

Performance of temporal artery temperature measurement in ruling out fever: implications for COVID-19 screening.

Adrian D. Haimovich MD, PhD¹, R. Andrew Taylor MD, MHS¹, Harlan M. Krumholz, MD SM^{2,3}, Arjun K. Venkatesh MD, MBA, MHS^{1,3}

1 Yale Department of Emergency Medicine, Yale School of Medicine, 464 Congress St. 260, New Haven, CT, 06516

2 Yale Department of Internal Medicine, Section of Cardiovascular Medicine, Yale School of Medicine, 333 Cedar Street, New Haven, CT 06520

3 Yale New Haven Hospital Center for Outcomes Research and Evaluation, 1 Church Street, New Haven, CT 06510

* adrian.haimovich@yale.edu

Background

Current CDC recommendations for mitigation of community COVID-19 transmission include temperature screening.¹ Port-of-entry symptom and temperature screenings are now common worldwide in community settings such as airports as well as healthcare settings such as Veterans Health Administration Hospitals and are being considered as part of return-to-work protocols.^{2,3} Due to low cost and ease-of-use, temporal artery thermometers (TATs) applied to the forehead are widely employed to screen for fever, but prior literature has suggested poor sensitivity and high variability.^{4,5} Published data are limited by small per-study sample sizes and a focus on pediatric, surgical, and intensive care settings that are not generalizable to real-world screening populations.

Objective

We sought to determine the real-world test performance of TATs for fever rule-out by utilizing a large electronic dataset of emergency department encounters for whom universal temperature screening was conducted as part of standard triage processes. For reference standards, we included rectal temperature, a widely recognized core temperature, as well as oral temperature, which benefits from wide clinical acceptance and robust specificity.⁵ Our primary objective was to determine TAT sensitivities and specificities across a range of temperatures in comparison to rectal and oral cutoffs of 100.4°F/38°C.^{4,5} Our secondary outcome was limit-of-agreement (LOA) by Bland-Altman analysis.

Methods and Findings

We extracted temperature measurements and method of temperature assessment from electronic health record (EHR) data (Epic, Verona, WI) collected between March 2013 and June 2019 within a large hospital system comprising ten acute care sites. These data were part of a quality improvement effort and exempted from review by the Yale University Institutional Review Board. We identified paired per-patient data where a TAT measurement was documented within 15 minutes of a rectal temperature measurement or oral temperature. When multiple measurements were taken with a single modality within the defined interval, the mean value was used.

We identified 1.84 million adult (age > 18 years) emergency department visits by 602,089 patients with over 4.6 million temperature readings; there were 1,293 paired readings from 1,276 encounters that met our inclusion for TAT versus rectal measurement and 16,132 readings from 16,031 encounters for TAT versus oral measurement (Table 1). Fever prevalence in the rectal and oral temperature populations were 34.4% and 4.3%, respectively. Using a threshold of 100.4°F, TAT measurement identified fever compared with the rectal reference with sensitivity 0.27 (95% CI: 0.27-0.31), specificity 0.98 (0.96-0.99), PPV 0.85 (0.79-0.91), NPV 0.72 (0.69-0.74). TAT measurement identified fever

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Table 1: Characteristics of Patients with Temperature Measurement in the ED

		Overall	Temperature Type	
			Oral	Rectal
n		17271	16031	1276
Age	mean [IQR]	57.0 [39.0,74.0]	56.0 [38.0,72.0]	76.0 [61.0,87.0]
Gender	Female	9008 (52.2)	8336 (52.0)	689 (54.0)
	Male	8263 (47.8)	7695 (48.0)	587 (46.0)
Race	American Indian or Alaska Native	41 (0.2)	39 (0.2)	2 (0.2)
	Asian	229 (1.3)	217 (1.4)	12 (0.9)
	Black or African American	4187 (24.2)	3937 (24.6)	252 (19.7)
	Native Hawaiian or Other Pacific Islander	52 (0.3)	49 (0.3)	3 (0.2)
	Other	642 (3.7)	600 (3.7)	44 (3.4)
	Unknown	1872 (10.8)	1796 (11.2)	79 (6.2)
	White or Caucasian	10248 (59.3)	9393 (58.6)	884 (69.3)
Insurance	Commercial	4727 (27.4)	4443 (27.7)	290 (22.7)
	Medicaid	4171 (24.2)	4020 (25.1)	155 (12.1)
	Medicare	4536 (26.3)	3976 (24.8)	578 (45.3)
	Other	3755 (21.7)	3522 (22.0)	241 (18.9)
	Self pay	82 (0.5)	70 (0.4)	12 (0.9)
Ethnicity	Hispanic or Latino	2770 (16.0)	2658 (16.6)	114 (8.9)
	Non-Hispanic	14399 (83.4)	13280 (82.8)	1153 (90.4)
	Unknown	102 (0.6)	93 (0.6)	9 (0.7)

