Office of Evidence Based Practice (EBP) – Critically Appraised Topic: Thermometers

Specific Care Question: In Pediatric patients, are temporal thermometers equivalent or better to axillary thermometers for measuring temperatures.

Question Originator
Shannon Lysaught, RN, BSN, CPN, MBA
Janet Franzen, RN, MSN, NE-BC
Whitney Pierce, RN, BSN, CPN
Megan McGurn RN, BSN, CPN

Literature Summary

Background. Fever is an atypical rise in body temperature that occurs as part of a particular biologic response (Ward, 2018). Children's Mercy's policy on Vital Signs recommends using the method to measure temperature that is most appropriate for age, development, and cognitive level of the patient (Vital Signs, 2018). The gold standard for measuring body temperature is core body temperature measurement (Sims, Patton, Williamson, & Ryan-Wenger, 2018). It is difficult to get an actual core body temperature because the procedures are invasive. The sites for core body temperature include the pulmonary artery, bladder, esophagus, or nasopharyngeal sites (Batra, Saha, & Faridi, 2012). Because measuring core temperature is invasive and not conducive to screening, non-core temperature techniques are used in hospital and ambulatory environments (Sim et al., 2018). Non-core temperature techniques include rectal, oral, axillary, tympanic, or temporal artery. The purpose of this review is to examine the accuracy of non-core thermometers used in hospitals and ambulatory environments. To more accurately reflect the body of literature, the review expanded from the original question that compared two non-core temperature techniques (temporal and axillary) to compare all non-core temperature techniques to core temperature to identify best practice for temperature assessment.

Study characteristics. The search for suitable studies was completed on September 17, 2018. Jeff Michael, DO reviewed the 21 titles and abstracts found in the search and identified five articles believed to answer the question. After an in-depth review, one high-quality systematic review (SR) with a meta-analysis (MA) answered the question (Ryan-Wenger, Sims, Patton, and Williamson, 2018).

Key results. Ryan-Wenger et al. (2018) examined 159 studies, of which 34 were included in the MA that compared the accuracy of core thermometers temperatures to non-core thermometer temperatures. Ryan-Wenger et al. (2018) employed the Quality Assessment of Diagnostic Accuracy Studies (QUADAS 2) criteria to analyze the risk of bias in the 34 studies (Whiting et al., 2011). The GRADE criteria for diagnostic tests and strategies (Schünemann, Brozek, & Oxman, 2013) was used to evaluate the quality of studies included in the meta-analysis (Ryan-Wenger et al., 2018). The authors rated the studies as high quality. Only studies that used core temperatures (pulmonary artery, esophageal, bladder, and nasopharyngeal) as the reference standard were included in the SR with MA.

Oral and rectal electronic thermometer devices were shown to be the only non-core thermometers that met accuracy criterion of remaining within ±0.5 °C of the core temperature 95% of the time (this accuracy was reported as Limits of Agreement or LOA). Axillary chemical, axillary electronic, tympanic, and temporal artery thermometers are not recommended due to the sizeable mean difference from the reference standard (see Table 1). It should be understood that when peripheral thermometers (other than oral or rectal) are used, there is a high risk of overestimating or underestimating temperatures and the results may not be reliable for clinical decision making. It is also important to point out, the mean difference varied too widely between the same types of thermometers to develop a standard conversion.

While the Ryan-Wenger et al. (2018) MA recommends only oral or rectal temperatures, there are patients where this may not be possible. Oral temperatures are difficult to get in children less than five years of age, and rectal temperatures are not recommended for premature infants, patients with potential for bleeding, altered immune systems, and rectal abnormalities (Children's Mercy, 2018). Also, evidence of patient comfort and preference is lacking in existing literature (Ryan-Wenger et al., 2018).

Summary of Evidence

Ryan-Wenger et al. (2018) reported the accuracy between core and non-core temperature measurement. Ryan-Wenger et al. (2018) defined the accuracy criterion to be if the tested thermometer reading remained within ±0.5 °C of core temperature 95% of the time. Temperature accuracy, for this report, is the pooled mean difference (*MD*) and the associated confidence intervals (*CI*). It is important to note that within some of the single studies reported in the Ryan-Wenger et al. (2018) SR the patient temperatures were taken through either multiple routes (such as left and right tympanic measurement) or different thermometers therefore, when this occurred the sample size reported will be larger than the actual number of single studies identified (see Table 1).

