

Esophageal Thermal Exposure to Hot Beverages: A Comparison of Metrics to Discriminate Distinct Consumption Habits



Daniel R.S. Middleton¹, Shuang-Hua Xie², Liacine Bouaoun¹, Graham Byrnes¹, Guo-Hui Song³, Joachim Schüz¹, Wen-Qiang Wei², and Valerie A. McCormack¹

Abstract

Background: Hot beverage consumption is a probable risk factor for esophageal squamous cell carcinoma (ESCC). No standardized exposure assessment protocol exists.

Methods: To compare how alternative metrics discriminate distinct drinking habits, we measured sip temperatures and sizes in an international group of hot beverage drinkers in France ($n = 20$) and hot porridge consumers ($n = 52$) in a high ESCC incidence region of China. Building on the knowledge that sip size and temperature affect intraesophageal liquid temperature (IELT), IELTs were predicted by modeling existing data, and compared with first sip temperature and, across all sips, mean temperature and sip-weighted mean temperature.

Results: Two contrasting exposure characteristics were observed. Compared with the international group, Chinese porridge consumers took larger first sips [mean difference +17 g; 95% confidence interval (CI), 13.3–20.7] of hotter

(+9.5°C; 95% CI, 6.2–12.7) liquid, and their mean sip size did not vary greatly across sips, but the former groups increased in size as temperature decreased. This resulted in higher predicted IELTs (mean 61°C vs. 42.4°C) and sip-weighted temperatures (76.9°C vs. 56°C) in Chinese porridge consumers, and compared with first sip and mean temperature, these two metrics separated the groups to a greater extent.

Conclusions: Distinguishing thermal exposure characteristics between these groups was greatly enhanced by measuring sip sizes. Temperature at first sip alone is suboptimal for assessing human exposure to hot foods and beverages, and future studies should include sip size measurements in exposure assessment protocols.

Impact: This study provides a logistically feasible framework for assessing human exposure to hot beverages.

Introduction

Consumption of hot beverages is a putative risk factor for esophageal cancer—the seventh commonest cancer worldwide, with the sixth highest number of cancer deaths in 2018 (1). Reports dated as early as 1939 speculate thermal irritation as a risk factor for esophageal cancer (2). In 2016, the International Agency for Research on Cancer (IARC), within their Monograph Program on the Identification of Carcinogenic Hazards to

Humans, classified drinking "very hot beverages" (>65°C) as "probably carcinogenic to humans" (Group 2A; ref. 3) and recently listed it as a high priority for reevaluation (4). Positive associations between hot beverage consumption and esophageal cancer—namely squamous cell carcinoma (ESCC)—have been reported in most of the world's notable high incidence regions, including prospective evidence from the Chinese Kadoorie Biobank cohort (5) and the Iranian Golestan cohort, the latter reporting an HR of 1.6 for first sip temperature of tea drinking >60°C compared with <60°C (6).

Thermal injury exposure assessment methods are expected to have a large impact on the magnitude of its association with ESCC; however, there is no "gold-standard" exposure assessment method or metric. Previous epidemiologic studies, with the exception of the Golestan cohort, share a common limitation—they relied on self-reported perceived drinking temperatures, e.g., "very hot" versus "warm" beverages. In the few studies to objectively measure temperatures, either for descriptive purposes or in relation to disease, temperature at first sip was commonly utilized as the metric of thermal exposure (6–8). However, a reliable metric of thermal exposure might quantify the heat transfer from the swallowed liquid to the esophageal mucosa, i.e., heat being a function not only of temperature of the liquid, but also its volume. Indeed, in a 1972 publication in *Gut* by de Jong and colleagues (9), the authors performed a unique study of the relationship between beverage temperature and temperature measured in the lower esophagus during drinking through placement of a temperature sonde within the esophagus. In addition to a positive relationship between what the authors called

¹Section of Environment and Radiation, International Agency for Research on Cancer (IARC), Lyon, France. ²National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. ³Cancer Institute/Hospital of Ci County, Cixian, China.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

D.R.S. Middleton and S.-H. Xie contributed equally to this article and share first authorship.

Corresponding Authors: Daniel R.S. Middleton, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France. Phone: 33-472-738-310; E-mail: middletond@fellows.iarc.fr; and Wen-Qiang Wei, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 17 Panjiayuan, Chaoyang District, Beijing, 100021 China. Phone: 8610-8778-8661; E-mail: weiqw@cicams.ac.cn

Cancer Epidemiol Biomarkers Prev 2019;28:2005–13

doi: 10.1158/1055-9965.EPI-19-0856

©2019 American Association for Cancer Research.

"intra-esophageal temperature" and initial drinking temperature, they observed the expected positive association with sip volume, for example, with intraesophageal temperature reaching approximately 46°C upon either drinking 15 mL of a 65°C coffee or 20 mL of 60°C coffee. Despite these important findings, we are not aware of a thermal exposure measurement protocol that has quantitatively incorporated both sip temperature and volume.

In terms of the practicalities of measuring beverage drinking temperatures, studies have employed a variety of methods. In a UK study (10), a single tea temperature was recorded at first sip—defined as the point at which the beverage had cooled enough to drink in successive gulps. In a study of coffee drinkers (11), consumers were given a choice of hot or cooler coffee and presented with options to alter the temperature. A single temperature was recorded at the desired sipping temperature. In a high ESCC incidence region of Brazil, maté temperatures were assessed by single temperature measurements of maté inside the gourd (12). In the abovementioned Golestan cohort (8), participants were asked to sip their tea at 5°C intervals—determined from a proxy cup—until the desired drinking temperature was reached and recorded. In another high ESCC risk area in Tanzania, a cross-sectional study (7) of tea drinking habits in Kilimanjaro employed the Golestan protocol, additionally recording the times and temperature when half and the entire cup had been drunk, which demonstrated large differences in cooling curves by type of drink. The Iranian study observed agreement between self-reported and measured drinking temperatures (8). First sip temperature was also employed in a recent Kenyan cross-sectional study (13) of tea drinking temperatures, evidence that this metric is beginning to be routinely adopted.

