Correction of human forehead temperature variations measured by non-contact infrared thermometer

Adrian Shajkofci, *Member, IEEE*

Abstract—Elevated body temperature (fever) can be a common symptom of a medical condition, such as a viral or bacterial infection, including SARS-CoV-2 or influenza. Non-contact infrared thermometers are able to measure forehead temperature in a timely manner and were used to perform a fast fever screening in a population. However, forehead temperature measurements differ greatly from basal body temperatures, and are the target of massive perturbations from the environment. Here we gathered a dataset of $N=18024$ measurements using the same precision infrared sensor in different locations while tracking both outside temperature, room temperature, time of measurement, and identity. Herein, we propose a method able to extract and remove the influence of external perturbations and set the threshold for fever based on local statistics to 37.38°C , after calibration and temperature correction. This method can help manufacturers and decision-makers to build and use more accurate tools so as to maximize both sensitivity and specificity of the screening protocol.

Index Terms—fever, ncit, infrared, screening, temperature, medical monitoring, thermopile

I. INTRODUCTION

FEVER, also called pyrexia, is one of the usual clinical features that appears during the course of several infectious diseases, such as influenza or SARS-CoV-2 viral infections [1]. It reduces viral replication and is a straightforward marker of immune response [2]. Body temperature can be measured in numerous ways. The traditional method for such a measurement is using contact thermometers placed in the mouth, ear, armpit or rectum. Contact thermometers are measuring temperature using the conduction of heat to a thermocouple, a resistance temperature detector (RTD) or a thermistor through a metallic element. While RTDs are among the most precise temperature sensors available, the process of conduction is slow (around one minute) and the device needs extensive disinfection after use to prevent cross-contamination [3]. On the other hand, Non-contact infrared thermometers (NCITs) allow the temperature to be taken without contact and therefore do not require constant disinfection, take fast measurements (less than one second) and allow for a comfortable and much less intrusive measurement process. NCITs were extensively reviewed in [4] and validated for their use for fever symptoms detection.

NCITs can be classified into two categories. The first category comprises infrared cameras, also called thermal scanners. These devices operate like normal digital cameras, and capture a map

of temperatures over a constant field-of-view (FOV). They are used to measure objects up to 10 meters in front of the camera. They focus heat radiation using coated optics made of germanium, zinc salts, or surface mirrors, towards a focal plane array (FPA) of heat-conductive pixels. The relative temperature of the FPA is then sequentially read by a bolometer that converts the temperature difference into resistance, further converted into electrical signals. Thermal imaging cameras require an extensive calibration using different sensitivity curves allocated to each pixel, their output being compared to the temperature of a black body. These devices usually need to be stabilized at a defined chip temperature [5]. Moreover, they are greatly dependent on the fluctuations of air temperature since the camera receives the radiation emitted from the targeted object, as well as heat coming from surrounding objects or the atmosphere [6]. According to Minkina and Klecha [7], the following parameters should be considered when calculating the temperature of the object: the emissivity of the object material (ϵ), the transmittance of the atmosphere, the distance between the object and the camera, the relative humidity, the atmospheric temperature T_{amb} and information about the surrounding heat sources (sun, halogen lamps, heated walls, etc.). These parameters are often difficult to gather and one must make approximations when designing a NCIT system. Moreover, these parameters can easily shift over time due to wind (or a draught), rain, or direct sunlight. Consequently, this type of sensor is known not to be accurate enough for medical temperature measurement. These devices are subject to an error

Manuscript received November 17, 2020; revised December 28, 2020.
Adrian Shajkofci is with the Idiap Research Institute, CH-1920 Martigny, Switzerland and École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland.

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

of 0.5 °C to 5 °C when compared to contact thermometers, depending on the calibration modality [8]. The second type of NCIT is the family of thermopiles, mostly found in handheld forehead or tympanic thermometers. They receive the radiant energy released by the object over a constant FOV and convert it into an electrical signal (using the Seebeck effect), which is calibrated to the relative temperature of the object. Thermopiles are more accurate than infrared cameras since the ambient air flow is minimized due to the smaller distance between the sensor and the imaged object (1 cm to 20 cm) [9]. However, their accuracy drops drastically when used outside of their recommended range [10].

