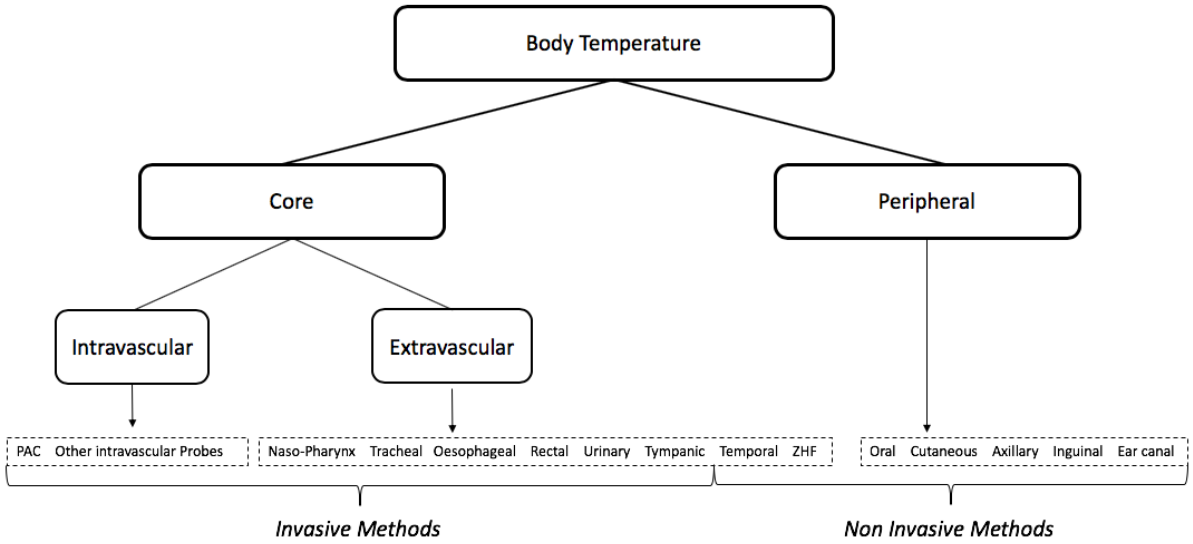
Appendix 1. Randomized trials investigating body temperature in different clinical conditions.

Main Pathology	Year	Journal	Authors	Body temperature		Body temperature measurement method
				Intervention, °C	*Control, °C*	
Cardiac Arrest	2002	New Engl J Med	Bernard et al.	33	37	Tympanic or bladder route, until a pulmonary-artery catheter was placed.
Cardiac Arrest	2002	New Engl J Med	The hypothermia after cardiac arrest study group	32 to 34	Normothermia	Infrared tympanic on admission, urinary bladder-temperature probe for further measurements.
Cardiac Arrest	2013	New Engl J Med	Nielsen et al.	33	36	Urinary bladder-temperature probe, or esophageal or intravascular probe in patients with low urinary output.
Cardiac Arrest	2019	New Engl J Med	Lascarrou et al.	33	37	Urinary bladder-temperature probe, an oesophageal probe, or another central probe (pulmonary arterial catheter, for example, if present)
Traumatic Brain Injury	2011	Lancet Neurol	Clifton et al.	35	37	Urinary bladder-temperature probe
Traumatic Brain Injury	2015	New England J Medicine	Andrews et al	32 to 35	Standard Care (within 37°)	Urinary bladder, esophageal, rectal-temperature probe, PA Catheter.
Traumatic Brain Injury	2019	JAMA	Cooper et al.	33 to 35	37 ± 0.5	Urinary bladder-temperature probe (preferred) or esophageal temperature
Convulsive Status Epilepticus	2016	New England J Medicine	Legriel et al	32 to 34	Standard Care (within 37°)	Esophageal probe
Severe Bacterial Meningitis	2013	JAMA	Mourvillier et al.	32-34	Standard Care (within 37°)	Esophageal probe
Infection	2015	New Engl J Med	Young et al.	NA	NA	Axillary thermometer
Fever	2019	Intensive Care Medicine	Young et al.	NA	NA	Continuous monitoring of core body temperature in mechanically ventilated. Tympanic thermometer was the preferred method when core temperature monitoring was deemed not appropriate by the treating physician. Temporal artery and axillary thermometers use was discouraged because of concerns about the accuracy of these devices.

