LJMU Research Online

Bongers, CCWG, Daanen, HAM, Bogerd, CP, Hopman, MTE and Eijsvogels, TMH

 Validity, Reliability, and Inertia of Four Different Temperature Capsule Systems.

http://researchonline.ljmu.ac.uk/id/eprint/7014/

Article

Validity, Reliability and Inertia of Four Different Temperature Capsule Systems

RUNNING TITLE: Comparison of Temperature Capsule Systems

Coen C.W.G. Bongers[1]

Hein A.M. Daanen[2]

Cornelis P. Bogerd[3]

Maria T.E. Hopman[1]

Thijs M.H. Eijsvogels[1,4]

[1]Radboud Institute for Health Sciences, Radboud university medical center, Department of Physiology, Nijmegen, The Netherlands

[2]Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, The Netherlands

[3]TNO, CBRN Protection / Training and Performance Innovations, P.O. Box 45, 2280 AA Rijswijk, The Netherlands

[4]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom

Total Word Count: 3,336

Abstract Word Count: 250

Total Number of Figures: 3

Total Number of Tables: 3

Author for correspondence:

Coen C.W.G. Bongers, Department of Physiology (392), Radboud university medical center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands, Tel +31 24 36 14468, Fax +31 24 36 16413, E-mail: Coen.Bongers@radboudumc.nl

34 **ABSTRACT**

35 **Purpose.** Telemetric temperature capsule systems are wireless, relatively non-invasive and

36 easily applicable in field conditions, and have therefore great advantages for monitoring core

37 body temperature. However, the accuracy and responsiveness of available capsule systems have

38 not been compared previously. Therefore, the aim of this study was to examine the validity,

39 reliability and inertia characteristics of four ingestible temperature capsule systems (i.e.

40 CorTemp, e-Celsius, myTemp and VitalSense).

41 **Methods.** Ten temperature capsules were examined for each system in a temperature controlled

42 water bath during three trials. The water bath temperature gradually increased from 33°C to

43 44°C during Trial 1 and 2 to assess the validity and reliability, and from 36°C to 42°C in Trial

44 3 to assess the inertia characteristics of the temperature capsules.

45 **Results.** A systematic difference between capsule and water bath temperature was found for

46 CorTemp (0.077°C±0.040°C), e-Celsius (-0.081°C±0.055°C), myTemp (-0.003°C±0.006°C)

47 and VitalSense (-0.017°C±0.023°C) ($p<0.010$), with the lowest bias for the myTemp system

48 ($p<0.001$). A systematic difference was found between Trial 1 and Trial 2 for CorTemp

49 (0.017°C±0.083°C, $p=0.030$) and e-Celsius (-0.007°C±0.033°C, $p=0.019$), whereas

50 temperature values of myTemp (0.001°C±0.008°C) and VitalSense (0.002°C±0.014°C) did not

51 differ ($p>0.05$). Comparable inertia characteristics were found for CorTemp (25±4 sec), e-

52 Celsius (21±13 sec) and myTemp (19±2 sec), while the VitalSense system responded more

53 slowly (39±6 sec) to changes in water bath temperature ($p<0.001$).

54 **Conclusion.** Although differences in temperature and inertia were observed between capsule

55 systems, an excellent validity, test-retest reliability, and inertia was found for each system

56 between 36°C and 44°C after removal of outliers.

57 **Key words:** Core body temperature, gastrointestinal temperature, thermoregulation,

58 thermometer

INTRODUCTION

Major sport events are increasingly organized in extreme environmental conditions, making it more important for athletes to perform well in hot and cold ambient conditions and to monitor their core body temperature from a safety perspective (Tc). Exercise-induced increases in metabolic heat production[1, 2] are known to induce a major physiological challenge to the thermoregulatory system[1, 3]. A disbalance between heat production and heat loss causes the core body temperature (Tc) to rise, which may lead to the development of exertional hyperthermia (Tc>40°C), heat related illnesses (i.e. heat exhaustion/heat stroke) and/or a reduction of athletic performance[2, 4, 5]. Alternatively, exercise in cold environments could lead to rapid heat loss due to conduction (water), convection (wind) and radiation, which may contribute to the development of hypothermia[6]. Hence, accurate assessment of an athlete's Tc is important to assess the presence and magnitude of thermoregulatory strain and to select and apply appropriate cooling or heating techniques for preservation of health and exercise performance[7-9].

The gastrointestinal temperature, measured with ingestible temperature capsules, has been established as a valid surrogate marker for Tc[10-12]. Temperature capsule systems are wireless, relatively non-invasive and easily applicable in field based conditions. Although the validity of these temperature capsule systems have been examined[11, 13, 14], different study designs were applied and a substantial variation in accuracy was found (i.e. -0.001-0.27°C). Hence, it is essential to determine which capsule system is superior for assessment of Tc in field conditions.

The aim of this study was to examine the validity, reliability and inertia characteristics of four commercially available ingestible telemetric temperature capsule systems (i.e. CorTemp, e-Celsius, myTemp and VitalSense) in well controlled ex-vivo circumstances using a water bath. Data from this study provide insight in which telemetric capsule system has the

84   most favorable characteristics for Tc assessment, which could enable researchers and trainers

85   to select the best temperature sensor for their scientific study and/or daily practice.