Measuring temperatures throughout the drinking episode is possible; and accurate exposure assessment is necessary for prospective studies, cross-sectional community surveys (i.e., to enable reliable inter- and intrapopulation comparisons), and studies investigating the as yet unknown carcinogenic mechanism of thermal injury. Therefore, as part of ongoing investigations into the etiology of ESCC, we:

(I) developed an improved measurement protocol of hot beverage/semiliquid thermal exposure which incorporates both sip volume and temperature;

(II) used this protocol to record thermal exposure measurements in two groups of people who were expected to differ substantially in their thermal exposures: Group A (low exposure) of hot beverage drinking by Europe-based international researchers; Group B (high exposure) of hot porridge consumption in a high ESCC incidence region of China;

(III) examined how the following metrics of thermal exposure discriminated Groups A and B: the temperature at first sip (the most commonly used metric), mean drinking temperature (across sips), sip-weighted mean drinking temperature, and, using de Jong and colleagues' (9) results, predicted intraesophageal liquid temperatures (IELT) as a function of beverage temperature and sip size.

Materials and Methods

Ethical approval and consent to participate

Ethical approval for the study was received in China from the Ethics Committee of the Cancer Institute and Hospital, Chinese Academy of Medical Sciences (approval number NCC2017YJZ-

001), and at the IARC (IEC Project No. 17-04-A1). Written-informed consent was provided by all participants, and the study was conducted in accordance with the International Ethical Guidelines for Health-related Research Involving Humans developed by the Council for International Organizations of Medical Sciences.

Study design and settings

Thermal exposure metrics in two groups of individuals, expected to differ considerably, were compared. Group A was comprised of volunteer staff, invited via internal email, at IARC in Lyon, France, during the months of December 2017 to January 2018 (8 am–6 pm) who drank a hot beverage in an indoor office with an ambient temperature of 23°C. Volunteers were asked to choose from a selection of beverages (black tea, green tea, or black coffee with or without the addition of sugar only). For Group B, a subset of 52 participants of an ongoing population-based cohort study (14), residing in two rural villages in Cixian County, Hebei Province, China (Supplementary Fig. S1), were recruited. During November 2017 and January 2018, participants were visited at their homes when taking their morning porridge. Ambient temperature was recorded where the porridge was being consumed (sometimes outdoors) and ranged from 5 to 23°C.

Participants in both groups underwent an interview-administered questionnaire to obtain age, sex, and self-reported perceived hot beverage consumption habits.

Beverage temperature and sip size measurements

In China (Group B), participants were asked to prepare two identical servings of porridge in their usual manner. This watery porridge is similar to the *congee* consumed in other regions of China, but made with maize instead of rice. One bowl was placed on a digital weighing balance for consumption and the other used for temperature measurement with a stainless steel temperature probe (Checktemp, HANNA Instruments). At the time of serving, a digital timer was started, and participants were instructed to begin eating their porridge when they normally would and in their usual manner. The time and temperature of the first sip were recorded, and the weight of the porridge before and after the sip to estimate sip volume (assuming 1 g of porridge \approx 1 mL). This was repeated in what fieldworkers judged to be the middle of the meal and just before the end—giving measurements for three sips in total, in addition to the time taken to complete the meal. At IARC (Group A), the same protocol was employed for tea/coffee, with the exception that measurements were recorded for every sip throughout the drink to enable a detailed investigation of exposure metrics. To ease the capture of temperature at every sip, a Bluetooth-enabled thermometer (BlueTherm One, ETI Ltd.) with a rapid response penetration probe and smartphone application (HACCP Mobile) were used—capable of logging timestamped temperature readings on demand. This device is comparable with the one used in China, with the addition of data logging capability, and near-perfect agreement was found between both instruments when used simultaneously for validation (mean absolute temperature difference $<0.3 \pm 0.42^\circ\text{C}$ between readings, $n = 15$ measurements; temperature range: 95–59°C). In addition, the procedure was repeated once more for each IARC participant on a different day to assess intraparticipant variability of different exposure metrics.

Prior to fieldwork, we preliminarily tested the performance of two alternative types of thermometer (Supplementary Fig. S2) on

boiling water left to cool—a noncontact infrared thermometer (TFI 250, Ebro, Xylem Analytics) and a submersible temperature sensor and data logger (Tempo Disc, Blue Maestro Ltd.). However, the former was susceptible to interference from evaporating vapor, and the latter had slow instrumental response due to insulation from a waterproof casing (Supplementary Fig. S3).

Exposure metrics and statistical analysis

In addition to the previously used exposure metrics of temperature and time at first sip and mean temperature across the drink, we investigated the utility of previously untested exposure metrics—mean temperature weighted for sip size and predicted IELT.