The accuracy of the measurement of human body temperature using NCITs is affected by individual factors, such as instrument placement, skin type and thickness, sweating, ethnicity, or body mass index [11], [12], as well as environmental factors such as outdoor temperature, wind, relative humidity, and level of direct sunlight. Indeed, these factors can either alter the temperature of the skin, or change the calibration parameters of the devices [13], [14].

During the 2020 COVID-19 pandemic outbreak, NCITs were used for fever symptom screening at quarantine stations. Their accuracy is discussed since skin temperature must become acclimatized after coming inside from the cold [15]. For example, in recent statistics from Kaohsiung Hospital [16], during the month of March, only 5 patients out of 40887 were identified with fever when the measurement was made outside. Yet, at a second indoor measurement checkpoint, further 37 people were identified with fever since their bodies have acclimatized to the indoor temperature.

Manufacturers and users of NCITs usually calibrate their sensors to the ambient temperature using blackbody calibration sources as a temperature of reference [17]. Nevertheless, this calibration method does not take into account the effect of the outside temperature on the human skin. Indeed, the skin acts as an insulative layer that, in the presence of cold, prevents heat loss in a both passive (fat layer) and active (vasoconstriction, arteriovenous shunts) manner [18]. For these reasons, without any correction, skin temperature becomes an erroneous marker for body temperature.

In our research, we consider the variation of human forehead temperature in different environments defined by separate locations, times of measurement within a day, outside temperatures, and room temperatures. We then examine if these perturbations influence the measurement accuracy in the context of infection symptom screenings. We believe that the analysis and corrections of external perturbations on human forehead temperature can be used to improve fever screening protocols by adapting the fever detection threshold to the variations.

After having described the acquisition modalities in Section II-A, we design a statistical analysis of temperature variation of groups and individuals across time, ambient temperature, and meteorological conditions. Then, we propose in Section II-B a method to remove the effect of external perturbations. We publish the results in Section III. Finally, we discuss the results and suggest a threshold for fever detection in Section IV.

II. METHODS

A. Acquisition modalities and sensor calibration

We collected the data from participating companies that used the infrared sensor module Coronasense (Coronasense, Martigny, Switzerland). The device embeds a thermopile-based sensor element (MLX90614-DCI, Melexis Technologies NV, Belgium) with a reported measurement accuracy of ≤ 0.2 °C. The sensor chip is certified to comply with the ASSTM standard section 5.4 (Designation: E1965-98/2009) - Standard Specification for Infrared Thermometers for Intermittent Determination of Patient Temperature. The devices were installed at the entrance of the buildings and measurements were taken approximately 30 seconds after entry. For each measurement, the subject was told to aim at a point 3 cm above the junction between the eyes, 10 cm away from the sensor. Exact distance between the device and the subject was controlled and recorded using a Light Detection And Ranging (LIDAR) detector. A burst of 10 measurements were taken at different positions on the forehead, and the maximal value was kept. For a subset of the data, we gathered an anonymous imprint of the subject's identity, so we could track their temperature over time. For another subset of the data, we could also associate the measurements to the local meteorological conditions using OpenWeatherMap (OpenWeather Ltd., London, UK).

All sensors were factory calibrated using a black body and automatically calibrated themselves using the ambient air temperature. We adjusted the sensor output to the human body emissivity $\epsilon = 0.98$ from Plank's law [19]:

$$T = \sqrt[4]{\frac{T_{amb}^4 - (1 - \epsilon)T_{ir}^4}{\epsilon}}, \quad (1)$$

where T_{amb} and T_{ir} are the ambient temperature of the environment, and the mirrored temperature of the object, respectively. We added to every measurement an offset of 2.2 degrees in order to match with the average temperature obtained using a contact or a tympanic thermometer, as observed and concluded by [4], [20].

The acquired data is expressed as a mean and a variance. We calculated the coefficient of variance (CV) to determine the accuracy of the measurement. A CV under 5 % indicates a sufficient precision [21].

B. Correction of environmental perturbations

We now turn to the task of measuring the correlation between the forehead temperature and environmental factors (inside ambient temperature, outside temperature and hours in the days) in the aim of removing the influence of these factors from the raw data. We model the relationship between the forehead temperature and the ambient and outside temperature data as a combination of affine functions $t_a(t) = a_a t + c_a$ and $t_o(t) = a_o t + c_o$ respectively. Furthermore, we model the relationship between the hours in a day and the forehead temperature as a second-degree polynomial function $t_p(t) = t^2 + t + c_p$. We correct the temperature by subtracting to every point of data its corresponding point in the modeled curve or line and multiply, for every environmental perturbation, by the mean of the raw data.