86

87   **METHODS**

88   **Experimental design**

89   Four different ingestible telemetric temperature capsule systems (CorTemp, e-Celsius,

90   myTemp and VitalSense) were tested in a custom made accurately controlled water bath. The

91   primary outcomes were the validity, test-retest reliability and inertia characteristics of the

92   capsule systems. A total of 10 temperature capsules from a single production batch of each

93   telemetry system were tested during three separate trials. The first and second trial consisted of

94   a similar study protocol and was used to assess the validity and test-retest reliability. The third

95   trial adopted a different protocol and was used to examine the inertia characteristics of the

96   temperature capsules. To reduce any bias caused by environmental factors and to ensure that

97   the capsule systems were evaluated in comparable circumstances, a single temperature capsule

98   for each capsule system was used simultaneously in each trial.

99

100   **Experimental Setup**

101   An overview of the experimental setup is presented in Supplementary Figure 1 (SDC 1,

102   Overview of the experimental setup) . A thermostat-controlled and distilled water-filled bath

103   (3.5 L) was used in which four highly sensitive and calibrated wired temperature probes (1529

104   Chube E-4 Thermometer Readout Thermistor, Fluke Hart Scientific, Everett, USA) measured

105   temperature up to 0.00035°C exactly. The average value of these wired temperature sensors

106   represented the temperature of the water bath. In addition, a heater (Fluke Hart Scientific 2100

107   Temperature Controller, Everett, USA) and stirrer (Heidolph Instruments D91126, type RZR1,

108   Schwabach, Germany) system ensured thermal homogeneity of the water bath. A custom made

109 holder prevented the sensor reaching the bottom of the water bath or coming into contact with

110 another sensor. The external monitors of each of the telemetric capsule systems were placed

111 around the water bath within a distance range of 0.2 m.

112

113 **Study protocol**

114 Prior to each experiment, the sensors and external monitors were synchronized to ensure that

115 the measurements occurred simultaneously. In the validity and reliability measurements the

116 water bath temperature gradually increased from 33°C to 44°C, exceeding the physiological

117 range between hypothermia (<35°C) and exertional hyperthermia (>40°C). An automated

118 protocol was programmed to induce a stepwise increase in water bath temperature, resulting in

119 twelve temperature plateaus (33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43 and 44°C). For each

120 temperature plateau, three conditions had to be achieved before the protocol could proceed: 1)

121 water bath temperature did not vary >0.02°C during fifty consecutive measurements (5

122 minutes), 2) the average value of the four independent probes did not vary >0.01°C during two

123 consecutive measurements, and 3) the change in heater power did not exceed 8% during two

124 consecutive measurements. These conditions ensured stability of the water bath temperature

125 and thereby reliable temperature measurements at each point of measurement. The study

126 protocol was performed twice for each temperature capsule (Trial 1/Trial 2), which allowed us

127 to calculate the validity and test-retest reliability. The water bath temperature was measured

128 every 6 seconds.

129   In the inertia experiment the water bath temperature gradually increased from 36°C to

130 42°C. At every temperature threshold (36, 37, 38, 39, 40, 41 and 42°C) the water bath

131 temperature was stabilized for five minutes. Then, the water bath temperature increased by 1°C

132 in a timeframe of five minutes. This timeframe was constructed to mimic the increase in Tc

133 during high intensity exercise in hot ambient conditions, if no heat can be removed from the

134   body(2). This study protocol allowed us to calculate the time delay of the temperature measured

135   by the temperature capsule compared to the actual temperature of the water bath during the

136   stepwise heating phase. This time delay is defined as the inertia of the temperature capsule.

137

138   **Telemetric temperature capsule systems**

139   Characteristics of the ingestible telemetric temperature capsule systems are shown in Table 1.

140   All capsule systems used an external wireless recorder to receive the signal from the

141   temperature capsule via a specific radio frequency. The temperature capsules of CorTemp (HQ

142   Inc., Florida, USA), e-Celsius (BodyCap, Caen, France) and VitalSense (Philips Respironics,

143   Bend, Oregon, USA) were delivered in standby modus and had to be activated before use. The

144   myTemp (myTemp, Nijmegen, Netherlands) capsule is automatically activated by the external

145   recorder, which is also the power supply for the temperature capsule. All temperature capsules

146   were activated directly prior to Trial 1. Furthermore, all measurements were performed in

147   accordance with the manual of the individual capsule systems and the highest sample frequency

148   was used throughout the protocol. The external recorders of all capsule systems stored the data,

149   which were exported to a computer for further analysis using the latest version of available

150   software.

151   **Data processing and Statistical Analysis**

152   The average capsule temperature during the final 150 seconds of each temperature

153   plateau was calculated per telemetric system. Due to differences in sample rate, capsule

154   temperature reflected the average of n=25 consecutive measurements for myTemp, n=15 for

155   CorTemp, n=6 for e-Celsius, and n=10 for VitalSense. Average capsule temperature and water

156   bath temperature were compared for each temperature plateau (33-44ºC). Outliers were defined

157   as observations with a difference >1°C between consecutive measurements and were excluded

158   from further analysis. Furthermore, we addressed the number of measurements with a

159 difference between consecutive data points between 0.2°C and 1.0°C to get more insight into

160 the consistency of the data.

161   In order to establish the validity, the Bland-Altman method for assessing the agreement

162 between two methods was used(15). In short, the mean difference (=systematic bias) between

163 the temperature capsule and water bath was assessed using a one-sample T-test. The systematic

164 bias and accompanying 95% Limits of Agreement (LOA) were derived from the Bland-Altman

165 plot(15). Furthermore, the Intraclass Correlation Coefficient (ICC) was calculated for the

166 average of all 10 capsules, to determine the inter-measure agreement(16). The Standard Error

167 of Measurement (SEM) was calculated based on the standard deviation (SD) of the difference

168 between temperature capsules and water bath temperature(17). Furthermore, we conducted a

169 Repeated Measures ANOVA to determine whether the accuracy of the capsule systems was

170 different across temperature plateaus (i.e. 33-44ºC). Differences in accuracy across capsule

171 systems were examined using one-way ANOVA. A similar approach was used to determine the

172 test-retest reliability.