This appendix was part of the submitted manuscript and has been peer reviewed. It is posted as supplied by the authors.

Appendix 2. Body temperature measurement methods.



Abbreviations: PAC, Pulmonary Artery Catheter; ZHF, Zero Heat Flux.

Appendix 3. Search strategies via Ovid- MEDLINE, Ovid-EMBASE and Cochrane Central.

Ovid – MEDLINE

1. exp Intensive Care Units/ or Intensive care units.mp/ or intensive care.mp.
2. exp Critical Care/ or Critical care.mp.
3. exp Critical Illness/ or Critical illness.mp.
4. ICU.mp.
5. 1 OR 2 OR 3 OR 4
6. exp Thermometers/ or Thermometer.mp or Thermometers.mp.
7. exp Thermography/ or Thermography.mp.
8. exp Monitoring, Physiologic/ or Monitoring, physiologic.mp
9. temperature measurement.mp.
10. body temperature measurement.mp.
11. body temperature monitoring.mp.
12. body temperature method*
13. 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
14. exp Body Temperature/ or Body Temperature.mp.
15. core temperature.mp.
16. temperature comparison.mp.
17. peripheral temperature.mp.
18. exp Fever/ or fever.mp.
19. Hyperthermia.mp.
20. normothermia.mp.
21. exp Hypothermia/ or exp Hypothermia, induced/ or hypothermia.mp.
22. 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21
23. 5 AND 13 AND 22
24. limit 23 to (human and English language and (adult <18 to 64 years> or aged <65+ years>))

Ovid – EMBASE

1. exp intensive care unit/ or Intensive care unit/ or ICU.mp.
2. exp intensive care/ or intensive care.mp.
3. exp critical illness/ or critical illness.mp. or critical care.mp.
4. 1 OR 2 OR 3
5. exp Thermometers/ or thermometers.mp.
6. temperature method*.mp.
7. exp temperature measurement/ or temperature measurement.mp.
8. exp body temperature measurement/ or body temperature measurement.mp.
9. body temperature monitoring.mp.
10. exp thermography/ or thermography.mp.
11. exp physiologic monitoring/ or Physiologic monitoring.mp
12. 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. exp body temperature/ or Body temperature.mp.
14. exp core temperature/ or Core temperature.mp.

15. Peripheral temperature.mp.
16. Temperature comparison.mp.
17. exp fever/ or Fever.mp.
18. exp hypothermia/ or Hypothermia.mp.
19. exp induced hypothermia/ or exp induced hypothermia/
20. exp hyperthermia/ or hyperthermia.mp.
21. 9 OR 10 OR 11 OR 12
22. 4 AND 8 AND 13
23. limit 14 to (human and English language and (adult <18 to 64 years> or aged <65+ years>))

Cochrane Central – Mesh terms

1. exp Intensive Care Units/
2. exp Critical Care
3. exp Critical Illness
4. 1 OR 2 OR 3
5. exp Thermometers
6. exp Thermography
7. exp Monitoring, Physiologic
8. 5 OR 6 OR 7
9. exp Body Temperature
10. exp Fever
11. exp Hypothermia
12. exp Hypothermia, induced
13. 9 OR 10 OR 11 OR 12
14. 4 AND 8 AND 13

Appendix 4. QUADAS-2 tool for the assessment of bias and applicability of four specific domains tailored to explore the accuracy of peripheral body temperature measurements compared to Pulmonary Artery Catheter.

Research question: What is the accuracy of non-invasive body temperature measurement compared to pulmonary artery catheter measurement in the assessment of core temperature in ICU patients?