Oral electronic thermometers to core temperatures. Seven temperature samples compared oral electronic thermometers to core temperatures. Oral electronic thermometers had a *MD* of - 0.05°C from core temperature, 95% *CI* [-0.38, 0.24]. Oral electronic thermometers underestimated core body temperature 43% (*n* = 3) of the time and overestimated it 57% (*n* = 4) of the time. Oral thermometer devices met accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Rectal electronic thermometers to core temperatures. Fourteen temperature samples compared electronic rectal thermometers to core temperatures. Rectal electronic thermometers had a *MD* of -0.04°C from core temperature, 95% *CI* [-0.63, 0.55]. Rectal electronic thermometers underestimated core body temperature 43% (*n* = 6) of the time, overestimated it 50% (*n* = 7) of the time, and one study reported *MD* as 0°C (7%). Rectal thermometer devices met accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Temporal artery thermometers to core temperatures. Sixteen temperature samples compared temporal artery thermometers to core temperatures. Temporal arterial thermometers had a *MD* of 0.25°C from core temperature, 95% *CI* [-0.99, 1.19]. Temporal artery thermometers underestimated core body temperature 19% (*n* = 3) of the time and overestimated it 81% (*n* = 13) of the time. Temporal artery thermometer devices did not meet accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Tympanic thermometers to core temperatures. Thirty-nine temperature samples compared tympanic thermometers to core temperatures. Tympanic thermometers had a *MD* of 1.05°C from core temperature, 95% *CI* [-0.26, 2.36]. Tympanic thermometers underestimated core body temperature 41% (*n* = 16) of the time, overestimated it 51% (*n* = 20) of the time, and in three studies the *MD* was 0°C (8%). Tympanic thermometer devices did not meet accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Axillary chemical thermometers to core temperatures. Five temperature samples compared axillary chemical thermometers to core temperatures. Axillary chemical thermometers had a *MD* of 0.27°C from core temperature, 95% *CI* [-1.00, 1.54]. Axillary chemical thermometers underestimated core body temperature 20% (*n* = 1) of the time and overestimated it 80% (*n* = 4) of the time. Axillary chemical thermometer devices did not meet accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Axillary electronic thermometers to core temperatures. Fifteen temperature samples compared axillary electronic thermometer temperatures to core temperatures. Axillary electronic thermometers had a *MD* of - 0.19°C from core temperature, 95% *CI* [-1.00, 1.54]. Axillary electronic thermometers underestimated core body temperature 53% (*n* = 8) of the time and overestimated it 47% (*n* = 7) of the time. Axillary electronic thermometer devices did not meet accuracy criterion of remaining within ±0.5 °C of core temperature 95% of the time.

Search Strategy and Results (see PRISMA diagram)
(("Thermometers"[Majr] OR "Thermometry"[Majr] OR "Body Temperature Changes"[Majr] OR "Body Temperature"[Majr]) AND ((temporal[tw] OR temple[tw] OR "temporal artery"[tw]) AND axillary[tw])) AND (children OR child OR infant OR adolescence OR pediatr* OR paediatr*) Total: 21

Studies Included in this Review (in Alphabetical Order)
Ryan-Wenger et al. (2018)

Studies <u>Not</u> Included in this Review with Exclusion Rationale (in Alphabetical Order)

Authors (YYYY)	Reason for exclusion
Forrest et al. (2017)	Rectal thermometer used as the reference standard
Niven et al. (2015)	Included in Ryan-Wenger et al. (2018)
Opersteny et al. (2017)	Rectal thermometer used as the reference standard
Sahin et al. (2012)	Included in Ryan-Wenger et al. (2018)

Method Used for Appraisal and Synthesis
[a]Higgins, J. P. T., & Green, S. e. (2011). *Cochrane Handbook for Systematic Reviews of Interventions [updated March 2011]* (Version 5.1.0 ed.): The Cochrane Collaboration, 2011.

Medical Librarian Responsible for the Search Strategy
Keri Swaggart, MLIS, AHIP
EBP Scholar's Responsible for Analyzing the Literature
Azadeh Wickham, FNP-C
EBP Team Member Responsible for Reviewing, Synthesizing, and Developing this Document
Jarrod Dusin, MS, RD, LD, CPHQ