Location	N	Average ($^{\circ}\text{C}$)	CV	Indiv. deviation ($^{\circ}\text{C}$)
All	18024	35.49 ± 0.80	2.25%	0.54 ± 0.29
A	3714	35.66 ± 0.89	2.50%	0.76 ± 0.25
B	4957	35.40 ± 0.67	1.89%	N/A
C	2508	35.19 ± 0.69	1.96%	0.39 ± 0.24

(a) without correction				
Location	N	Average ($^{\circ}\text{C}$)	CV	Indiv. deviation ($^{\circ}\text{C}$)
All	10078	35.49 ± 0.67	1.89%	0.49 ± 0.23
A	2992	35.64 ± 0.70	1.96%	0.57 ± 0.20
B	1608	35.38 ± 0.60	1.70%	N/A
C	2503	35.42 ± 0.54	1.52%	0.34 ± 0.19

(b) with corrections

TABLE I: Statistics of the dataset of forehead temperature measurements (average, standard deviation, coefficient of variance (CV)) before and after correction. We also tracked temperatures of individuals over time and reported the mean of the deviation of temperature for one person over time. Because of the limited availability of meteorological data, we could compute such corrections on a subset of the original data.

In the same way, one can instead correct the threshold for fever detection by subtracting the offset between the regression curve and the acquired raw data to the fever threshold $t_{\text{max_corr}}$. For example, to correct the threshold as a function of the outside temperature, we set:

$$t_{\text{max_corr}}(t) = t_{\text{max}} - a_o t - c_o, \quad (2)$$

with a_o and c_o trained from the previous measurements using a least-squares fit.

When it is possible to track measurements of a unique individual over time, we propose a personalized threshold when the temperature measurements can be associated with an individual identifier, computed using the K -th last measurements of the specific person i :

$$t_{\text{max_corr}}(t^i) = \bar{t}^i + \alpha \sqrt{\frac{\sum^K (\bar{t}^i - t_k^i)^2}{K}}, \quad (3)$$

where $\bar{t}^i$ is the mean of K temperature acquisitions for the individual i and $0 < \alpha < 5$ a coefficient for the standard variation. We use $\alpha = 3$ since it is the coefficient so that 99.7 % of the data remains detected as healthy under the normal approximation of the temperature distribution curve [22].

In the following section, for illustration purposes, instead of adapting the threshold dynamically for every data point, we adapted the individual temperature values against perturbations with a fixed threshold.

III. RESULTS

A. Population analysis

We gathered $N = 18024$ measurements of temperature during a period of five weeks, using the same device, but in 6 different locations (see Figure 1 (a)). We computed statistics of the ensemble data for all locations as well as individual locations (Table I). We observed an average forehead temperature of 35.49 ± 0.80 $^{\circ}\text{C}$. The mean and standard deviation of these

(a) without correction

(b) corrected against all perturbations

Fig. 1: Dataset of forehead temperature measurements relatively to time, taken during a period of five weeks. Statistics are shown in Table I. (a) without correction ($N = 18024$) (b) with the corrections computed in Section II-B (ambient, outside and time corrections) ($N = 10078$)

(a) histogram without correction

(b) histogram corrected against all perturbations

Fig. 2: Histogram of forehead temperature measurements, taken during a period of five weeks. Statistics are shown in Table I. (a) without correction ($N = 18024$) (b) with the corrections computed in Section II-B (ambient, outside and time corrections) ($N = 10078$)

Fig. 3: Bland-Altman plot of differences between uncorrected and corrected forehead temperature measurements ($N = 10078$). A negative value indicated that the corrected temperature is lower than the base temperature. The colors are the relative densities of measurements (the yellow color represent a higher density). (a) temperature correction values against mean temperature of corrected and uncorrected data points (b) temperature correction values against outside temperature

measurements did not change significantly across different locations and among different populations (see Table I for locations A-C). We suppose that this consistency is the result of measuring among the population of the same country (French-speaking Switzerland), with similar meteorological conditions and ethnicity. The distribution of temperatures seems to be skewed toward low temperatures (see Figure 2 (a)).