Domains	**Signalling questions**	**Risk of bias**	**Applicability concerns**
Patient Selection	1.Was a consecutive or random sample of patients enrolled? 2.Was a case-control design avoided? 3.Did the study avoid inappropriate exclusions?	Could the selection of patients have introduced bias?	Are there concerns that the included patients and patients do not match the review question?
Index Test	1.Were the index test results interpreted without knowledge of the results of the reference standard? *2. If a threshold was used, was it pre-specified?* *3. Did the investigators controlled for possible confounders[1]?*	Could the conduct or interpretation of the index test have introduced bias?	Are there concerns that the index test, its conduct or interpretation differ from the review question?
Reference Standard	*1.Is the reference standard likely to correctly classify the target condition?* *2.Were the reference standard results interpreted without knowledge of the results of the index test?*	Could the conduct or interpretation of the reference standard have introduced bias?	Are there concerns that the target condition as defined by the reference standard does not match the review question?
Flow and Timing	*1.Was there an appropriate interval between index test and reference standard?* 2. *Did all patients receive the same reference standard?* *3.Were all patients included in the analysis?*	Could the patient flow have introduced bias?	

Table S1. Characteristics of 13 studies meeting inclusion criteria.

Author	Year	Country	Design	Setting	Device	Patients	Readings	Age (years)	Temperature (°C)[#]
Dahyo-Fizelier	2017	France	Prospective	Surgical ICU	Zero Heat Flux (forehead)	54	61,298	54	36.6
					Oesophageal probe				
					Intravascular probe				
Eshranghi	2014	US	Prospective	Cardiac surgery	Zero Heat Flux (forehead)	103	21,787	67	37.0
					Zero Heat Flux (neck)				
					Self-adhesive skin probe				
					Pulmonary artery catheter				
Farnell	2005	UK	Prospective	Mixed ICU	Axillary chemical thermometer	25	153	65	37.2
					Tympanic infrared thermometer				
					Pulmonary artery catheter				
Fulbrook	1997	UK	Prospective	Mixed ICU	Axillary chemical thermometer	60	60	63	38.0
					Tympanic infrared thermometer				
					Pulmonary artery catheter				
Haugk	2010	Austria	Retrospective	Cardiac arrest	Oesophageal probe	21	2,000	61	34.7
					Endotracheal tube probe				
					Pulmonary artery catheter				
Lefrant	2003	France	Prospective	Medical ICU	Oesophageal probe	42	529	73	37.0
					Rectal probe				
					Gallium-in-glass thermometer (axillary, inguinal)				
					Temperature-sensing urinary catheter				

					Pulmonary artery catheter				
Moran	2007	Australia	Prospective	Mixed ICU	Mercury-in-glass thermometer	110	2,118	65	37.4
					Tympanic infrared thermometer				
					Temperature-sensing urinary catheter				
					Pulmonary artery catheter				
Myny	2005	Belgium	Prospective	Medical ICU	Electronic thermometer (axillary)	57	106	60	37.1
					Exergen temporal scanner				
					Pulmonary artery catheter				
Nierman	1991	US	Prospective	Medical ICU	Tympanic infrared thermometer	15	21	78	37.7
					Temperature-sensing urinary catheter				
					Pulmonary artery catheter				
Nonose	2012	Japan	Prospective	Mixed ICU	Electronic thermometer (axillary)	73	1,793	68	36.2
					Tympanic infrared thermometer				
					Temperature-sensing urinary catheter				
					Pulmonary artery catheter				
Shin	2013	Korea	Prospective	Cardiac arrest	Tympanic infrared thermometer	21	1,479	50	35.3
					Rectal probe				
					Temperature-sensing urinary catheter				
					Pulmonary artery catheter				
Smith	2004	US	Prospective	Cardiac surgery	SolarTherm (oral, axillary)	35	35	66	37.0
					Axillary DataTherm				
					Pulmonary artery catheter				
Stavem	1997	Norway	Prospective	Mixed ICU	Tympanic infrared thermometer	16	66	62	37.5
					Oesophageal probe				

	Rectal probe
	Pulmonary Artery Catheter

[#]Mean temperature as measured by the pulmonary arterial catheter.

Abbreviations: ICU, intensive care unit; UK, United Kingdom; US, United States.

This appendix was part of the submitted manuscript and has been peer reviewed. It is posted as supplied by the authors.