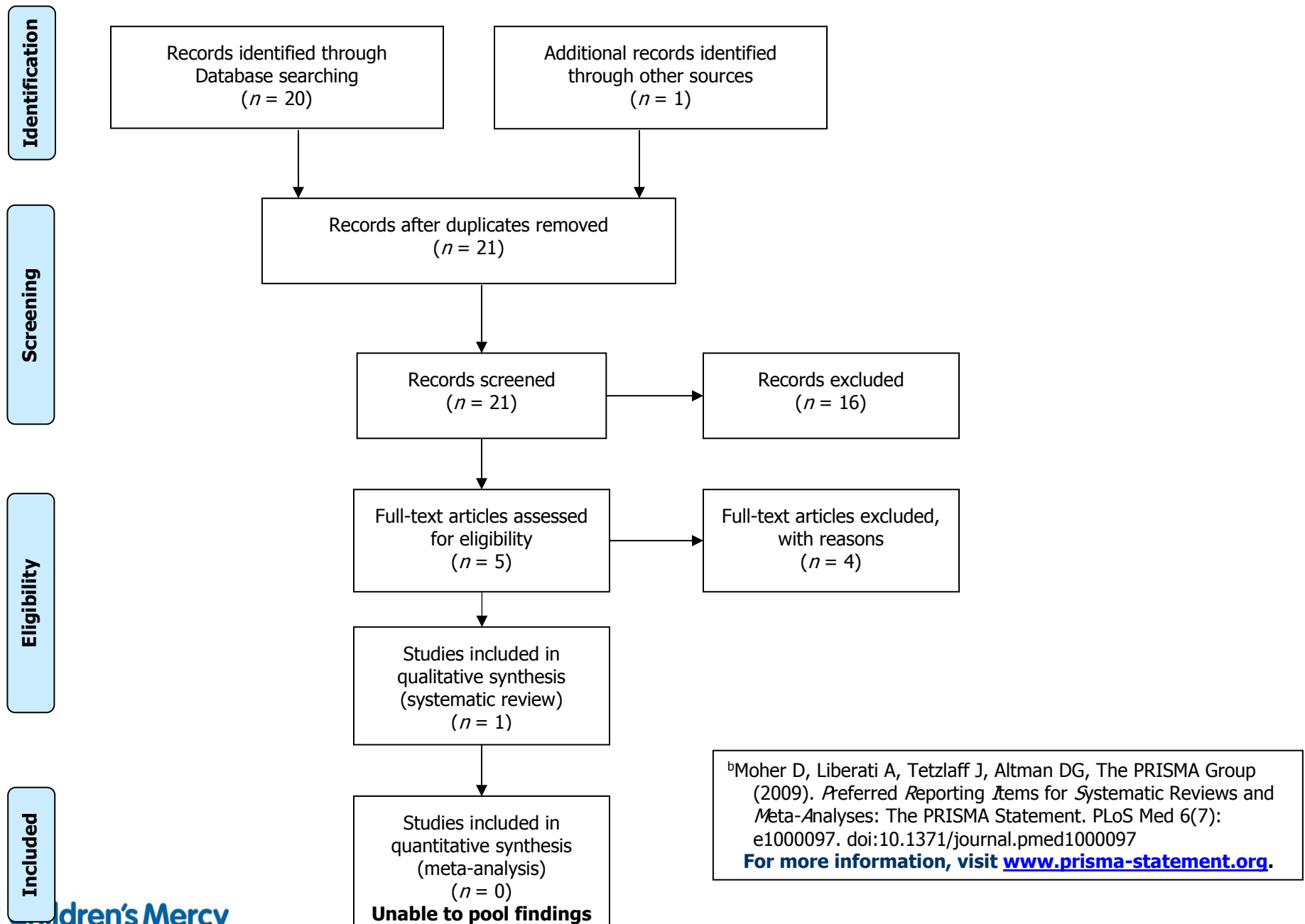**Acronyms Used in this Document**

Acronym	Explanation
LOA	Limits of Agreement
MD	Mean Difference
ES	Effect Size
CI	Confidence Intervals
CL	Confidence Limits
SD	Standard Deviation
C	Celsius
QUADAS	Quality Assessment of Diagnostic Accuracy Studies

Date Developed/Updated November 2018

Office of Evidence Based Practice (EBP) – Critically Appraised Topic: Thermometers

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRIMSA)[b]

Identification

Records identified through Database searching
($n = 20$)

Additional records identified through other sources
($n = 1$)

Screening

Records after duplicates removed
($n = 21$)

Records screened
($n = 21$)

Records excluded
($n = 16$)

Eligibility

Full-text articles assessed for eligibility
($n = 5$)

Full-text articles excluded, with reasons
($n = 4$)

Studies included in qualitative synthesis (systematic review)
($n = 1$)

Included

Studies included in quantitative synthesis (meta-analysis)
($n = 0$)
Unable to pool findings

[b]Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097
For more information, visit www.prisma-statement.org.