173   Inertia was assessed as the time delay of the telemetric capsule to reach the same

174 temperature as the water bath after a sudden temperature increase. Inertia was determined at

175 50% (P50) and 90% (P90) of the increase to each temperature plateau, and the time at which

176 the first observation of the capsule and the water bath exceeded the P50 or P90 temperature was

177 taken. Subsequently, the time to reach P50 and P90 of the capsule system was compared with

178 the time of the water bath to reach P50 and P90, and was defined as the time delay (inertia). As

179 the time delay may be influenced by the accuracy and sample frequency of the capsule, we

180 applied two different correction methods: 1) the systematic bias of the telemetric capsule (i.e.

181 sensitivity data) was subtracted from the recorded values, 2) temperature data was interpolated

182 between subsequent samples to determine the exact time at which P50 and P90 were exceeded.

183 Inertia characteristics were presented as: I) raw data, II) corrected for differences in accuracy,

184   and III) corrected for differences in accuracy and sample frequency. To examine whether there

185   was an inertia difference per temperature plateau across telemetric capsule systems, a two-way

186   repeated measures ANOVA was performed. One-way ANOVA was used to assess the

187   differences in inertia characteristics at P50 and P90 between the four telemetric capsule

188   systems. Furthermore, time constants of the systems response were determined by exposing a

189   single capsule three times to a step change in temperature between two water baths of 7°C (30

190   – 37°C). Differences in the systems sampling rates did not allow a very precise determination,

191   however by interpolation of the data the time constants can be determined.

192   All statistical analyses were performed using SPSS Statistics (Version 20), in which the

193   level of significance was set at $p < 0.05$. The systematic bias was reported as mean difference ±

194   SD, unless indicated otherwise.

195

196   **RESULTS**

197   *Missing data and outliers.* A total of 40 temperature capsules were investigated: 10 sensors per

198   telemetric capsule system. We experienced difficulties with the activation of n=4 VitalSense

199   telemetric capsules, although the provided instructions were carefully followed. Moreover, n=1

200   of these VitalSense temperature capsules could not be activated at all and 1 temperature capsule

201   stopped measuring after 43ºC during Trial 2, meaning that data of the 44°C temperature plateau

202   of 44°C is not reported for that temperature capsule. As a result, data from 39 temperature

203   capsules was used for our analyses.

204   In n=6 from n=9 VitalSense temperature capsules, data was randomly missed

205   throughout the protocol (Trial 1 + 2), representing 1.0% of the total data. n=2 CorTemp capsules

206   and n=1 e-Celsius capsule randomly missed 0.1% of the data, whereas no missing data was

207   reported for the myTemp system (Supplementary Table 1, SDC 2, Missing data and outliers).

208   The CorTemp system appeared to be the only system with outliers ($\Delta T_{capsule} > 1$ºC), which was

209   randomly present in 4.0% of the total data, ranging from a difference of 1°C to 62.1°C.

210   CorTemp also showed error measurements ($0.2°C < \Delta T_{capsule} < 1°C$) in 4.4% of the total data,

211   whereas these error measurements were not present in the other systems. Outliers and error

212   measurements were both found in all CorTemp capsules.

213   *Validity.* After exclusion of outliers, mean differences between capsule and water bath

214   temperature for Trial 1 were 0.077±0.040°C (CorTemp), -0.081±0.055°C (e-Celsius), -

215   0.003±0.006°C (myTemp) and -0.017±0.023°C (VitalSense) (Figure 1), which were

216   significantly different from zero (all p-values≤0.01). Additionally, the myTemp system

217   demonstrated the smallest mean difference, followed by VitalSense, CorTemp and e-Celsius

218   ($p_{capsule\ system}$<0.001). The 95% LOA were ±0.079°C (CorTemp), ±0.108°C (e-Celsius),

219   ±0.013°C (myTemp) and ±0.046°C (VitalSense). The SEM was 0.028°C for CorTemp,

220   0.039°C for e-Celsius, 0.005°C for myTemp and 0.017°C for the VitalSense system. All capsule

221   systems demonstrated an excellent agreement between capsule and water bath temperature

222   based on the significant ICC of 1.00 (all p-values <0.05). The data of Trial 2 revealed similar

223   outcomes with respect to the mean differences, LOA, SEM and ICC (Table 2). A repeated-

224   measures ANOVA indicated that the mean difference between the e-Celsius, myTemp and

225   VitalSense system and water bath temperature did not drift across temperature plateaus

226   (p<0.05). In contrast, a significant decrease in mean difference was found across increasing

227   water bath temperatures for the CorTemp system (p=0.002, Figure 2).

228   *Test-retest reliability.* Mean difference between Trial 1 and Trial 2 appeared to be significantly

229   different from zero for CorTemp (0.017±0.083°C, LOA= ±0.162°C, p=0.030) and e-Celsius (-

230   0.007±0.033°C, LOA= ±0.064°C p=0.019) (Figure 3). For myTemp (0.0001±0.008°C, LOA=

231   ±0.016°C) and VitalSense (0.002±0.014°C, LOA= ±0.028°C) the mean difference did not differ

232   significantly from zero (both p-values>0.05). Furthermore, the CorTemp system demonstrated

233   the highest mean difference between Trial 1 and Trial 2 (p=0.001), whereas the other systems

234   had a comparable mean difference between both trials ($p > 0.05$). The SEM was 0.058°C for

235   CorTemp, 0.023°C for e-Celsius, 0.006°C for myTemp and 0.010°C for the VitalSense system.

236   An excellent agreement between Trial 1 and Trial 2 was found for all four capsule systems

237   (ICC=1.00, $p < 0.05$).