Table S2. DerSimonian and Laird random effects meta-analysis of 8 studies reporting the mean temperature difference between thermometer measurements in critically ill patients, excluding studies which did not which did not fully account for repeated measures within individual subjects was performed. Reported values include pooled mean bias (mean index measurement subtracted from mean pulmonary artery catheter measurement), pooled standard deviation, and pooled 95% limits of agreement.

Body temperature measurement methods	Comparisons (n)	Studies (n)	Pooled mean bias (95% CI, I^2)	Pooled SD (95% CI, I^2)	Pooled 95% limits of agreement
PAC - urinary bladder	2	2	-0.02 (-0.06 to 0.01)	0.23 (0.17 to 0.28)	-0.47 to 0.43
PAC - oesophagus	0	0	N/A	N/A	N/A
PAC - rectum	0	0	N/A	N/A	N/A
PAC - axillary	5	4	0.54 (0.44 to 0.63; 85%)	0.36 (0.22 to 0.51; 97%)	-0.17 to 1.25
PAC - tympanic	3	3	0.34 (-0.30 to 0.97; 98%)	0.65 (0.16 to 1.15; 98%)	-0.93 to 1.61
PAC - Zero Heat Flux	3	2	0.24 (0.04 to 0.44; 100%)	0.33 (0.20 to 0.46; 100%)	-0.41 to 0.89

Abbreviations: CI, confidence interval; SD, standard deviation; PAC, Pulmonary Artery Catheter; N/A, not applicable.

Table S3. Summary of the QUADAS-2 risk-of-bias assessment for included studies.

Author, year	Patient selection bias	Patient selection applicability	Index test bias	Index test applicability	Reference standard bias	Reference standard applicability	Flow and time bias
Dahyot-Fizelier, 2017	Unclear	Unclear	Unclear	Low	Low	Low	Low
Eshraghi, 2014	Unclear	Unclear	Unclear	Low	Low	Low	Unclear
Farnell, 2005	Low	Low	Unclear	Low	Low	Low	Low
Fulbrook, 1997	Low	Low	Unclear	Low	Low	Low	Low
Haugk, 2010	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low
Lefrant, 2003	Low	Low	Unclear	Low	Low	Low	Low
Moran, 2007	Low	Low	Unclear	Low	Low	Low	Low
Myny, 2005	Low	Low	Unclear	Low	Low	Low	Low
Nierman, 1991	Low	Low	Unclear	Low	Low	Low	Low
Nonose, 2012	Low	Low	Unclear	Low	Low	Low	Low
Shin, 2013	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low
Smith, 2004	Unclear	Unclear	Unclear	Low	Low	Low	Low
Stavem, 1997	Low	Low	Unclear	Low	Low	Low	Low

Table S4. Source of bias for index test-associated measurements.

Author, year	Check for users training before thermometer use	Check for calibration before thermometer use	Check and control for confounders [1] before thermometer use
Dahyot-Fizelier, 2017	No	Yes	No
Eshraghi, 2014	No	Yes	No
Farnell, 2005	Yes	Yes	No
Fulbrook, 1997	Yes	Yes	Yes (axillary only)
Haugk, 2010	No	No	No
Lefrant, 2003	No	No	No
Moran, 2007	Yes (tympanic infrared – ear tug only)	No	No
Myny, 2005	Yes	No	No
Nierman, 1991	No	No	No
Nonose, 2012	Yes (tympanic infrared only)	No	Yes (axillary only)
Shin, 2013	No	No	Yes (tympanic infrared only)
Smith, 2004	Yes	Yes	No
Stavem, 1997	No	No	No

[1] Possible confounders:
Oral: endotracheal tube/mechanical ventilation; probe not in the posterior sublingual pocket; hot/cold fluids food; oral mucositis/stomatitis; no calibration of the thermometer.
Oesophageal: probe not at the same level of the heart; no measures to prevent displacement of the probe; contemporary enteral nutrition; no calibration of the thermometer.
Naso-Pharyngeal: no measures to prevent displacement of the probe; endotracheal tube; hot/cold fluids or food; inflammation; no calibration of the thermometer.
Temporal Artery: diaphoresis; dirty lens; air flowing across the face; vasopressors; no calibration of the thermometer.
Tympanic scanner: cerumen or inflammation of the ear canal; ear of measurement against the pillow; repeated measurements with a down time lesser than 2 mins; dirty lens; no calibration of the thermometer.