Children's Mercy

Table *1*

Devices	Number of studies	Number of samples	Mean Difference between non-core temperature and core temperature.	95% CI
Oral Electronic	*n* = 7	*n* = 7	- 0.05	- 0.38, 0.24
Rectal Electronic	*n* = 13	*n* =14*	- 0.04	- 0.63, 0.55
Axillary Electronic	*n* = 14	*n* = 15*	- 0.19	- 0.69, 1.37
Temporal Artery	*n* = 16	*n* = 16	0.25	- 0.99, 1.19
Axillary Chemical	*n* = 4	*n* = 5*	0.27	- 1.00, 1.54
Tympanic	*n* = 24	*n* = 39*	1.05	- 0.26, 2.36

*Number of samples is increased from the number of studies due to different thermometers or temperature sites within one study (such as right and left axillary temperatures).

Table 2
Ryan-Wenger 2018

Design	Quantitative Synthesis (meta-analysis)
Objective	The purpose of this meta-analysis was to review and synthesize research on the accuracy of thermometer devices commonly used in hospitals and clinics, to make recommendations for use in clinical practice. For patients greater than one month of age to adults are temperatures from non-core devices as accurate as core temperatures?
Methods	**Protocol and registration.** The protocol was not registered. **Eligibility Criteria.** Studies that compared core body temperatures versus body temperatures from non-core thermometer devices in children older than one month of age to adulthood. **Information sources.** • Medline • Cumulative Index of Nursing and Allied Health Literature • Cochrane Database • Clinical Key • National Guideline Clearinghouse **Search.** See study for search strategy used. **Study Selection.** There were 244 articles retrieved for initial review by institution's Research and Evidence-Based Practice team. There was a secondary review with the following criteria: Concurrent or sequential core and non-core temperature method, reports of mean differences, standard deviation, and 95% confidence intervals between core and non-core temperatures, or sufficient data to calculate these statistics. Reported sample sizes needed to be > 10. The authors did not disclose who screened the studies or how disagreements were resolved. **Data collection process.** The authors indicate that the research and evidence-based practice team individually reviewed the articles. **Risk of bias (RoB) across studies.** The Quality Assessment of Diagnostic Accuracy Studies (QUADAS 2) criteria were used for risk of bias. **Summary measures.** Effect size (ES), Standard error (SE), confidence limits (CL), and confidence intervals (CI), along with inter-rater and intra-rater reliability were reported.

Office of Evidence Based Practice (EBP) – Critically Appraised Topic: Thermometers

	Synthesis of results. Confidence limits were reported at 95%; the GRADE criteria for diagnostic tests and strategies were used to evaluate the studies. **Additional analyses.** Tests for precision and heterogeneity of six non-core thermometer devises based on variation of temperatures between and within studies was completed.

Results	**Study Selection.** Give numbers of studies screened, assessed for eligibility, and included in the review.
	    Number of articles identified: *N* = 244
	        o **Full-text articles assessed for eligibility:** *n* = 159
	        o **Studies included in quantitative synthesis:** *n* = 34
	
	Synthesis of results.
	Temporal Artery Thermometers Assessment: *n* = 16
	• Mean difference range - 0.44 to + 1.3°C.
	• Core body temperature was overestimated in 81.3% (*n* = 13).
	• Effect sizes ± 0.02 and ± 1.8°C
	• Lower CL range - 2.99 to + 4.1°C
	• Upper CL range - 0.41 to + 3.74°C
	
	Tympanic Thermometers Assessment: *n* = 39
	• Mean difference range - 1.06 to + 0.98°C
	• Core body temperature was underestimated in 41% (*n* = 16) and overestimated in 51% (*n* = 20). Three studies with mean differences of 0°C (8%)
	• Effect sizes ± 0.03 and ± 1.02°C
	• Lower CL range - 2.98 to - 0.24°C
	• Upper CL range - 0.38 to + 3.8°C
	
	Axillary Chemical Thermometers: *n* = 5
	• Mean difference range - 0.01 to + 0.50°C
	• Core body temperature was overestimated in 80% (*n* = 4)
	• Effect sizes ± 0.35 and ± 0.53°C
	• Lower CL range - 0.75 to - 0.49°C
	• Upper CL range + 0.73 to + 1.51°C
	
	Axillary Electronic Thermometers: *n* = 15
	• Mean difference range - 1.25 to + 0.60°C
	• Core body temperature was underestimated in 53% (*n* = 8) and overestimated in 47% (*n* = 7).
	• Effect sizes ± 0.26 and ± 1.00°C
	• Lower CL range - 2.94 to - 0.27°C
	• Upper CL range - 0.33 to + 2.17°C
	