compared with the oral reference with a sensitivity 0.23 (95% CI: 0.20-0.26), specificity 0.99 (0.99-0.99), PPV 0.53 (0.48-0.59), NPV 0.97 (0.96-0.97). We did not observe significant differences in fever sensitivity when limiting our paired analysis to five or ten minute windows. Decreasing the threshold for fever using TAT improved sensitivity with respect to the rectal temperature gold standard of 100.4°F (Figure 1). Using a cutoff of 99°F, for example, increased sensitivity to 0.63 (0.58-0.67) and NPV to 0.82 (0.79-0.84), while specificity decreased to 0.86 (0.83-0.88) and PPV to 0.7 (0.65-0.74). The mean differences and LOA between TAT and rectal temperatures and TAT and oral temperatures were -1.13 (LOA: -6.69 to 4.42)°F, and -0.45 (-3.89-2.99)°F, respectively (Figure 2).

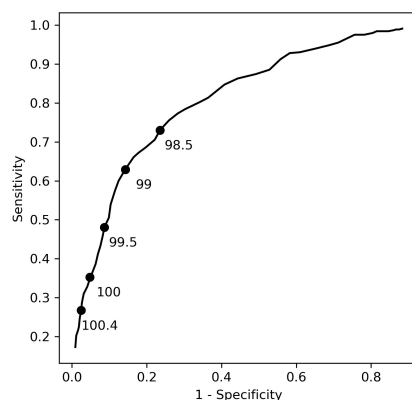


Figure 1: Specificity/sensitivity plot for fever screening by TAT as compared to rectal temperature of 100.4°F. Select temperature cutoffs for TAT are labeled.

Discussion

To our knowledge, this is the largest study to date comparing TAT measurement to rectal or oral measurements in an adult population. Consistent with prior work,^{4,5} we found TATs had poor test performance, identifying less than one in three positive cases using either rectal or oral reference measurements. We observed that decreasing TAT temperature thresholds yielded significantly improved

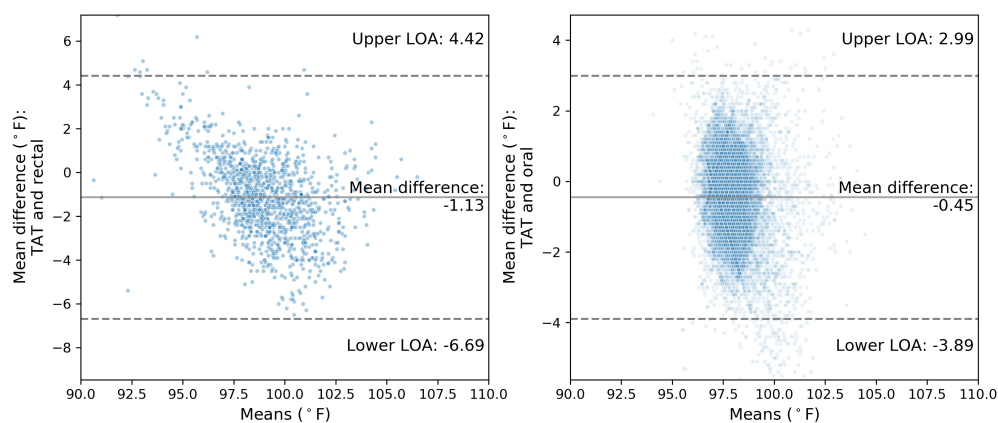


Figure 2: Bland-Altman plots of TAT and (A) rectal and (B) oral measurements. Mean difference is shown with the solid line, limits of agreement as defined by two standard deviations are shown with dashed lines.

test sensitivity with modest losses in specificity. Our analysis is limited by its reliance on selection of patients who had paired measures, which may limit the generalizability. Nevertheless, these data raise questions about the sensitivity of TAT screening for the detection of people with fever amidst the current COVID-19 pandemic. Using a TAT cutoff of 100.4°F, the majority of people who would meet criteria for fever would be wrongly classified as afebrile.

References

1. Implementation of Mitigation Strategies for Communities with Local COVID-19 Transmission. Centers for Disease Control and Prevention; 2020.
2. Outbreak of acute respiratory syndrome associated with a novel coronavirus, China: third update. European Centre for Disease Prevention and Control; 2020.
3. Normile D. Why airport screening won't stop the spread of coronavirus. Science; 2020. Available from: <https://www.sciencemag.org/news/2020/03/why-airport-screening-wont-stop-spread-coronavirus>.
4. Geijer H, Udumyan R, Lohse G, Nilsagård Y. Temperature measurements with a temporal scanner: systematic review and meta-analysis. *BMJ open*. 2016;6(3):e009509.
5. Niven DJ, Gaudet JE, Laupland KB, Mrklas KJ, Roberts DJ, Stelfox HT. Accuracy of peripheral thermometers for estimating temperature: a systematic review and meta-analysis. *Annals of internal medicine*. 2015;163(10):768–777.