B. Tracking of individual temperature data

After having examined the data from an ensemble perspective, we now turn to the tracking of temperatures of individuals over time. For a subset of the data ($N = 12882$), we were able to gather individual anonymous tagging for every temperature measurement, along with all other features already analyzed in Section III-A. Results shown in Table I for variations of temperature of an individual over time indicates that the individual variation (0.54 ± 0.29 °C) was lower than the average variation of temperature in a population (0.80 °C).

C. Correction of environmental perturbations

We were able to retrieve, on a subset of the dataset ($N = 10078$), the meteorological data of the exact measurement location at the time when the measurement was taken. We

Fig. 4: Dataset of forehead temperature measurements relatively to external perturbations ((a) outside temperature, (c) room temperature, (e) hour of measurement), taken during a period of five weeks, before ((a), (c), (e)) and after correction for environmental effects ((b), (d)). The number of measurements after outdoor temperature correction is smaller than without correction due to the availability of meteorological data. The color represents the relative point density (yellow is a higher density).

then trained the regression models from Section II-B. There was a significant correlation between outside temperature ($R^2 = 0.19$) and forehead temperature in terms of Pearson correlation coefficient. Room temperature had a lower influence on the data ($R^2 = 0.12$), since some measurements were taken when the individual entered a particular building. A similar correlation ($R^2 = 0.19$) was found by comparing the measurements with their acquisition times within a 24-hour day. From the trained regression curves, we were able to subtract the differences representing the coupling of features for ambient, outside temperature and time of the day with the forehead temperature ($R^2 = 0$) (see Figure 1 (b)). After correction, the variation of temperatures across both population and time were reduced, lowering the CV by 0.36 % (Table I (b)). The histogram of the corrected temperature density appeared to follow a Gaussian curve (Figure 2 (b)). The corrections were in the order of 0.02 ± 0.51 °C, with most of the corrections being in the $[0$ °C; $+0.5$ °C] interval (see Figure 3 (a)). By design, the corrections did not alter the mean of the temperature measurements. The corrections do not depend on the subject's temperature but rather on the outside temperature (see Figure 3 (b)). For example, when the outside temperature was 5 °C, the applied correction ranged from 0.65 °C to 1.47 °C.

IV. DISCUSSION AND CONCLUSION

NCITs are remarkably noisy sources of data. Indeed, they exhibit variations due to different types of noise: intrinsic noise from the electronics; temperatures disparities due to sensor placement, skin type, skin color, and thickness, moisture, fat content; and finally, environmental conditions such as outside temperature, humidity, and sunlight exposure. Out of all these parameters, we chose to model the relationship between outside temperatures, ambient temperatures, acquisition time within the day and the measured forehead temperature as first and second-degrees polynomials. Using these models, we were able to subtract the influence of external perturbations and drastically reduce the co-linearity between these features. Thanks to the data analyzed in this study, we found out that the average forehead temperature was 35.49 ± 0.80 °C. This number was consistent across different places and matched to the values found in [23], [24], after proper scaling. Similarly to [4], we proposed a fever detection temperature threshold of the mean plus 3 times the standard deviation of the temperatures in our dataset (37.81 °C with the offset we applied, or 35.68 °C without calibration). After correction of external components, this threshold was reduced to 37.38 °C, which is consistent with the results from [22]. Using the same calculation, this threshold could even be lowered due to the smaller deviation of measurements in specific environments (see Table I for Location B-C). Our results were not in agreement with the research by [4], who did not find a significant correlation between air temperature and body temperature. However, it appears that, in that study, the subjects were already acclimatized to a constant room temperature. We believe that, in a real-world case, temperature screening devices are installed at the entrance of a building and the security policy do not give the incoming people enough time to acclimatize inside the building before measuring their forehead temperature.