Axillary: hair; probe not placed with adduction of the arm to the body along the period of temperature measurement; active cooling/warming; no calibration of the thermometer.
Inguinal: probe not placed with adduction of the leg to the body along the period of temperature measurement; active cooling/warming; no calibration of the thermometer.
Urinary: oliguria.
Rectal: stool; no measure to prevent displacement of the probe; no calibration of the thermometer.
Pulmonary Artery Catheter: Extra Corporeal Membrane Oxygenation; contemporary administration of fluids via central line.

Figure S1. DerSimonian and Laird random effects meta-analysis of 13 studies reporting the mean temperature difference between pulmonary artery catheters and peripheral thermometers. Reported values include mean bias (index measurement subtracted from pulmonary artery catheter measurement) and 95% confidence intervals. Heterogeneity indices include I^2 and the p value of Cochran's Q. Lefrant[1] used an inguinal device; Lefrant[2] used an axillary device; Smith[3] used a DataTherm axillary device; Smith[4] used an axillary SolarTherm device.

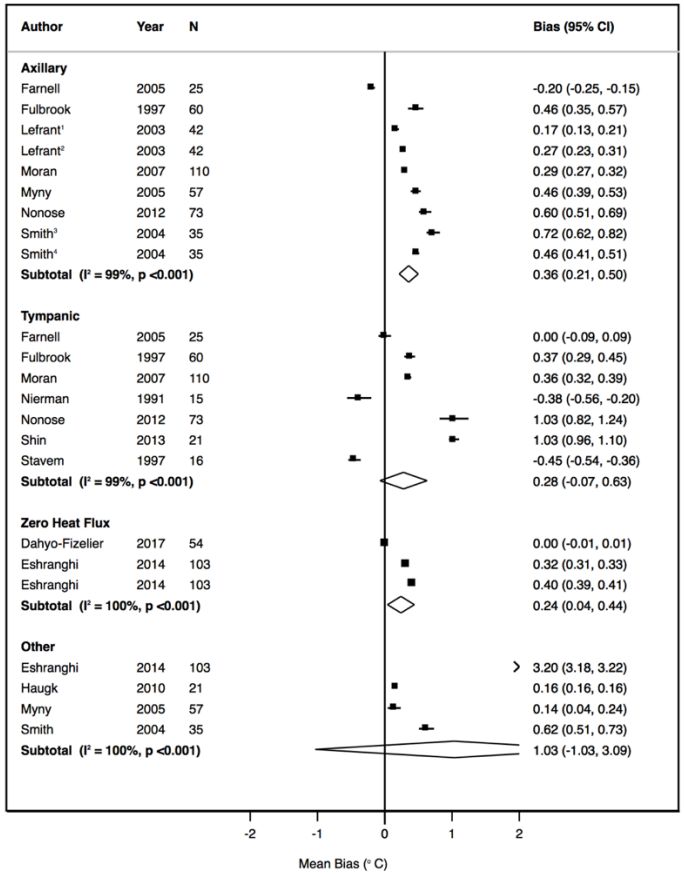


Figure S2. DerSimonian and Laird random effects meta-analysis of 13 studies reporting the mean temperature difference between pulmonary artery catheters and peripheral thermometers. Reported values include the standard deviation of the mean bias and 95% confidence intervals. Heterogeneity indices include I^2 and the p value of Cochran's Q. Lefrant[1] used an inguinal device; Lefrant[2] used an axillary device; Smith[3] used a DataTherm axillary device; Smith[4] used an axillary SolarTherm device.

