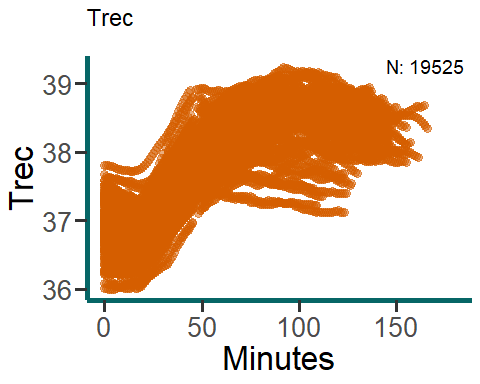
CORE results

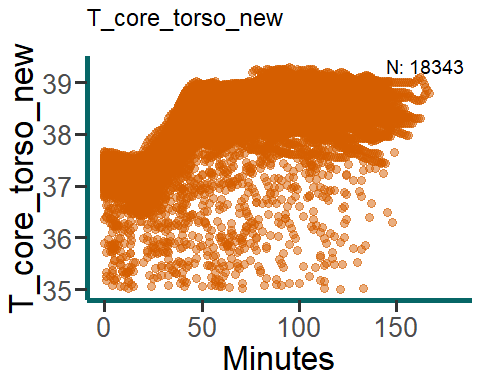
## Additional info (not for paper)

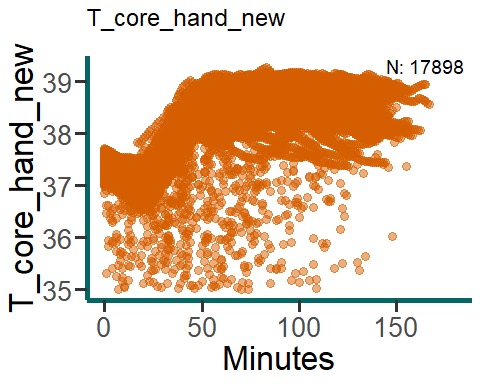
Within this section I highlight some additional information for us, as team, to discuss.

### First, we start with all the 1-minute averaged data of Trec, Tcore, Tsk, HR.

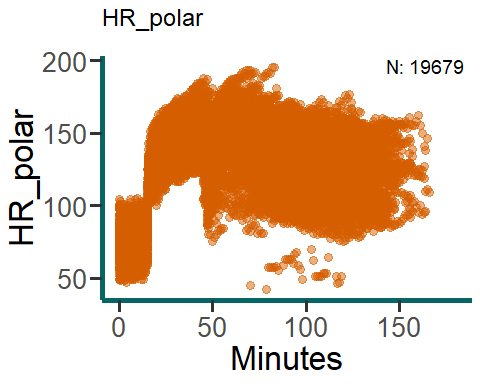


p25 is the one with the very high Trec. Most of the low Trec after min 50 are p7, who was incapable of cycling enough.

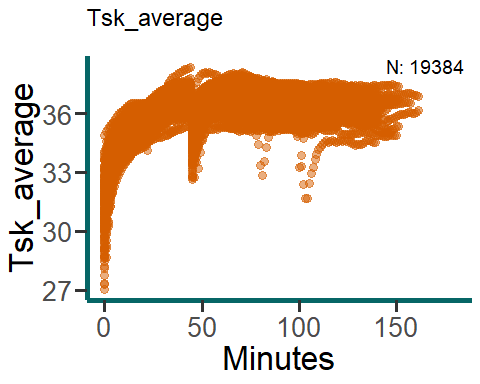




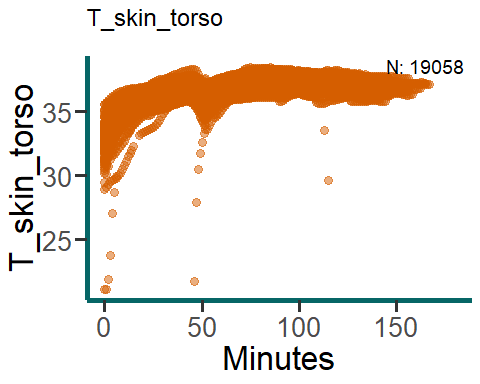
Both torso and hand show very low values sometimes. Not sure why. For now I leave it in.