238   *Inertia.* Inertia characteristics are summarized in Table 3. The raw data revealed that the

239   CorTemp system had a significant lower time delay to reach p50 (9±5 seconds) and p90 (10±5

240   seconds) compared to the other capsule systems, whereas the VitalSense system demonstrated

241   the slowest response (p50= 54±12 seconds, p90= 35±3 seconds; $p < 0.001$). After correction for

242   the systematic bias of each capsule system, the myTemp system demonstrated the lowest p50

243   and p90, followed by the CorTemp and e-Celsius system. The p50 and p90 remained the highest

244   for the VitalSense system ($p < 0.001$). Additional correction for sample frequency did not alter

245   inertia characteristics (Table 3). Time constants of the systems response were 22 seconds for

246   myTemp, 28 seconds for e-Celsius, 47 seconds for CorTemp and 48 seconds for VitalSense.

247

248   **DISCUSSION**

249   This is the first study to compare the validity, reliability and inertia characteristics of all

250   commercially available ingestible telemetric temperature capsule systems. Our well controlled

251   ex-vivo water bath study demonstrates that all temperature capsule systems, are valid and

252   reliable to measure (water) temperature, evidenced by their small systematic biases and a low

253   LOA and SEM after removal of outliers (CorTemp). Furthermore, we found that the CorTemp,

254   e-Celsius and myTemp capsule system demonstrated comparable inertia characteristics,

255   whereas the VitalSense system demonstrated a lower responsiveness to changes in water bath

256   temperature. These findings enable researchers and clinicians to select the telemetric capsule

257   system that best suits their goal, which can improve the safety aspect of doing exercise in a hot

258   and cold environment.

259   An excellent validity and reliability of a temperature measurement technique is

260   characterized by a 1) low systematic bias (<0.1°C), 2) narrow 95% LOA (maximal ±0.4°C), 3)

261   high ICC (>0.80) with the reference temperature, and 4) low SEM(10, 13, 18). We found a

262   significant systematic bias for all four capsule systems, but the validity and reliability of every

263   capsule system complied with reference criteria for an excellent acceptable level of agreement.

264   Nevertheless, we observed a substantial prevelence of outliers in our raw CorTemp data (4.0%),

265   leading to a high LOA (2.3°C) and violation of accuracy criteria (<0.1°C). Data verification

266   and cleaning are, therefore, needed before CorTemp data can be used appropriately.

267   Furthermore, the decreasing systematic bias with increasing temperatures suggests that the

268   CorTemp system is mainly accurate in normothermic and hyperthermic conditions (36-44°C),

269   but less accurate for hypothermic conditions (33-35°C). Although, the CorTemp system did not

270   met the criteria for an excellent validity for hypothermic conditions, the systematic bias (0.1 -

271   0.2°C) is still physiologically acceptable. e-Celsius, myTemp and VitalSense were more

272   constant and performed well across the whole temperature range. Furthermore, the intraclass

273   correlation coefficient (ICC) and the standard error of measurement (SEM) were used to assess

274   the reliability(17, 18). An ICC of 1.00 was found for all capsule systems, whereas an ICC of

275   >0.80 is typically considerd as acceptable, with higher values respresenting a better

276   reliability(18). The high ICC of the four capsule systems suggests that the error variance

277   between water bath and capsule temperature and between Trial 1 and Trial 2 are negligible

278   compared to the normal variance of the measurement(19). Addionally, the low SEM for all

279   capsule systems is another indication that there is an excellent agreement between water bath

280   and capsule temperature and between Trial 1 and Trial 2. Therefore, all capsule systems are

281   valid and reliable methods to measure temperature after outliers have been removed.

282   The responsiveness of the temperature capsules was quantified by the inertia

283   characteristics at p50 and p90. We found that the VitalSense system had the slowest response

284 (38-39 seconds) to acute changes in temperature compared to the other systems (range: 18-26

285 seconds). Nevertheless, all systems demonstrated an acceptable responsiveness to changes in

286 temperature. A previous study reported a maximal Tc increase of 1°C per 5 minutes if no heat

287 can be removed from the body(2). An inertia of 18 to 39 seconds is, therefore, physiologically

288 irrelevant. Moreover, the underestimation of Tc measured with a temperature capsule in

289 dynamic and/or quick changing situations is marginal and hardly influences final Tc.

290 Furthermore, the order of the results of the time constants matches the results of the p50 and

291 p90 times corrected for sample frequency. The observed time constants are considered

292 appropriate for the physiological signals measured.

293 Even though the results of our study may be promising, practical considerations must

294 be taken into account. First, the activation of the VitalSense temperature capsules was hard and

295 one of the capsules (10%) could not be activated at all. Anecdotal evidence from our research

296 groups and our collaborators, confirm the infrequent non-activation problem of VitalSense

297 capsules in other studies, whereas similar problems were occasionally experienced for

298 CorTemp capsules. The sample frequency is also an important distinction between the capsule

299 systems, since the sample frequency can be adjusted for CorTemp and myTemp, while it is

300 fixed and relatively low frequent for e-Celsius and VitalSense. Furthermore, 4% of the raw

301 CorTemp data consisted of outliers (>1°C) and another 4.4% of error measurements (0.2-

302 1.0°C). The CorTemp system is therefore less consistent and the use of the raw data with large

303 intervals between measurements might result in inaccurate values. Finally, the present study

304 used capsules from a single production batch from each capsule system, which limited us to

305 assess batch differences within capsule systems.