Furthermore, we found that the deviation of the forehead temperature of one subject over time is lower than the deviation of the mean temperature in a population, due to disparities between subjects. These findings confirm the idea that every individual is biologically distinct in terms of temperature and skin type due to demographic and physiological factors [25]. Consequently, the limit threshold for fever can be set individually as well. Personalized thresholds can be lower than the static threshold and might have a significant advantage over the latter for fever screening. Consequently, we propose a personalized fever detection threshold of the mean plus 3 times the standard deviation of the temperature record for a given subject, after correction for external perturbations. Such correlations between perturbations and the measured forehead temperature reveal that corrections to the temperature or fever detection threshold can lead to a more accurate recognition of high temperature symptoms. Eventually, even after corrections, our method does not take into account temperature perturbations caused by factors such as occasional sports, sweating and moisture on the skin. However, other data sources and modalities can be integrated to the computation, such as humidity measurements, tracking data or personal schedule of the subject to compensate for other perturbations. Our method might as well exaggerate false positives (type I error). For example, the alert might be triggered if the temperature outside is very low and the subject is warmed by an external source like a car heating. Hence, the challenge presented in [15], [16] is only partially settled. Finally, like with any fever screening method, we cannot exclude some form of examiner bias. Indeed, feverish subjects would be prone to either stay at home or try to avoid the screening process and therefore would not appear in the statistics. In a time when disease outbreaks can spread very easily rapidly worldwide, fast disease screening solutions, such as fever screening, appear to be necessary. We hope that this research helps decision-makers and manufacturers to build more robust and intelligent tools that use multi-dimensional data in order to maximize both sensitivity and specificity of the screening protocol.

V. DECLARATIONS

Code and datasets are available online: <https://github.com/ashajkofci/CorrectionForeheadTemperature>. AS work for Coronasense, Martigny, Switzerland, who gathered and released the raw data.

VI. ETHICS STATEMENT

The authors used freely available anonymous data that was gathered with the consent of the owners of the measurement devices.

VII. REFERENCES

- [1] S. S. Evans, E. A. Repasky, and D. T. Fisher, "Fever and the thermal regulation of immunity: The immune system feels the heat," *Nature Reviews Immunology*, vol. 15, no. 6, pp. 335–349, 6 2015.