To estimate IELT, we extracted data from de Jong and colleagues' previous investigation of measured "intra-esophageal temperature" (9) and used it to model IELT. We note that the nomenclature "intra-esophageal liquid temperature" is used henceforth so as not to imply that the metric refers to tissue temperature, but likely reflecting the temperature of the liquid passing through the esophagus. We modeled IELT as a function of sip size (V_s in g) and sip temperature (T_s in °C), constraining IELT to be body temperature (37°C) when no sip was taken ($V_s = 0$), and having a theoretical asymptote at temperature T_s for large volumes, as follows: $IELT = 37 - (T_s - 37) + 2(T_s - 37) / (1 + \exp(\beta V_s (T_s - 37)))$ for which predicted values are shown in Supplementary Table S1. However, its fit (Supplementary Fig. S4) could only be examined in the range of V_s from 5 to 20 g and T_s from 55 to 65°C, thus, as most sipping temperatures (T_s) exceed 65°C, many predictions of IELT are beyond the observed volume and

temperature range. Random errors were then added to point IELT predictions, drawn from a normal distribution with mean = 0 and SD = 2.69, where 2.69 is the estimated between-participant SD in IELT for a given T_s and V_s , based on that observed for Jong's measured IELT values. A schematic of the different steps of deriving IELT is presented in Fig. 1. Sip-weighted mean temperature for each person was calculated across a single drinking episode as: $(\sum T_s \times V_s) / \sum V_s$.

We estimated differences in the aforementioned metrics of drinking habits between the two exposure groups (China and IARC), using an unequal variance *t* test (Welch test) to compare sample means for continuous variables, and a χ^2 test for categorical variables. For the analysis of changes in sip weight over a drinking sitting, a multilevel model was fitted with sips nested within an individual's sitting, and taking a natural log-transformation to normalize the distribution of sip weights, thus the change in sips by temperature is reported on a relative scale. We also examined the correlation [Pearson (*r*)] of drinking metrics (natural log-transformed) with each other (separately for the two groups), and, for IARC participants only, correlations of a given metric between repeat sittings for the same participant.

Results

Study participant characteristics and self-reported hot food/beverage consumption habits

Group A at IARC had 20 participants, each of whom had two drinks measured, giving 406 time-temperature-sip size measurements in total, and Group B from China had 52 participants who

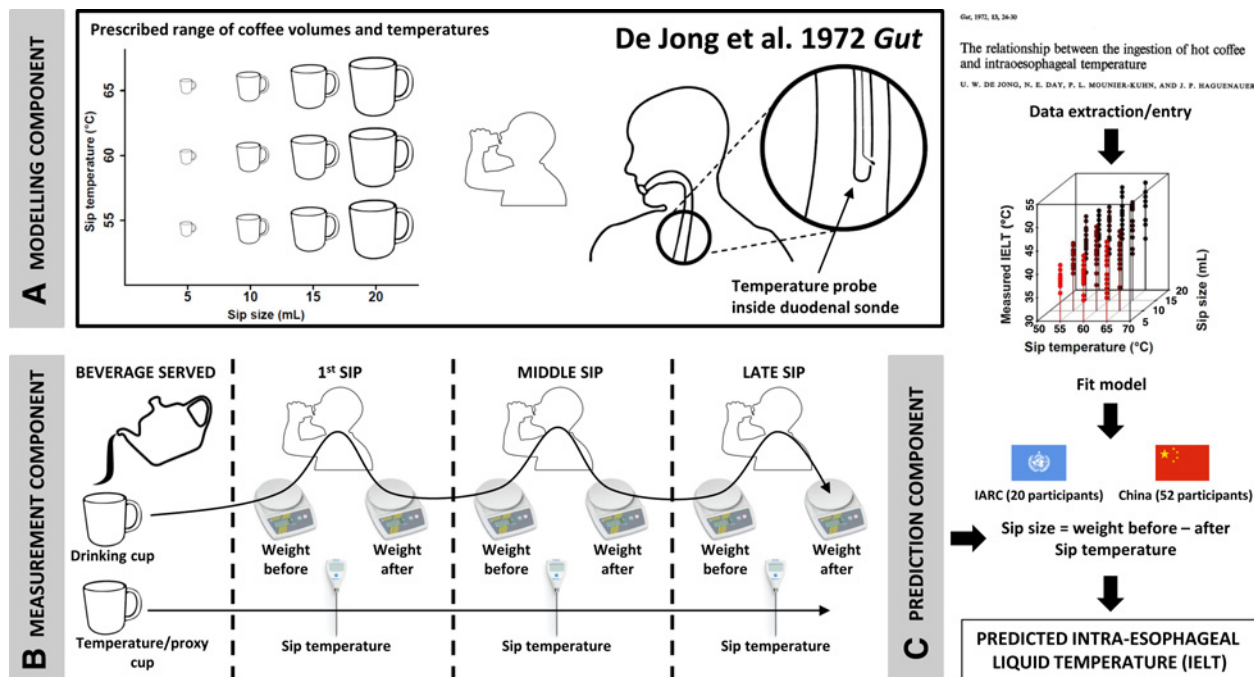


Figure 1. Schematic of the different study components. **A**, Existing data (De Jong and colleagues; ref. 9) were modeled to derive the relationship between measured IELT and sip temperature and volume (fit examined for volumes 5–20 mL and temperatures 55–65°C). **B**, Porridge (China) and tea/coffee (IARC) temperatures and sip sizes were measured in two study groups (at every sip for IARC participants). **C**, The fitted model derived in **A** was used to predict IELT from measurements made in **B**.