306 For human use, other aspects than the investigated accuracy, test-retest reliability and

307 inertia, also play a role. Tc is the result of the local thermal balance affected by tissue properties

308 and local blood flow(20). Studies comparing different measurement location in the digestive

309 system showed that absolute temperatures and inertia differ between locations(21, 22).

310 Moreover, the esophageal temperature is ~0.2°C lower during moderate intensity exercise

311 compared to both the gastrointestinal and rectal temperature(21). Additionally, the response

312 time of the esophageal temperature is faster than the gastrointestinal temperature, which in turn

313 was faster than the rectal temperature(21). Ideally, the capsule should be located in the

314 gastrointestinal tract and not in the stomach, which can be achieved by timely swallowing the

315 capsule(12, 23).

316   In conclusion, significant but small differences were observed across telemetric

317 temperature capsule systems. CorTemp demonstrated outliers and error measurements in 4.0%

318 of the recorded data, while this was virtually absent in all other systems. Nevertheless, an

319 excellent validity and test-retest reliability was found for all systems after removal of outliers.

320 The best test-retest reliability was found for the myTemp and VitalSense system, whereas

321 CorTemp and e-Celsius demonstrated a small, but negligible, systematic difference between

322 Trial 1 and Trial 2. Furthermore, the VitalSense system showed the slowest response to

323 increases in water bath temperature, while the other systems had a comparable time delay.

324

332

REFERENCES

1. Kenefick RW, Cheuvront SN, Sawka MN. Thermoregulatory function during the marathon. *Sports Med*. 2007;37(4-5):312-5.
2. American College of Sports M, Armstrong LE, Casa DJ et al. American College of Sports Medicine position stand. Exertional heat illness during training and competition. *Med Sci Sports Exerc*. 2007;39(3):556-72.
3. Hargreaves M. Physiological limits to exercise performance in the heat. *J Sci Med Sport*. 2008;11(1):66-71.
4. Gonzalez-Alonso J, Teller C, Andersen SL, Jensen FB, Hyldig T, Nielsen B. Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *J Appl Physiol (1985)*. 1999;86(3):1032-9.
5. Tatterson AJ, Hahn AG, Martin DT, Febbraio MA. Effects of heat stress on physiological responses and exercise performance in elite cyclists. *J Sci Med Sport*. 2000;3(2):186-93.
6. Burtscher M, Kofler P, Gatterer H et al. Effects of Lightweight Outdoor Clothing on the Prevention of Hypothermia During Low-Intensity Exercise in the Cold. *Clin J Sport Med*. 2012;22(6):505-7.
7. Casa DJ, DeMartini JK, Bergeron MF et al. National Athletic Trainers' Association Position Statement: Exertional Heat Illnesses. *J Athl Train*. 2015;50(9):986-1000.
8. Bongers CC, Hopman MT, Eijsvogels TM. Cooling interventions for athletes: An overview of effectiveness, physiological mechanisms, and practical considerations. *Temperature (Austin)*. 2017;4(1):60-78.
9. Bongers CC, Thijssen DH, Veltmeijer MT, Hopman MT, Eijsvogels TM. Precooling and percooling (cooling during exercise) both improve performance in the heat: a meta-analytical review. *Br J Sports Med*. 2015;49(6):377-84.
10. Byrne C, Lim CL. The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *Br J Sports Med*. 2007;41(3):126-33.
11. Challis GG, Kolb JC. Agreement Between an Ingestible Telemetric Sensor System and a Mercury Thermometer Before and After Linear Regression Correction. *Clin J Sport Med*. 2010;20(1):53-7.
12. Gant N, Atkinson G, Williams C. The validity and reliability of intestinal temperature during intermittent running. *Med Sci Sports Exerc*. 2006;38(11):1926-31.
13. Travers GJ, Nichols DS, Farooq A, Racinais S, Periard JD. Validation of an ingestible temperature data logging and telemetry system during exercise in the heat. *Temperature (Austin)*. 2016;3(2):208-19.
14. Bongers C, Hopman MTE, Eijsvogels TMH. Validity and reliability of the myTemp ingestible temperature capsule. *J Sci Med Sport*. 2017.
15. Bland JM, Altman DG. Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement. *Lancet*. 1986;1(8476):307-10.
16. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-8.
17. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30(1):1-15.
18. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998;26(4):217-38.
19. de Vet HCWT MC, Knol LB. *Measurement in Medicine*. 3rd ed. New York: Cambridge University Press; 2011.
20. Taylor NA, Tipton MJ, Kenny GP. Considerations for the measurement of core, skin and mean body temperatures. *J Therm Biol*. 2014;46:72-101.

381   21.   Mundel T, Carter JM, Wilkinson DM, Jones DA. A comparison of rectal, oesophageal and gastro-
382   intestinal tract temperatures during moderate-intensity cycling in temperate and hot
383   conditions. *Clin Physiol Funct Imaging*. 2016;36(1):11-6.
384   22.   Teunissen LP, de Haan A, de Koning JJ, Daanen HA. Telemetry pill versus rectal and esophageal
385   temperature during extreme rates of exercise-induced core temperature change. *Physiol
386   Meas*. 2012;33(6):915-24.
387   23.   Bongers CC, Hopman MT, Eijsvogels TM. Using an Ingestible Telemetric Temperature Pill to
388   Assess Gastrointestinal Temperature During Exercise. *J Vis Exp*. 2015;(104).
389   24.   Bongers CCWGH, M.T.E.; Eijsvogels, T.M.H. . Validity and reliability of the myTemp ingestible
390   temperature pill. *Unpublished data*. 2017.