	Oral Electronic Thermometers: *n* = 7
	• Mean difference bias range - 0.25 to + 0.12°C
	• Core body temperature was underestimated in 43% (*n* = 3) and overestimated in 69% (*n* = 4).
	• Effect sizes ± 0.15 and ± 0.45°C

- Lower CL range -1.07 to -0.30°C
- Upper CL range + 0.29 to + 1.02°C

Rectal Electronic Thermometers: *n* = 14
- Mean difference range - 0.69 to + 0.54°C
- Core body temperature was underestimated in 40% (*n* = 6) and overestimated in 47% (*n* = 7), with one mean difference reported to be 0°C
- Effect sizes ± 0.10 and ± 1.00°C
- Lower CL range - 2.36 to + 0.19°C
- Upper CL range - 0.11 to + 1.56°C

Risk of bias across studies.
- The authors used the QUADAS 2 criteria to report the risk of bias.
- Three of four criteria (selection of patients, conduct, and interpretations between core and non-core devices) were rated as low risk.
- The remaining criteria, patient flow, was rated as low quality as the studies that comprised the meta-analysis poorly described this component.

Additional analysis. The meta-analysis authors employed the GRADE criteria to analyze the body of literature reviewed. The authors identified their confidence in the estimate of effects was high based on:
- Consistency of study design
- Discrepancies in precision were explained by the individual study authors
- Publication bias was rated as low

Discussion	**Summary of evidence.** "The meta-analysis findings indicate that only oral and rectal electronic thermometer devices should be used to measure temperature of individuals for screening, monitoring, diagnostic, and treatment decisions. Tympanic temporal, axillary chemical, and axillary electronic thermometer devices are not recommended to use in clinical practice." **Limitations.** There are no limitations discussed for this meta-analysis
Funding	**Funding.** The authors did not identify any outside source of funding was used for this systematic review/meta-analysis.

References

Batra, P., Saha, A., & Faridi, M. M. A. (2012). Thermometry in children. *Journal of emergencies, trauma, and shock, 5*(3), 246.

Forrest, A. J., Juliano, M. L., Conley, S. P., Cronyn, P. D., McGlynn, A., & Auten, J. D. (2017). Temporal artery and axillary thermometry comparison with rectal thermometry in children presenting to the ED. *American Journal of Emergency Medicine, 35*(12), 1855-1858. doi: 10.1016/j.ajem.2017.06.017

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097 For more information, visit www.prisma-statement.org.

Niven, D. J., Gaudet, J. E., Laupland, K. B., Mrklas, K. J., Roberts, D. J., & Stelfox, H. (2015). Accuracy of peripheral thermometers for estimating temperature: A systematic review and meta-analysis. *Annals of internal medicine, 163*(10), 768-777. doi: 10.7326/M15-1150

Opersteny, E., Anderson, H., Bates, J., Davenport, K., Husby, J., Myking, K., & Oron, A. P. (2017). Precision, Sensitivity and Patient Preference of Non-Invasive Thermometers in a Pediatric Surgical Acute Care Setting. *Journal of Pediatric Nursing, 35*, 36-41. doi: 10.1016/j.pedn.2017.02.003

Ryan-Wenger, N. A., Sims, M. A., Patton, R. A., & Williamson, J. (2018). Selection of the most accurate thermometer devices for clinical practice: Part 1: Meta-analysis of the accuracy of non-core thermometer devices compared to core body temperature. *Pediatric Nursing, 44*(3), 116-133.

Sahin, S. H. M. D., Duran, R. M. D., Sut, N. P., Colak, A. M. D., Acunas, B. M. D., & Aksu, B. M. D. (2012). Comparison of temporal artery, nasopharyngeal, and axillary temperature measurement during anesthesia in children. *Journal of Clinical Anesthesia, 24*(8), 647-651. doi: 10.1016/j.jclinane.2012.05.003

Schünemann, H., Brozek, J., & Oxman, A. (2013). GRADE handbook for grading quality of evidence and strength of recommendations. *Updated October*.

Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., . . . Bossuyt, P. M. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine, 155*(8), 529-536.

Vital Signs, (September, 2018), *CMH Patient Care Services Standards Manuel*. Children's Mercy Hospital, Kansas City, Missouri.

Ward, M., (2018). https://www.uptodate.com/contents/fever-in-infants-and-children-pathophysiology-and-management?search=temperature&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1