Downloaded from <http://aacrjournals.org/cebp/article-pdf/28/12/2005/2285822/2005.pdf> by guest on 07 October 2024

Table 1. Study participant characteristics and self-reported hot porridge/ beverage exposure variables and consumption habits

Characteristic		Group A Hot beverage, IARC, n (%)	Group B Hot porridge, China, n (%)
Number of participants		20	52
Sex	Male	10 (50)	14 (27)
	Female	10 (50)	38 (73)
Age (years)	Mean (SD)	32 (7.8)	59 (7.8)
	20–29	7 (35)	—
	30–40	13 (65)	—
	40–49	—	8 (16)
	50–59	—	12 (23)
	60–64	—	12 (23)
	65–70	—	19 (37)
Broad region of ethnic origin	Asia or Middle East	6 (30)	52 (100)
	Europe	11 (55)	—
	North America	2 (10)	—
	North Africa	1 (5)	—
	Village	Guoxiaotun	—
Food/beverage being consumed	Liueying	—	37 (71)
	Porridge	—	52 (100)
	Black coffee	3 (15)	—
	Green tea	13 (65)	—
Beverage/porridge drinking temperature ^a	Black tea	4 (20)	—
	Extremely hot	0 (0)	0 (0)
	Very hot	4 (20)	1 (2)
	Hot	12 (60)	30 (58)
Beverage/porridge drinking speed ^a	Warm	4 (20)	21 (40)
	Extremely fast	0 (0)	1 (2)
	Fast	4 (20)	32 (62)
	Normal	13 (65)	19 (36)
Burned mouth on beverage/porridge ^a (occurrences in last month)	Slow	3 (15)	0
	0	2 (10)	52 (100)
	1	8 (40)	0
	2	7 (35)	0
	4	3 (15)	0

^aSelf-reported/perceived.

each consumed one bowl of porridge, with 156 measurements in total. The study groups differed in their demographic characteristics (Table 1). An equal sex and younger age distribution (all <40 years) was present for the IARC volunteers compared with those from China, who were >70% female and all >40 years old. There was a greater ethnic and cultural diversity among the IARC group, which included 55% White Europeans and volunteers originating from 4 different continents. Conversely, all Chinese participants were residents of the study area. Self-reported perceived consumption temperatures were higher at IARC. Although among Chinese participants approximately 60% categorized themselves in "hot" and approximately 40% in "warm," 20% of the IARC volunteers categorized themselves in "very hot," 60% in "hot," and 20% selected "warm." Similarly, no participant in China recalled having burned their mouth in the last month, whereas 90% of IARC volunteers selected at least once. Chinese participants opted toward faster self-reported consumption speeds compared with IARC—62% "fast" and 20% "fast," respectively.

Description of thermal exposures

Visual observations were made of the manner in which individuals consumed their beverages. After pouring into mugs, IARC beverage drinkers typically waited before testing with their lips how hot drinks were before proceeding to take sips when comfortable. In China, porridge was served into bowls directly from

the heat source. The porridge was of a watery consistency and drunk directly from the bowl like a beverage. Chopsticks were used to sweep the liquid into the path of the mouth, particularly toward the end of the meal. The visible temperature of the liquid at first sip, sometimes immediately after serving, was notable, as porridge could still be seen bubbling/steaming when visibly large sips were taken. Following these observations, a subjective assessment was made that thermal exposures were much higher in Chinese porridge consumers.

Measured hot food/beverage consumption habits and predicted intraesophageal temperatures

Agreeing with visual observations, statistically significant differences were found in objectively measured characteristics of thermal exposure between IARC beverage drinkers and Chinese porridge consumers (Table 2). IARC beverage drinkers started drinking after a mean 7.7 ± 3.5 minutes after pouring, at a mean temperature of $70.5 \pm 7.8^\circ\text{C}$, 9.5°C [95% confidence interval (CI), 6.2–12.7] cooler than Chinese porridge consumers ($79.9 \pm 8.9^\circ\text{C}$), who took their first sip after a mean of just 2.1 ± 1.8 minutes after serving, 5.6 (95% CI: 6.8, 4.3) minutes sooner than the IARC volunteers. Chinese drinkers also finished their porridge more quickly, when it had cooled to a mean temperature of $75 \pm 8.3^\circ\text{C}$, still 4.5°C hotter than the mean first sip temperature at IARC. Mean drinking durations (from first sip) were 6.3 ± 2.6 and 13.9 ± 4.4 minutes for China and IARC, respectively. This is of further note given the large difference in mean total serving size between both types of beverage, which was approximately 380 g for a serving of porridge and 190 g for a mug of tea or coffee. This resulted in a crude mean across-sip temperature of $77.1 \pm 8.4^\circ\text{C}$ in China, compared with $58.9 \pm 4.3^\circ\text{C}$ at IARC, i.e., a difference of 18.3°C (95% CI, 15.6–20.9), double the difference in the mean temperature at first sip.

Despite consuming at a higher temperature range, on average Chinese porridge consumers also took larger sips than IARC beverage drinkers, both at first sip (mean 20.9 g vs. 3.9 g) and across all sips (mean 22.3 vs. 16.1 g), but the variability around these mean sip sizes was also larger (e.g., SD 8.5 vs. 4.5 g). This resulted in significantly higher predicted IELTs (e.g., mean $61.0 \pm 8.3^\circ\text{C}$ vs. $42.4 \pm 2.1^\circ\text{C}$; mean difference 18.3, 95% CI, 16.1–20.9) for Chinese porridge consumers compared with IARC beverage drinkers. Sip-weighted mean temperatures ($76.9 \pm 8.3^\circ\text{C}$ vs. $56 \pm 4^\circ\text{C}$; mean difference 21.0°C, 95% CI, 18.3–23.5) were also higher in this group. A characteristic drinking behavior was observed among IARC volunteers, whereby sip size increased by 11.4% (95% CI, 10.8–11.9) per 1°C decrease in temperature. This is illustrated in Fig. 2A and B, in which sip sizes can be seen to increase with decreasing sip temperatures, resulting in stable predicted IELTs throughout the drink as IARC volunteers took small sips in reaction to hot liquid early on. This tendency of increasing sip size was weaker in magnitude [1.8% (95% CI, 0.8–2.8) increase per degree decrease] among Chinese porridge consumers (Fig. 2C).