391

392

393 **FIGURE LEGENDS**

394 **Figure 1.** Raw data **(A)** and data after outlier removal **(B)** mean difference between temperature

395 capsule and water bath temperature for the capsule systems. Data were presented as mean

396 difference ± LOA. * indicates a significant systematic bias.

397

398 **Figure 2.** An overview of the mean difference between capsule and water bath temperature for

399 the twelve discrete temperature plateaus. A separate line was plotted for each temperature

400 capsule system. Data were presented as mean difference ± SD, and * represents a drifted

401 response over the temperature plateaus.

402

403 **Figure 3.** Raw data **(A)** and data after outlier removal **(B)** mean difference between

404 temperatures measured during Trial 1 and Trial 2 for the capsule systems. Data were presented

405 as mean difference ± LOA. * indicates a significant systematic bias.

406

407   **SUPPLEMENTAL FILES**

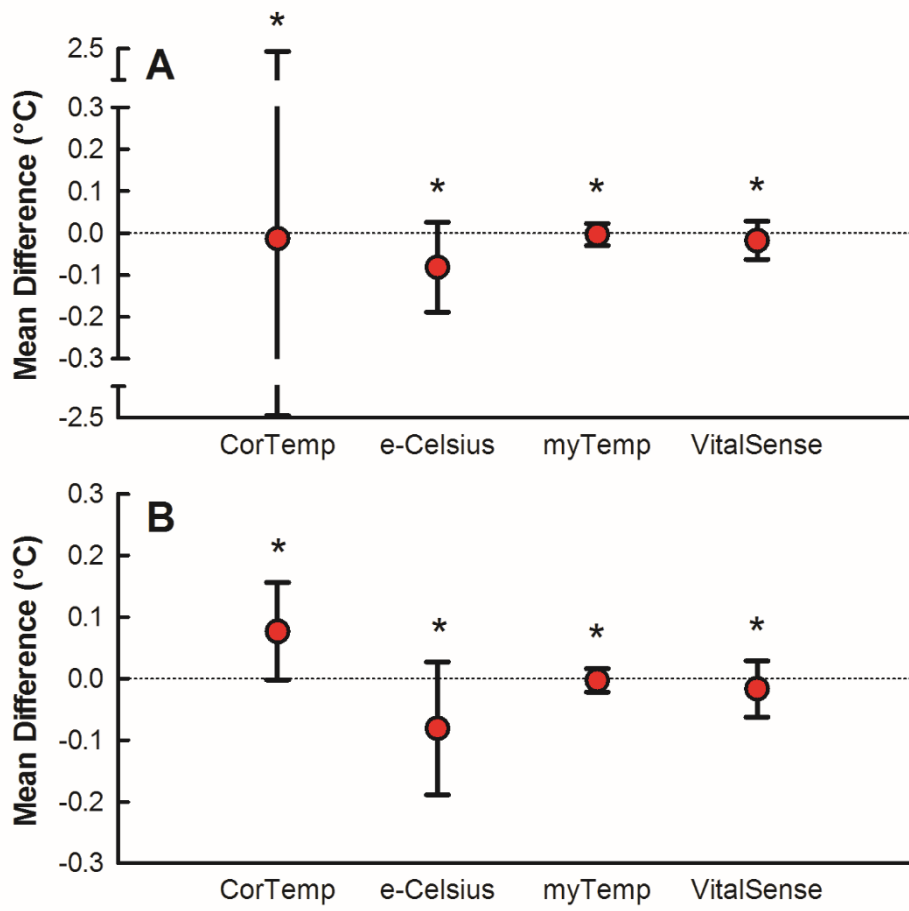408   **Supplementary Table 1.** Missing data and outliers (Supplementary Table 1.doc)