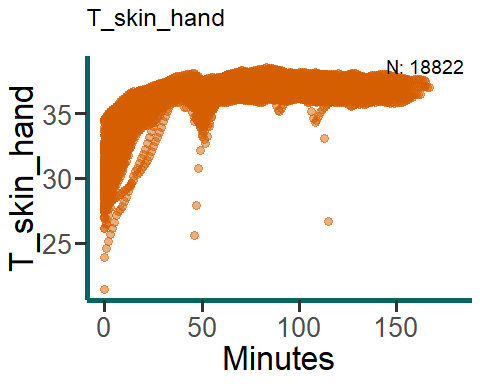


After removal of the very low HR values, there only are some low values left towards the end of the protocol. I suggest to leave it in because, at that time people might be sitting (so not cycling) for a long time already.



Some people start with cold skins, but this is very possible given the relatively cold lab outside of the climatic chamber. The drop somewhere before 50 represents the brief period that participants were out of the climatic chamber. The drops at 75 and 100 might be that someone stepped outside for a short period of time. I have left this data in for now.





Same goes for the skin temperatures measured by the CORE sensor, sometimes start low.

### Second, which to exclude?

As 2 participants dropped-out after 2 and 5 sessions due to injury (knee, p3) and sickness (maagzuur, p4), 16 participants finished the entire HA-protocol. One participant kept changing power output during the exercise phase and is therefore excluded (p18). Due to measurement error using the rectal probe (very high values, see figure Trec) , HA-1 for pp25 had to be excluded from Tc related research questions involving HA1 (Main research questions about HA-adaptations). Due to measurement error in CORE during HA1, p16 and p20 have to be excluded from Tc related research questions involving HA1 (Main research questions about HA-adaptations).

This leaves 12 participants who might be included in the mixed-models ANOVA.

Additionally, after calculation of Tc averages for the resting and constant cycling phase, 4 participants had one (N = 3) or two (N = 1) missing values for one of the measurement devices due to measurement error. I propose to replace the missing value with the average of the HA-session before and after the missing session. This causes the sample size for this research question to be 12 (instead of 8). ***Do we agree on this method?***

One participant (pp7) did not reach Tre > 38.5 any day. I suggest excluding here from the time > 38.5 RQ. Interestingly she did show large changes in resting (-0.6 a -0.7) and exercising (-0.5) Tre.

For the other research questions, where we have only Blant-Altman analysis and ICC, I suggest to use all the available data. That includes the drop-outs. Because why not? More data = more better ;)

### Third, Normality of data.

We have very large sample sizes (see graphs) for the RQ regarding wrist vs torso and overall validity, therefore we can assume normality in those cases. However, for the main research question and the time > 38.5 degr C question our sample is small (N = 9 or 14, see above) and moderate (N > 50) so we have to check.

Below are the results of my visual inspection of data normality. We have rest and ex(ercising) data which is Y(es) or N(ot) normaly distributed. There are some sessions which are not normaly distributed. However, in general, data is normally distributed. Therefore, I conclude we can proceed with parametric tests. ***do you agree?***

***dTre: rest & ex YES (excl pp25) - ALL DATA*** -

* rest Y: HA2,4-9 -
* rest N: HA3 -
* ex Y: HA2-9

***dTCore\_torso: rest & ex YES (excl pp25) - ALL DATA*** -

* rest Y: HA2-9 -
* ex Y: HA2-4,7 -
* ex N: HA5, 6, 8, 9

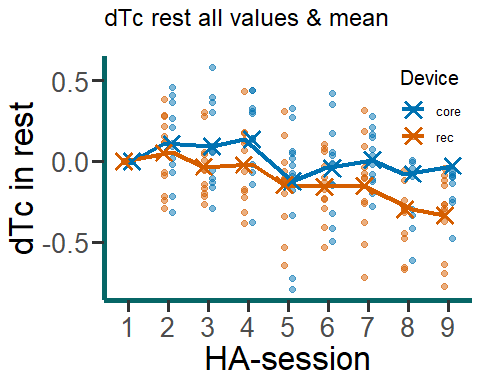
***dTCore\_hand: rest & ex NO (excl pp25)*** -

