. Dexcom g6 cgm system: No fingersticks, no scanning. <https://www.dexcom.com/g6-cgm-system> (2022). [Online; accessed 14-February-2023].
43. Mira. Mira fertility us. <https://usd.miracare.com/products/fertility-plus-starter-kit> (2023). [Online; accessed 14-February-2023].
44. Song, C., Zeng, P., Wang, Z., Zhao, H. & Yu, H. Wearable continuous body temperature measurement using multiple artificial neural networks. *IEEE Trans. Ind. Inform.* **14**, 4395–4406 (2018).
45. Sunden, L., Deshpande, A. S. & Lafon, B. Methods, systems, and devices for improved skin temperature monitoring (U.S. Patent 11,737,675 B2, Aug. 29, 2023).
46. Maijala, A., Kinnunen, H., Koskimäki, H., Jämsä, T. & Kangas, M. Nocturnal finger skin temperature in menstrual cycle tracking: ambulatory pilot study using a wearable oura ring. *BMC Womens. Health* **19**, 1–10 (2019).
47. Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).
48. Marshall, J. *Planning for a Family: An Atlas of Mucothermic Charts* (Faber & Faber, 1979).
49. Clark, L. H., Schlosser, P. M. & Selgrade, J. F. Multiple stable periodic solutions in a model for hormonal control of the menstrual cycle. *Bull. Math. Biol.* **65**, 157–173 (2003).

## Acknowledgements

This research was funded in part by NSERC Discovery Grants RGPIN-2021-03457 and RGPIN-2021-04268, Google PhD Fellowship, and an unrestricted gift from Google.

## Author contributions

G.L., A.M., K.T., and S.P. conceived the experiment(s), G.L. conducted the experiment(s), G.L., H.L., and K.C. analyzed the results. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44294-024-00037-9>.

**Correspondence** and requests for materials should be addressed to Georgianna Lin.

**Reprints and permissions information** is available at <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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024