409   **Supplementary Figure 1.** Overview of the experimental setup (Supplementary Figure 1.tiff)
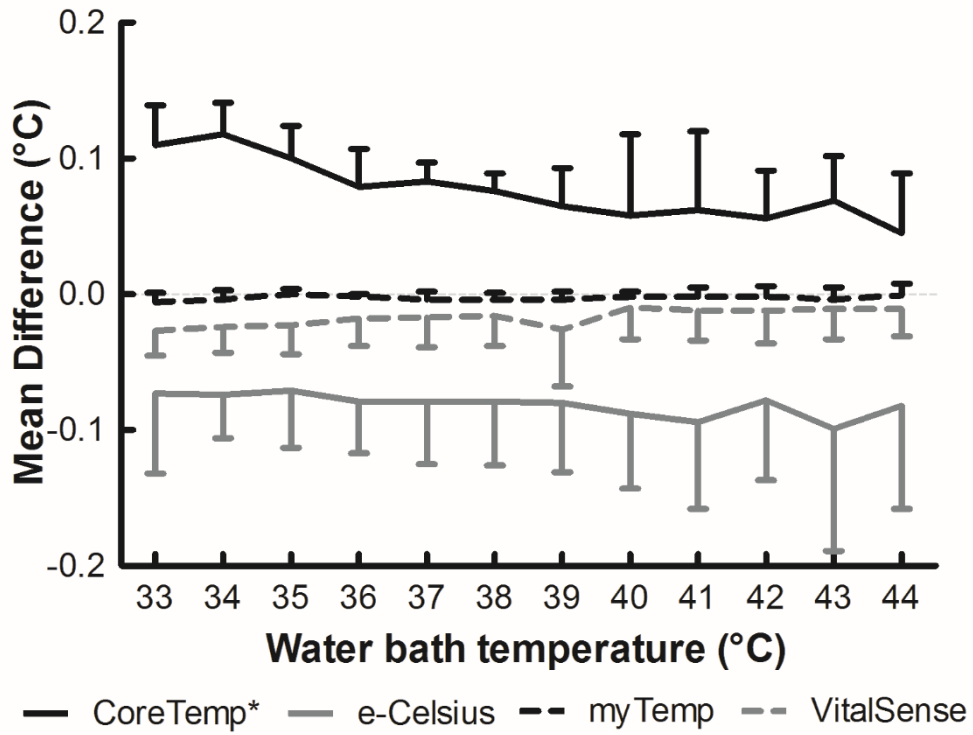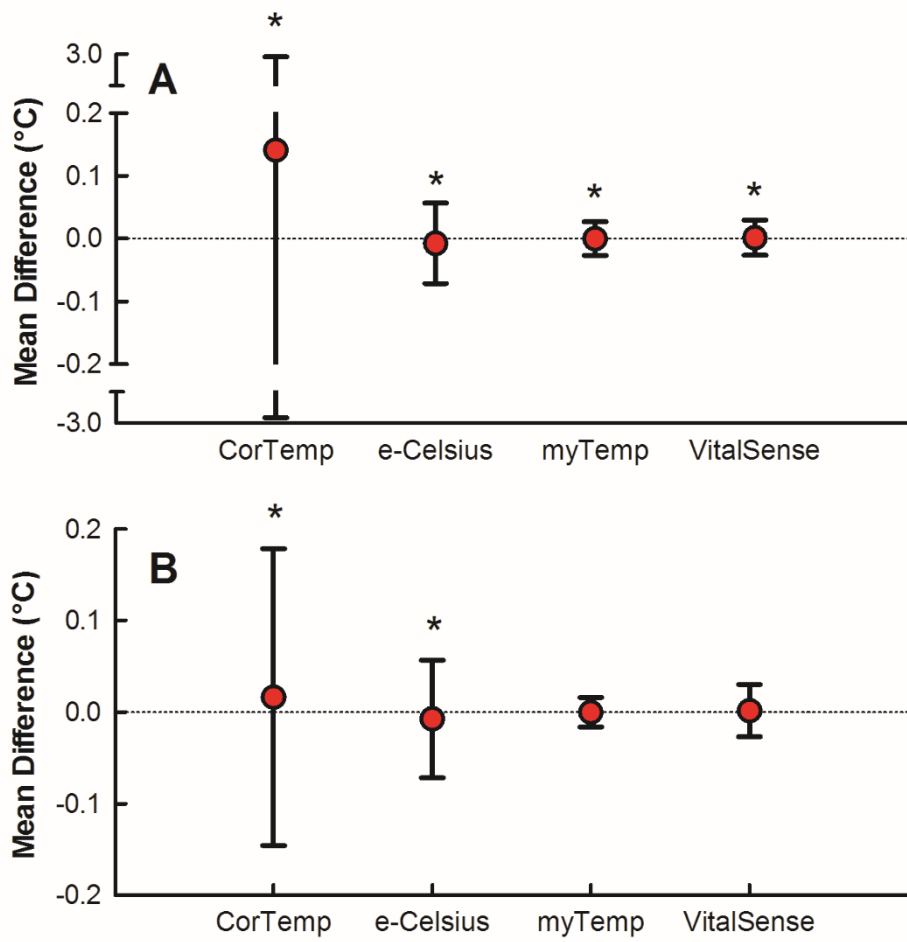
410

411   **Figure 1.**

412

413

414   **Figure 2.**

Mean Difference (°C) vs. Water bath temperature (°C). Legend: CoreTemp*, e-Celsius, myTemp, VitalSense.

415

416

417

418 **Figure 3.**

419

420

421 **Table 1.**

Table 1. Physical and technical characteristics of the telemetric capsule systems

	CorTemp	e-Celsius	myTemp	VitalSense
Capsule characteristics				
Length (mm)	22.4	17.7	20.0	23.0
Diameter (mm)	10.9	8.9	8.0	8.7
Weight (g)	2.8	1.7	1.3	1.5
Operating range (ºC)	30 to 45	0 to 50	30 to 45	-10 to 60
Accuracy (ºC)	0.27(11)	0.23(13)	0.001(24)	0.17(13)
Battery lifetime	7-10 days	20 days	Infinite	10 days
Power supply	Silver-oxide battery	Zinc-silver oxide battery	Self-induction	Battery
Sample frequency	Adjustable	Fixed	Adjustable	Fixed
Lowest sample rate (sec)	10	~30	6	~15
Software version	CorTrack II	e-Performance manager (v01.01.00.0C)	myTemp Manager (v01.08)	Equivital Manager (v1.2.39.4600)

422

423

424 **Table 2.**

Table 2. Validity of the four temperature capsule systems

		CorTemp	e-Celsius	myTemp	VitalSense
Trial 1	**ICC – raw data**	0.94	1.00	1.00	1.00
	ICC – after outlier removal	1.00	1.00	1.00	1.00
	SEM (°C) – raw data	0.836	0.039	0.005	0.017
	SEM (°C) – after outlier removal	0.028	0.039	0.005	0.017
Trial 2	**MD (°C) – raw data**	-0.154	-0.073	-0.002	-0.018
	MD (°C) – after outlier removal	0.061	-0.073	-0.002	-0.018
	LOA (°C) – raw data	1.466	0.105	0.013	0.037
	LOA (°C) – after outlier removal	0.167	0.105	0.013	0.037
	ICC – raw data	0.98	1.00	1.00	1.00
	ICC – after outlier removal	1.00	1.00	1.00	1.00
	SEM (°C) - raw data	0.529	0.038	0.005	0.013
	SEM (°C) - after outlier removal	0.060	0.038	0.005	0.013

425 **ICC=** Intraclass Correlation Coefficient, **SEM=** Standard Error of the Measurement, **MD=** Mean
426 Difference, **LOA=** Limits of Agreement.