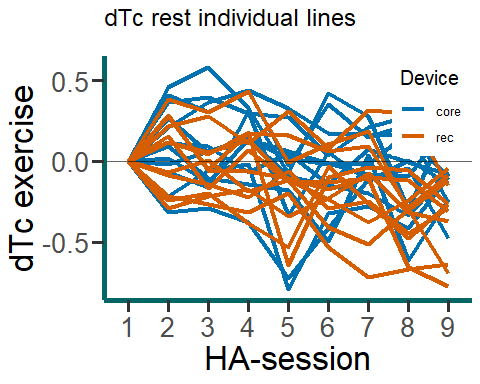
* rest Y: HA2-7
* rest N: HA8,9
* ex N: HA2, 5-8
* ex N: HA3, 4, 9

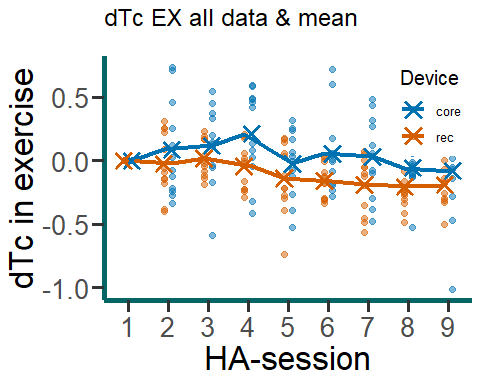
## HA-adaptations

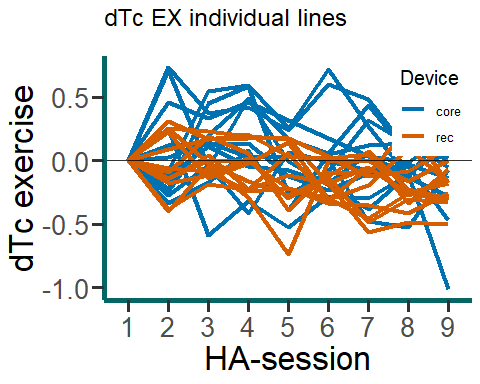
A significant interaction effect of HA and Device on dTc during rest was found, F(8 , 88 ) = 2.25 , P = 0.031 , ges = 0.04 . A significant main effect of HA on dTc during exercise was found F(8 , 88 ) = 7.62 , P = 1.08^{-7} , ges = 0.15 .

Specifically, dTc in rest measured using the rectal probe was significantly reduced for HA6 (P = 0.0471), HA8 (P = 2.1^{-4}), and HA9 (P = 3^{-5}) compared to HA1, while dTc in rest estimated using the CRS did not significantly change throughout the HA-protocol (P > 0.05). Additionally, dTc during exercise measured using the rectal probe was significantly reduced for HA8 (P = 0.03541) and HA9 (P = 3^{-5}) compared to HA1, while dTc during exercising estimated using the CRS was only significantly reduced during HA4 (P = 0.03146) compared to HA1.



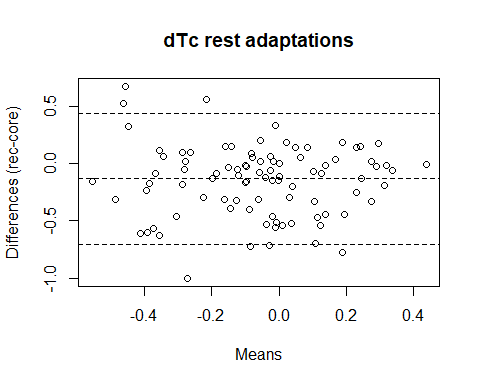




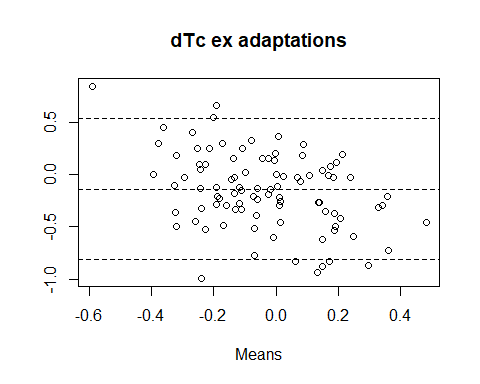


The ICC, mean bias and lower- and upper limit of dTc measured by the rectal probe and CRS are presented in the table below. The Bland-Altman plots are displayed in figure X.

| Variable | Mean Diffs | Lower Limit | Upper Limit | ICC Value | ICC Interpretation |
| --- | --- | --- | --- | --- | --- |
| dTc\_rest\_mean | -0.15 | -0.76 | 0.45 | 0.31 | poor |
| dTc\_ex\_mean | -0.16 | -0.87 | 0.56 | 0.10 | poor |