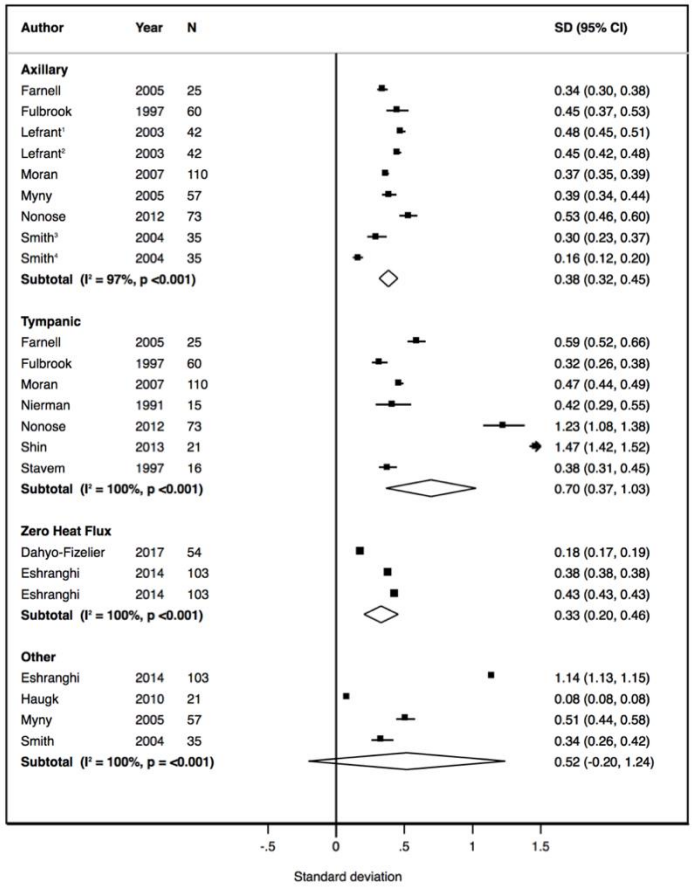


Figure S3. DerSimonian and Laird random effects meta-analysis of 7 studies reporting the mean temperature difference between pulmonary artery catheters and central thermometers. Reported values include mean bias (index measurement subtracted from pulmonary artery catheter measurement) and 95% confidence intervals. Heterogeneity indices include I^2 and the p value of Cochran's Q.