To visualize the difference in group separation according to alternative metrics, we plotted each individual's sip temperature in proportion to its sip size, with each person represented horizontally (Fig. 3A). For IARC beverage drinkers (blue points), sip sizes (graduated by point size) can be seen increasing as the drink cools, whereas the effect is small in China (red points). In Fig. 3A, individuals on the y axis are ordered by the commonly used metric of temperature at first sip, where there is appreciable overlap in

Table 2. Measured and derived characteristics of thermal exposures during consumption and their comparisons between two contrasting thermal exposure circumstances: hot beverage drinking in IARC volunteers and porridge consumption in Chinese participants

	Group A Hot beverage, IARC Mean ± SD unless stated otherwise	Group B Hot porridge, China Mean ± SD unless stated otherwise	Groups B-A China-IARC Difference (95% CI)	Difference	
				P	t ^a
Number of individuals	20	52	—	—	—
Number of drinking sittings	40	52			
Number of sip + temperature measurement sets	406	156			
Drinking characteristics					
Temperatures (°C)					
At pouring/serving	92.7 ± 2.4	82.0 ± 9.6	-10.7 (-13.6, -7.9)	<0.0001	-7.6
At first sip temperature	70.5 ± 6.8	79.9 ± 8.9	9.5 (6.2-12.7)	<0.0001	5.8
Mean across sips	58.9 ± 4.3	77.1 ± 8.4	18.3 (15.6-20.9)	<0.0001	13.6
At last sip	52.0 ± 3.9	75.0 ± 8.3	23.0 (20.4-25.6)	<0.0001	17.7
Sip size (g)					
At first sip temperature	3.9 ± 2.8	20.9 ± 13.1	17.0 (13.3-20.7)	<0.0001	9.1
Mean across sips	16.1 ± 4.5	22.3 ± 8.5	6.0 (3.3-8.7)	<0.0001	4.4
At last sip	26.6 ± 9.3	28.0 ± 21.7	1.4 (-5.3-8.1)	0.67	0.42
Drinking times (minutes)					
First sip time since pouring	7.7 ± 3.5	2.1 ± 1.8	-5.6 (-6.8, -4.3)	<0.0001	-9.1
Drink duration since pouring	21.6 ± 4.4	8.5 ± 3.5	-13.2 (-14.9, -11.5)	<0.0001	-15.4
Drink duration since first sip	13.9 ± 4.4	6.3 ± 2.6	-7.6 (-9.2, -6.0)	<0.0001	-9.7
Derived metrics					
Sip-weighted mean temperature (°C)					
Across sips	56.0 ± 4.0	76.9 ± 8.3	21.0 (18.3-23.5)	<0.0001	15.9
Predicted IELT (°C) ^b					
At first sip	40.6 ± 3.0	61.9 ± 11.0	21.3 (18.1-24.5)	<0.0001	13.3
Mean across sips	42.4 ± 2.1	61.0 ± 8.3	18.5 (16.1-20.9)	<0.0001	15.4
At last sip	42.1 ± 3.8	61.9 ± 11.0	19.8 (16.6-23.1)	<0.0001	12.1
Max during drink	47.0 ± 2.7	68.2 ± 9.7	21.3 (18.4-24.1)	<0.0001	15.1
IELT categories, mean across sips [N (column %)]					
37-<40	6 (15)	0	—	<0.0001 ^c	—
40-<42	11 (28)	0			
42-<44	13 (33)	0			
44-<46	9 (22)	0			
46-<48	1 (2)	1 (2)			
48-<50	0	3 (6)			
50-<53	0	6 (11)			
53+	0	42 (81)			
Variation in sip size by descending drinking temperature					
Drinking temperature (°C)	Sip size (g)	Sip size (g)			
80+	2.9 (1.4)	19.7 (16.1)	—	—	—
75-<80	2.6 (1.3)	24.3 (22.7)			
70-<75	4.5 (2.6)	22.2 (9.8)			
65-<70	7.6 (3.6)	23.0 (11.0)			
60-<65	11.5 (5.8)	27.0 (3.5)			
55-<60	16.8 (8.6)	30.9 (8.7)			
<55	20.8 (10.0)	—			
% increase in sip size per 1°C decrease in sip temperature	11.4 (10.8-11.9)	1.8 (0.8-2.8)	—	<0.0001	11.4

NOTE: The exposure metrics predominantly assessed are shown in bold type in the left column.

^aAccounts for unequal variance (Welch test).^bPredicted IELT values >~53°C are above the fitted range and should be interpreted with caution.^cχ² test.

exposure rank between the two groups. Alternatively, when ranked by the predicted IELT at first sip (Fig. 3B), the two groups are almost completely distinguishable from one another. The two metrics give different ranks as many IARC first sips were very hot, but they were too small in volume to have an appreciable effect on thermal exposure as measured by IELT or a sip-weighted temperature. The same degree of separation between the two groups was achieved by ranking by mean predicted IELT across sips, but to enable like-for-like comparison, first sip IELT is plotted in Fig. 3B.

Another reflection of the greater discrimination after incorporation of both size sip and temperature into metrics is that among the key metrics (bold in Table 2), the largest between-group difference was for sip-weighted temperature [China-IARC difference 21.0°C (18.3-23.5), *t* = 15.9], similar to the difference in mean IELT [18.5°C (16.1-20.9), *t* = 15.4] or mean temperature (18.3°C, *t* = 13.6), and all were double that of the temperature at first sip (9.5°C, *t* = 5.8).