NULL

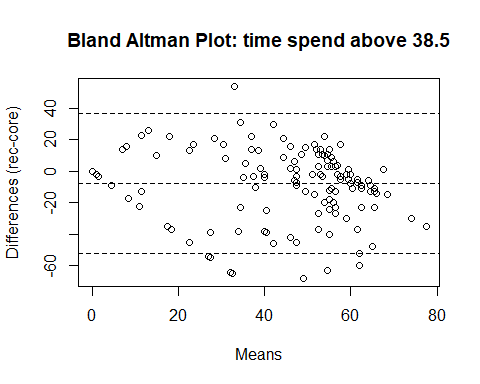


NULL

## Time spend above 38.5

One participant did not spend any time with a Tre > 38.5 and was therefore excluded from the analysis. The average time spend with a Tc at or above 38.5 was 40 +/- 21 min when measured using the rectal probe and 48 +/- 23 min when estimated using the CRS.

137 observations were included. The ICC of the time spend above 38.5 is 0.44 (95% CI: 0.28-0.57), F(136, 91) = 2.74, p = 0, indicating poor reliability. The Bland-Altman plot is shown below. The mean bias was -8 min with LoA of +/- 45 (range: -52 to 37 min). The negative value of the mean bias indicates that the CRS typically overestimated the time spend above 38.5 compared to the rectal probe.

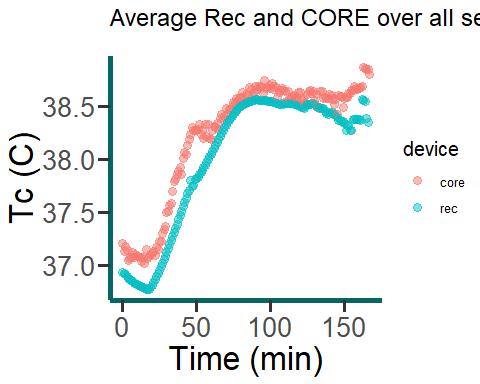


NULL

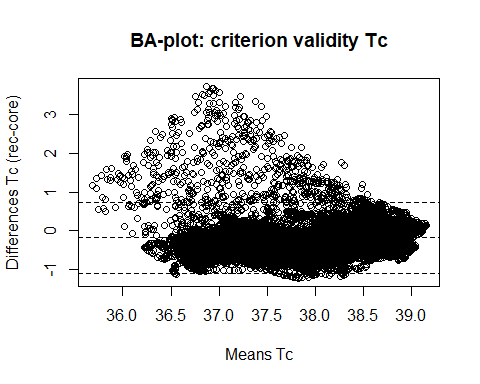
## Criterion validity CORE vs REC

The average Tc profile for the CRS and rectal probe are presented in the figure below.

Warning: Removed 1 rows containing missing values (`geom\_point()`).



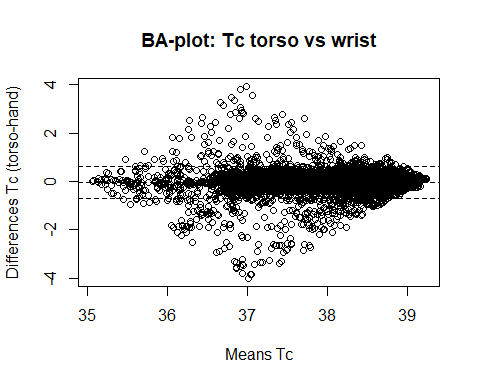
17668 observations were included. The ICC of the CRS to estimate Tc compared to the rectal probe is 0.77 (95% CI: 0.69-0.82), F(1.7667^{4}, 76) = 8.68, p = 0, indicating good reliability. The Bland-Altman plot is shown below. The mean bias was -0.18 C with LoA +/- 0.93 (range: -1.11 to 0.74 C). The negative value of the mean bias indicates that the CRS typically overestimated Tc compared to the rectal probe.



NULL

## Criterion validity torso vs wrist

17655 observations were included. The ICC of the CRS on the wrist to estimate Tc compared to the CRS on the torso is 0.89 (95% CI: 0.88-0.89), F(1.7654^{4}, 7591) = 16.73, P = 0, indicating good reliability. The Bland-Altman plot is shown below. The mean bias was -0.03 C with LoA 1 (range: -0.69 to 0.62 C).



NULL