427

428   **Table 3.**

Table 3. Inertia characteristics of the four temperature capsule systems.

		CorTemp	e-Celsius	myTemp	VitalSense	p-value
Raw data	**p50 (s)**	9±5 [b,c,d]	41±17 [a,c]	23±2 [a,b,d]	54±12 [a,c]	**<0.001**
	p90 (s)	10±5 [b,c,d]	27±9 [a,d]	23±3 [a,d]	35±3 [a,b,c]	**<0.001**
Correction I	**p50 (s)**	28±8 [d]	33±12 [c]	22±2 [b,d]	44±7 [a,c]	**<0.001**
(accuracy)	**p90 (s)**	30±6 [d]	33±11 [c,d]	21±1 [b,d]	45±8 [a,b,c]	**<0.001**
Correction II	**p50 (s)**	25±4 [d]	21±13 [d]	19±2 [d]	39±6 [a,b,c]	**<0.001**
(accuracy + sample frequency)	**p90 (s)**	26±7 [d]	21±9 [d]	18±1 [d]	38±9 [a,b,c]	**<0.001**

Data were presented as the delay of capsule systems to reach p50 and p90 compared to the water bath. [a] represents significantly different from CorTemp, [b] different from e-Celsius, [c] different from myTemp and [d] different from VitalSense.

429

430

431 **Supplementary Figure 1.**

Supplementary Figure 1. Overview of the experimental setup
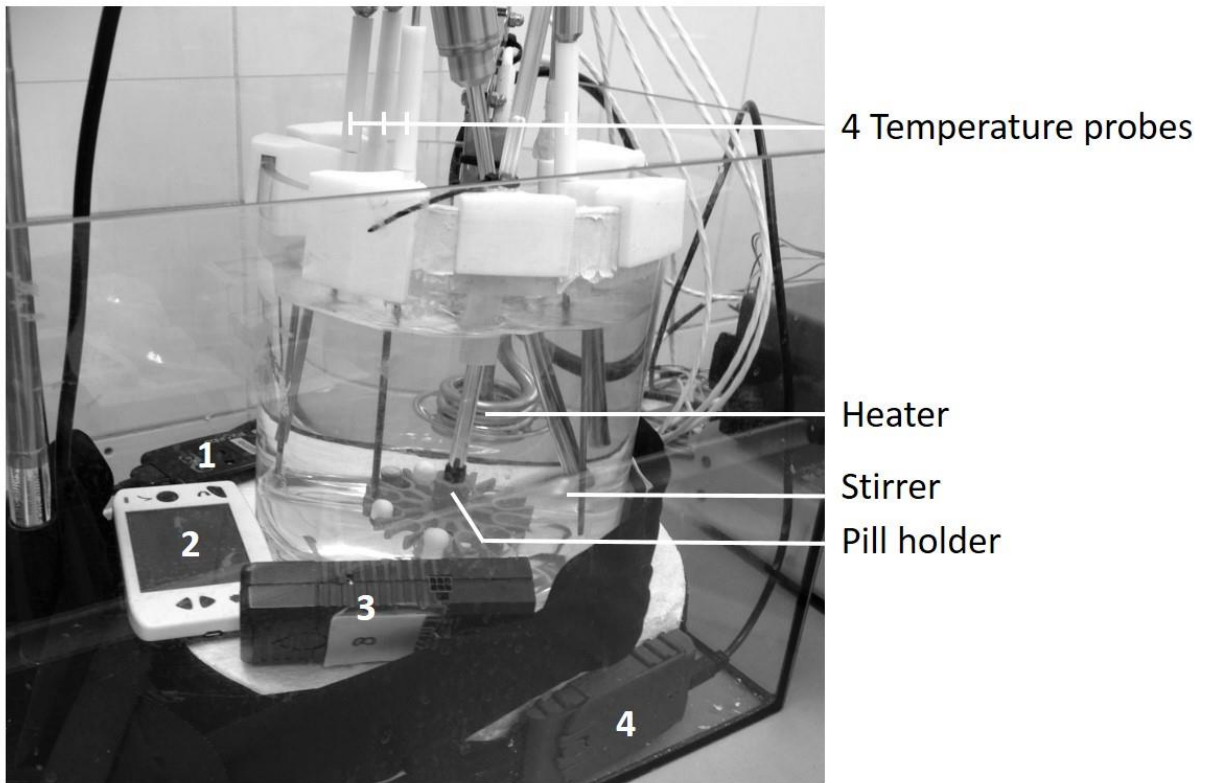
1= VitalSense recorder   2= e-Celsius recorder
3= CorTemp recorder   4= myTemp recorder

432

433

434 **Supplementary Table 1.**

Supplementary Table 1. Missing data and outliers

		CorTemp	e-Celsius	myTemp	VitalSense
Trial 1	**Missing data**	0.1%	0%	0%	0.4%
	Outliers > 1ºC	3.1%	0%	0%	0.1%
	Error measurements 0.2ºC < ΔT$_{capsule}$ < 1ºC	4.1%	0%	0%	0%
Trial 2	**Missing data**	0.1%	0.3%	0%	1.5%
	Outliers > 1ºC	4.9%	0%	0%	0.3%
	Error measurements 0.2ºC < ΔT$_{capsule}$ < 1ºC	4.7%	0%	0%	0%

435