To assess how alternative exposure metrics varied between-sittings (within-person), correlation was investigated with repeat

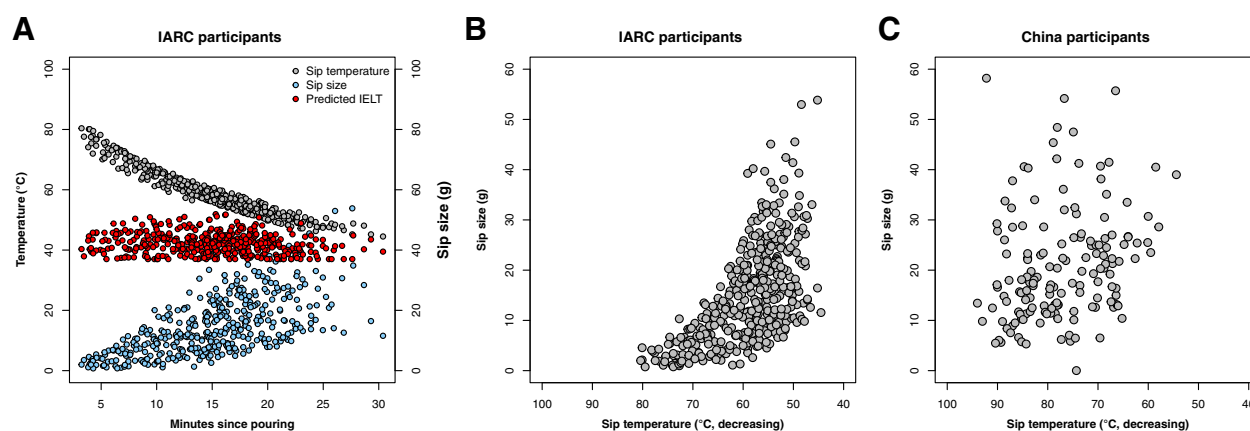


Figure 2.

Drinking habits and exposure metrics throughout the course of drinking. **A**, Sip temperatures, sizes, and predicted IELT changes throughout drinking durations (IARC hot beverage drinkers). Sip sizes in response to decreasing sip temperatures throughout the course of consumption for IARC beverage drinkers (**B**) and Chinese porridge consumers [**C**; two points omitted ($x = 78.2^{\circ}\text{C}$, $y = 121\text{ g}$; $x = 81.3^{\circ}\text{C}$, $y = 116.9\text{ g}$) to allow matching axis limits].

sitting measurements, which was possible only for IARC volunteers. Overall, metrics incorporating (i) sip size and (ii) measurements from multiple sips in their derivation had slightly stronger correlations than those with only temperature (e.g., within-person correlation of mean predicted IELT: $r = 0.76$ for both vs. mean temperature: $r = 0.71$) or measurements for a single sip (e.g., $r = 0.37$ for first sip predicted IELT and $r = 0.69$ for first sip temperature). Furthermore, intraperson differences in first sip temperature of up to 10°C were observed between sittings. Intraperson correlations are shown in Table 3 and plotted in Supplementary Fig. S5. Between-metric correlations are also reported in Table 3 and plotted in Supplementary Fig. S6; they tended to be stronger for the Chinese sample because of the smaller within-person variation in sip size.

The relationships between perceived self-reported drinking temperatures and objective metrics are presented separately for IARC beverage drinkers and Chinese porridge consumers in Supplementary Table S2. Although temperatures were perceived to be higher among IARC beverage drinkers (Table 1) and the contrary was measured objectively, within each group, agreement was found between perceived and objective metrics. Relative to "warm," participants reporting "hot" had 1.9°C and 5.1°C higher mean IELTs at IARC and in China, respectively; 2.8°C and 6.9°C higher for those reporting "very hot" (P for trend: 0.047 and 0.03 in Groups A and B, respectively).

Discussion

Despite longstanding and growing evidence of the probable carcinogenicity of hot beverage consumption to the squamous cells of the esophagus, there exists no validated protocol for routine exposure assessment. There are no studies having comparatively assessed alternative metrics of thermal exposure or having verified whether a commonly used metric, temperature at first sip size, accurately measures thermal exposure in the esophagus. To our knowledge, our study is the first to objectively measure both sip temperatures and sip sizes throughout the duration of hot beverage/semiliquid consumption and use them to derive and comparatively assess different exposure metrics. In summary, although the commonly used

temperature at first sip indeed discriminates two groups of hot beverage/porridge drinkers to some extent, use of mean temperature, and more so, sip-weighted mean temperature or mean IELT demonstrated much greater discriminatory ability. Hence, these novel metrics provided a more comprehensive thermal exposure assessment, and their application will reduce exposure misclassification in future studies. However, it should be noted that these two metrics do not account for sip size and temperature in similar manners. IELT is a metric that incorporates differentials in both average sip size and temperature between individuals. In contrast, the sip-weighted temperature is an average temperature alone (weighted to the larger sip sizes relative to all sips for that individual) and does not capture differences in average sip size between individuals. Thus, had the IARC and Chinese groups been consuming at similar temperatures, sip-weighted temperature would not have captured the large difference in thermal exposure experiences brought about by the larger sip sizes of the latter group.

We observed two contrasting thermal exposures in our study by comparing a group of Chinese hot porridge consumers, who took larger sips of hotter liquid sooner after serving, with a mixed group of international hot beverage drinkers, who waited longer to start drinking and took smaller sips which increased in size as temperature decreased, thus exhibiting a protective behavior which reduced their predicted esophageal thermal exposure.

In this study, we found that the previously employed metric of temperature at first sip is not sufficient in capturing relevant exposure contrasts, depicted by the large degree of overlap between both groups in Fig. 3A. In the knowledge that sip size is an important determinant of esophageal thermal exposure (9), by incorporating it, either by deriving predicted IELT, or calculating sip-weighted temperatures, it was possible to better characterize thermal exposures and make more meaningful comparisons between populations. A first-sip IELT discriminated the two groups almost as well as mean IELT or sip-weighted temperature; however, the comparatively weak within-person correlation ($r = 0.37$ at IARC) found for first sip IELT suggests that, to reflect habitual exposures, measurements at several time points throughout the drink/meal should be taken.