- [2] M. Yamaya, H. Nishimura, N. Lusamba Kalonji, X. Deng, H. Momma, Y. Shimotai, and R. Nagatomi, "Effects of high temperature on pandemic and seasonal human influenza viral replication and infection-induced damage in primary human tracheal epithelial cell cultures," *Heliyon*, vol. 5, no. 2, 2019.
- [3] V. Betta, F. Cascetta, and D. Sepe, "An assessment of infrared tympanic thermometers for body temperature measurement," *Physiol Meas*, vol. 18, no. 3, pp. 215–225, 1997.
- [4] H.-Y. Chen, A. Chen, and C. Chen, "Investigation of the Impact of Infrared Sensors on Core Body Temperature Monitoring by Comparing Measurement Sites," *Sensors (Basel)*, vol. 20, no. 10, 2020.
- [5] C. Trouilleau, B. Fièrque, S. Noblet, F. Giner, D. Pochic, A. Durand, P. Robert, S. Cortial, M. Vilain, J. L. Tissot, and J. J. Yon, "High-performance uncooled amorphous silicon TEC less XGA IRFPA with 17um pixel-pitch," in *SPIE Defense, Security, and Sensing*, Orlando, Florida, USA, 2009, 72980Q.
- [6] Q. H. Tran, D. Han, C. Kang, A. Haldar, and J. Huh, "Effects of Ambient Temperature and Relative Humidity on Subsurface Defect Detection in Concrete Structures by Active Thermal Imaging," *Sensors*, vol. 17, no. 8, p. 1718, 2017.
- [7] W. Minkina and D. Klecha, "Atmospheric transmission coefficient modelling in the infrared for thermovision measurements," *J. Sens. Sens. Syst.*, vol. 5, no. 1, pp. 17–23, 2016.
- [8] A. Curran, M. Klein, M. Hepokoski, and C. Packard, "Improving the accuracy of infrared measurements of skin temperature," *Extrem Physiol Med*, vol. 4, no. S1, A140, 2046-7648-4-S1-A140, 2015.
- [9] I. Pušnik and J. Drnovšek, "Infrared ear thermometers—parameters influencing their reading and accuracy," *Physiol. Meas.*, vol. 26, no. 6, pp. 1075–1084, 2005.
- [10] T. Fletcher, A. Whittam, R. Simpson, and G. Machin, "Comparison of non-contact infrared skin thermometers," *Journal of Medical Engineering & Technology*, vol. 42, no. 2, pp. 65–71, 2018.
- [11] M. Chudecka, A. Lubkowska, and A. Kempieńska-Podhorodecka, "Body surface temperature distribution in relation to body composition in obese women," *J Therm Biol*, vol. 43, pp. 1–6, 2014.
- [12] C. B. Mogensen, L. Wittenhoff, G. Fruerhøj, and S. Hansen, "Forehead or ear temperature measurement cannot replace rectal measurements, except for screening purposes," *BMC Pediatrics*, vol. 18, no. 1, p. 15, 2018.
- [13] M. S. Ganio, C. M. Brown, D. J. Casa, S. M. Becker, S. W. Yeargin, B. P. McDermott, L. M. Boots, P. W. Boyd, L. E. Armstrong, and C. M. Maresh, "Validity and Reliability of Devices That Assess Body Temperature During Indoor Exercise in the Heat," *J Athl Train*, vol. 44, no. 2, pp. 124–135, 2009.
- [14] A. Dante, I. Franconi, A. R. Marucci, C. M. Alfes, and L. Lancia, "Evaluating the Interchangeability of Forehead, Tympanic, and Axillary Thermometers in Italian Paediatric Clinical Settings: Results of a Multicentre Observational Study," *Journal of Pediatric Nursing*, vol. 52, e21–e25, 2020.
- [15] U. Erenberk, E. Torun, E. Ozkaya, S. Uzuner, A. D. Demir, and R. Dundaroz, "Skin temperature measurement using an infrared thermometer on patients who have been exposed to cold," *Pediatr Int*, vol. 55, no. 6, pp. 767–770, 2013.
- [16] S.-H. Hsiao, T.-C. Chen, H.-C. Chien, C.-J. Yang, and Y.-H. Chen, "Measurement of body temperature to prevent pandemic COVID-19 in hospitals in Taiwan: Repeated measurement is necessary," *Journal of Hospital Infection*, vol. 105, no. 2, pp. 360–361, 2020.
- [17] J. Goldberg and R. J. Rosati, "System and method for temperature determination and calibration in a biomedical probe," U.S. Patent 5150969A, 1992.
- [18] E. Lehmuskallio, J. Hassi, and P. Kettunen, "The skin in the cold," *International Journal of Circumpolar Health*, vol. 61, no. 3, pp. 277–286, 2002.
- [19] V. Bernard, E. Staffa, V. Mornstein, and A. Bourek, "Infrared camera assessment of skin surface temperature – Effect of emissivity," *Physica Medica*, vol. 29, no. 6, pp. 583–591, 2013.
- [20] E. M. Williams, A. I. Heusch, and P. W. McCarthy, "Thermal screening of facial skin arterial hot spots using non-contact infrared radiometry," *Physiol. Meas.*, vol. 29, no. 3, pp. 341–348, 2008.
- [21] R. H. Myers, *Classical and Modern Regression With Applications*, 2nd edition. Australia; Pacific Grove, CA: Brooks/Cole, 2000, 488 pp.
- [22] J. S. Hausmann, R. Berna, N. Gujral, S. Ayubi, J. Hawkins, J. S. Brownstein, and F. Dedeoglu, "Using Smartphone Crowdsourcing to Redefine Normal and Febrile Temperatures in Adults: Results from the Feverprints Study," *J GEN INTERN MED*, vol. 33, no. 12, pp. 2046–2047, 2018.
- [23] D. K.-k. Ng, C.-h. Chan, E. Y.-t. Chan, K.-l. Kwok, P.-y. Chow, W.-F. Lau, and J. C.-S. Ho, "A brief report on the normal range of forehead temperature as determined by noncontact, handheld, infrared thermometer," *Am J Infect Control*, vol. 33, no. 4, pp. 227–229, 2005.
- [24] H. Apa, S. Gözmen, N. Bayram, A. Çatkoğlu, F. Devrim, U. Karaarslan, İ. Günay, N. Ünal, and İ. Devrim, "Clinical accuracy of tympanic thermometer and noncontact infrared skin thermometer in pediatric practice: An alternative for axillary digital thermometer," *Pediatr Emerg Care*, vol. 29, no. 9, pp. 992–997, 2013.
- [25] Z. Obermeyer, J. K. Samra, and S. Mullainathan, "Individual differences in normal body temperature: Longitudinal big data analysis of patient records," *BMJ*, vol. 359, 2017.