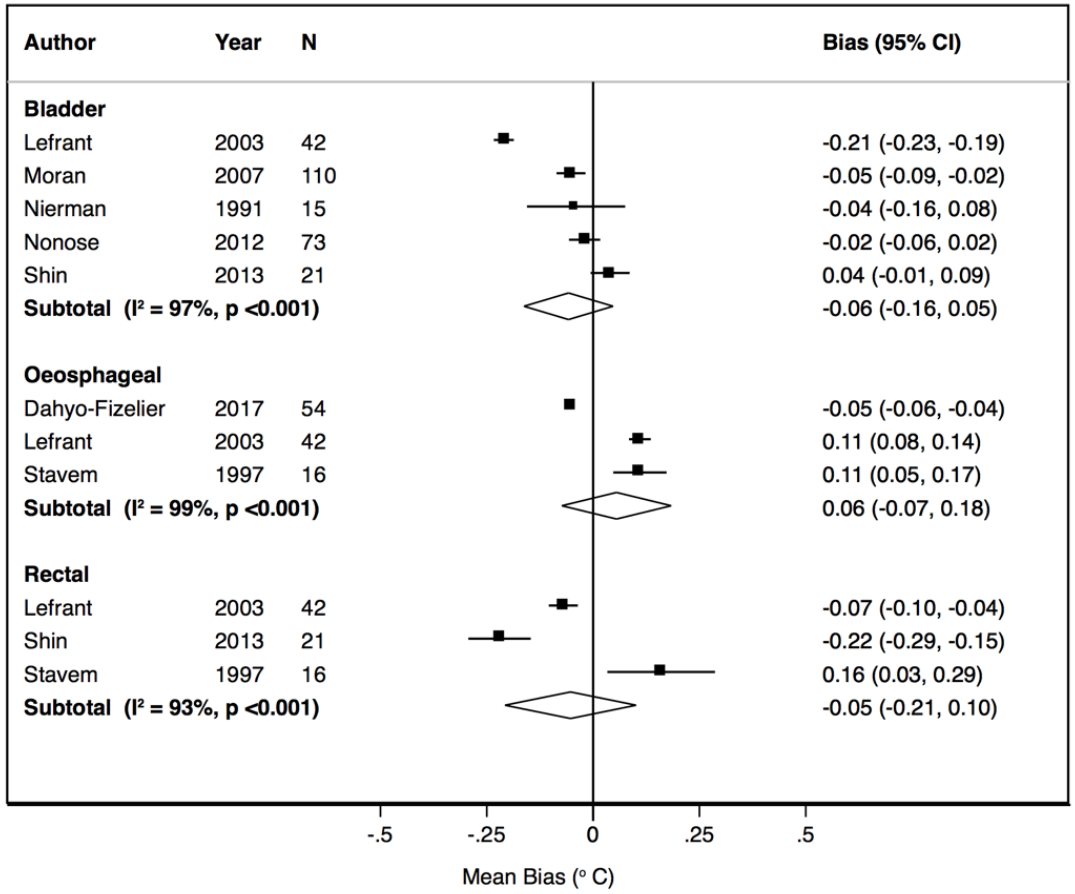


Figure S4. DerSimonian and Laird random effects meta-analysis of 7 studies reporting the mean temperature difference between pulmonary artery catheters and central thermometers. Reported values include the standard deviation of the mean bias and 95% confidence intervals. Heterogeneity indices include I^2 and the p value of Cochran's Q.

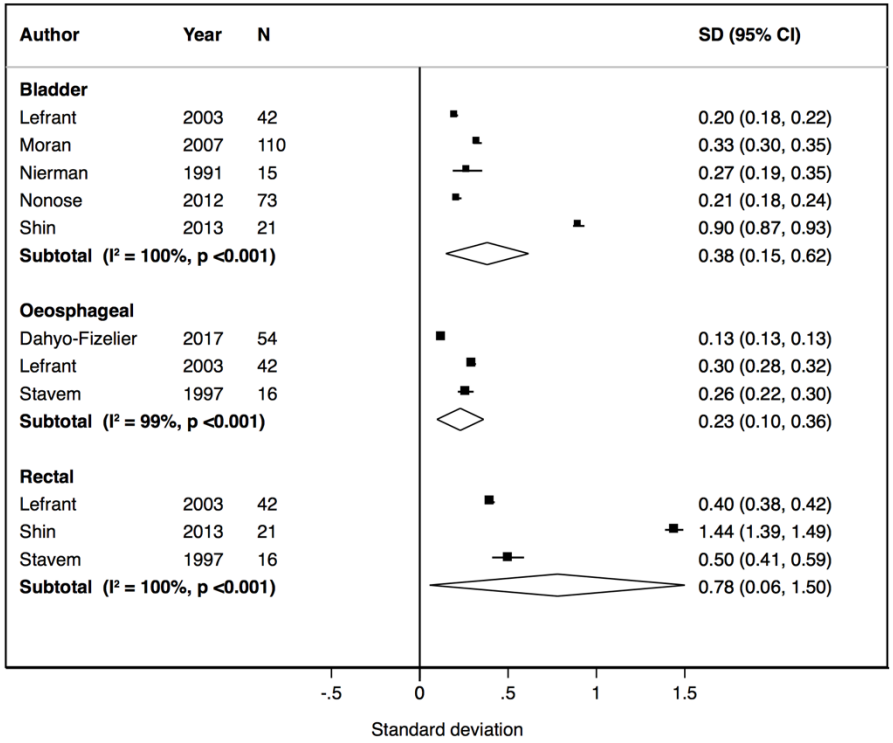


Figure S5. Box plot of the QUADAS-2 risk-of-bias assessment for included study.