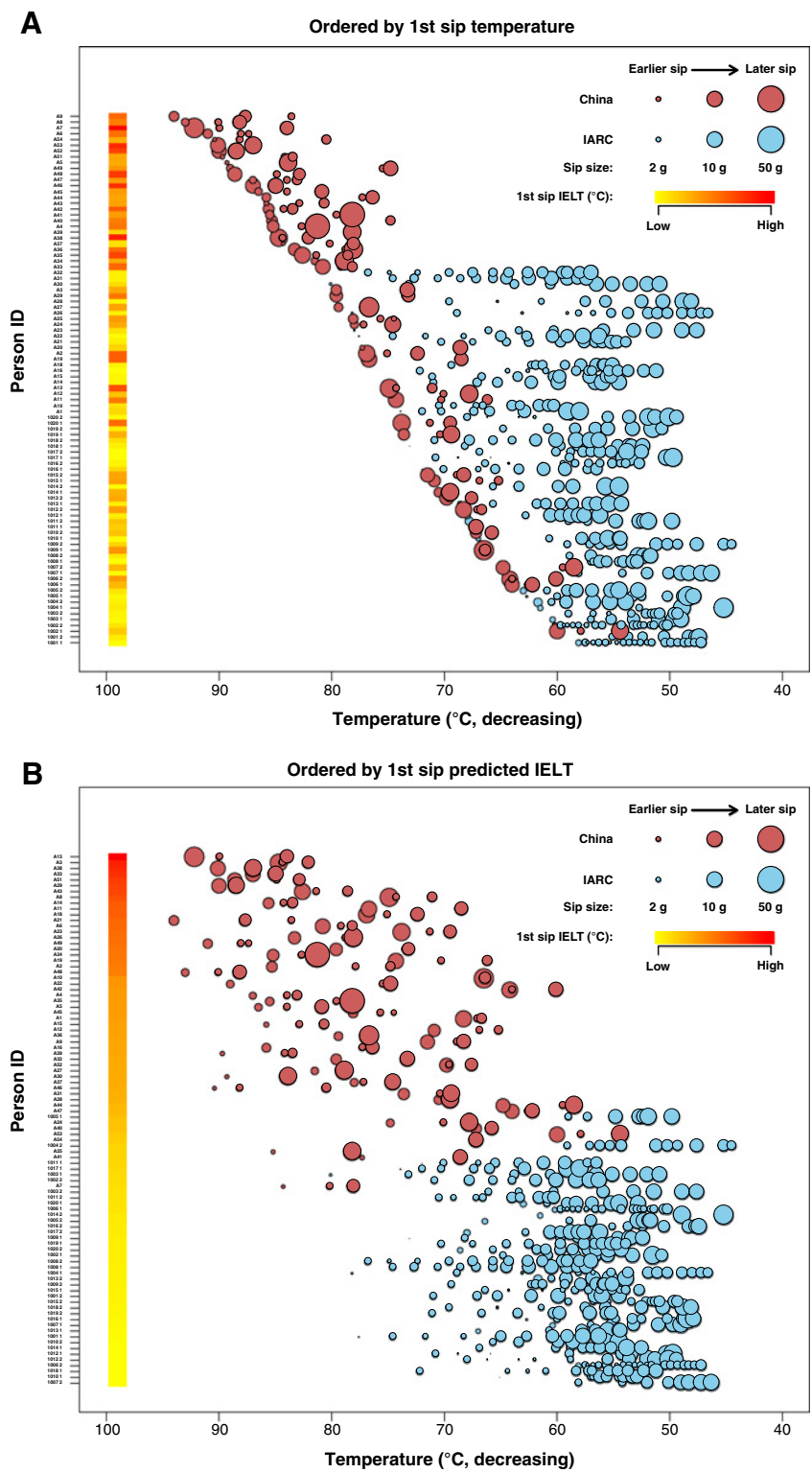


Figure 3. Graphical representations of exposure classifications for alternative metrics. Individual sip temperatures are plotted throughout the course of consumption—i.e., decreasing temperature—for each person (two sittings per IARC volunteer) and graduated by sip size. Individuals are ranked by first sip temperature (**A**) and first sip-predicted IELT (**B**), illustrating the different exposure classification achieved by each metric.

This is a small logistical/time addition to the measurement protocol, while providing valuable information—as shown here—on the behavior of people to modify, or not, sip sizes in response to temperature. In addition, although mean tem-

perature also discriminated well, the thermal exposure groups in this example were extremes apart, thus the improved sip temperature + size protocol may offer advantages when drinking temperature ranges are similar but sipping or gulping habits

Downloaded from <http://aacrjournals.org/cebip/article-pdf/28/12/2005/2285822/2005.pdf> by guest on 07 October 2024

Table 3. Pearson correlation coefficients between (A) different thermal exposure metrics and (B) repeat sittings for IARC participants