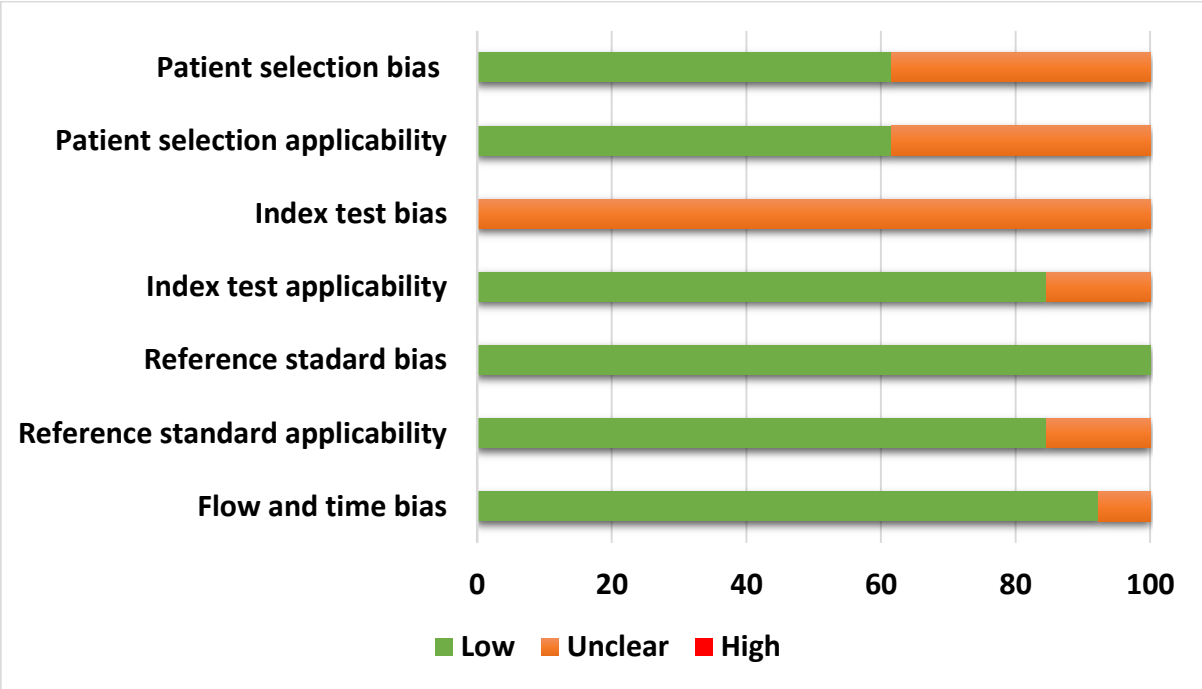


Figure S6a. Funnel plot assessment of the risk of publication bias in 7 studies comparing pulmonary artery catheters and central thermometers. Egger's regression test did not support the presence of small-study effects (p = 0.71).

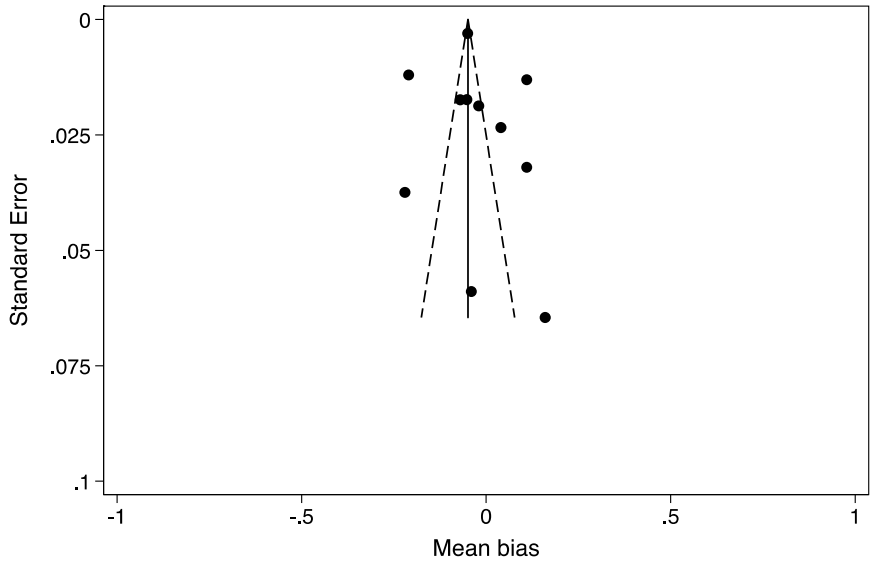


Figure S6b. Funnel plot assessment of the risk of publication bias in 13 studies comparing pulmonary artery catheters and peripheral thermometers. Egger's regression test did not support the presence of small-study effects (p = 0.44)