(A) Intermetric Pearson correlations coefficient (95% CI)		
	Group A hot beverage, IARC	Group B hot porridge, China
Mean temp vs. first sip temp	0.76 (0.59–0.87)	0.98 (0.97–0.99)
First sip IELT vs. first sip temp	0.04 (-0.28, 0.35)	0.45 (0.20–0.64)
Mean IELT vs. first sip temp	0.39 (0.09–0.63)	0.73 (0.56–0.83)
Sip-weighted temp vs. first sip temp	0.57 (0.32–0.75)	0.97 (0.95–0.98)
First sip IELT vs. mean temp	0.08 (-0.23, 0.38)	0.50 (0.27–0.68)
Mean IELT vs. mean temp	0.74 (0.55–0.85)	0.78 (0.64–0.87)
Sip-weighted temp vs. mean temp	0.94 (0.88–0.97)	0.995 (0.99–0.997)
Mean IELT vs. first sip IELT	0.35 (0.04–0.60)	0.71 (0.55–0.82)
Sip-weighted temp vs. first sip IELT	0.16 (-0.15, 0.45)	0.56 (0.34–0.72)
Sip-weighted temp vs. mean IELT	0.79 (0.64–0.89)	0.79 (0.65–0.87)
(B) Intraperson (between repeat sitting) Pearson correlation coefficient (95% CI) in descending order of strength		
Mean IELT	0.76 (0.48–0.90)	—
Sip-weighted temp	0.76 (0.48–0.90)	—
Mean temp	0.71 (0.39–0.88)	—
First sip temp	0.69 (0.36–0.87)	—
First sip IELT	0.37 (-0.08, 0.70)	—

Abbreviation: temp, temperature.

differ. Furthermore, a measurement at the beginning, middle, and just prior to the end of the drink is logistically feasible. Therefore, with no previous studies incorporating sip size into exposure assessment protocols, we have presented both a practical methodology for doing so and proposed exposure metrics to be derived from the data generated.

Our study had some limitations. Although we benefited from previous data (9) on the relationship between liquid temperature, volume, and IELT, thus enabling temperatures relevant to the target organ to be predicted, we were limited by the narrow range of sip temperatures (55–65°C) and volumes (5–20 mL) assessed by de Jong and colleagues. Therefore, we could only examine the fit of our model in this range, which only 49% and 24% of IARC and Chinese participants were within, respectively (Supplementary Table S3). Predicted IELTs for sips beyond this range of are subject to assumptions and should be interpreted cautiously. Further, we assumed that this model—based on liquid (coffee) drinking—equally applied to the liquid and porridge consumption in our study. The higher viscosity of porridge may however lead to different transfer and cooling rates through the esophagus, but with no porridge-specific IELT experiment, we assumed the same IELT prediction for both substances. For these reasons, researchers might prefer to use the sip-weighted mean temperature which simply reflects the temperature at which most volume is consumed; and our protocol affords this option. Further uncertainty of drinking temperature accuracy comes from the need to measure temperature in a proxy drinking vessel (Fig. 1B)—a necessary trade-off to allow inconspicuous measurements while not interfering with normal drinking behavior or contaminating the beverage. Temperature discrepancies between proxy and drinking vessels likely depend on several factors, including drinking duration, vessel material, ambient temperature, and volume change. It is advised to fill two vessels of the same material, at the same time with the same volume of liquid in the absence of a more direct approach. We acknowledge that the sample of individuals at IARC was not representative of a specific population; however, we emphasize that the two groups were primarily chosen to capture two samples with contrasting esophageal thermal exposures, for which both were adequate. Our study does not overcome the challenge of temperature measurement in case-

control designs as, due to their dysphagia, ESCC cases cannot be measured with our protocol or any of the others.

Our approach has the potential to strengthen future descriptive cross-sectional studies and those seeking to identify biomarkers of exposure and effect, as well as prospective studies. We recently highlighted deficiencies of the self-reported/perceived approach to drinking temperatures (e.g., "very hot"/"hot"/"warm"), which is vulnerable to interviewer and recall bias (15). In the present study, self-reported temperatures were higher for IARC beverage drinkers, who had much lower objectively determined thermal exposures. However, prospective studies conducted in Iran (6) and China (5) have found significantly elevated ESCC risks for self-reported drinking temperatures. Indeed, in the present study, we report within-group agreement between perceived and measured exposures (Supplementary Table S2), particularly for mean IELT. This suggests that responses (i) are relative to the norm within a given population and (ii) reflect heat as opposed to temperature alone, and may be of value to within-population but not between-population comparisons.

Given the demonstrated findings of the importance of both sip size and sip temperature, it is unfortunate that the only large cohort to have measured drinking temperatures, in Golestan (6), did not also include sip sizes, ideally multiple sip sizes, but at least the size of the first sip. If an inverse sip size-temperature relationship was present in Golestan, misclassification of thermal exposures may have been large based on first sip temperature alone, perhaps contributing to the relatively low reported HR (~1.4). Interestingly, the HR for self-reported "very hot" drinkers reported prospectively, which is not subject to recall bias, was 2.4, perhaps capturing a subgroup with truly high thermal exposure that was not captured in the measured first-sip temperature. The most profound implications of our study are how the findings of previous cross-sectional studies reporting first sip temperatures are interpreted; and how subsequent study designs are improved to yield meaningful descriptive data and interpopulation comparisons. Such information is essential to elucidating the carcinogenic mechanism of thermal injury and its lower threshold (i.e., maximum safe temperature at which any size sip can be taken), which is needed to design effective communication for cancer prevention strategies.

In conclusion, our study shows that temperature at first sip is suboptimal for assessing human exposure to hot beverages/semisolids, and future studies should include sip size measurements in exposure assessment protocols. In particular, comparisons across populations are questionable as drinking habits following first sip can be very distinct and result in different thermal exposure to the esophagus not captured in the information on the temperature at first sip alone. This finding endorses in general the responsibility of epidemiologists to question, verify, and continuously improve their applied exposure assessment methods, as the detection of especially moderate associations requires as little exposure misclassification as possible.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

Where authors are identified as personnel of the IARC/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy, or views of the IARC/World Health Organization.

Authors' Contributions

Conception and design: D.R.S. Middleton, W.-Q. Wei, V.A. McCormack
Development of methodology: D.R.S. Middleton, G. Byrnes, V.A. McCormack
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.R.S. Middleton, S.-H. Xie
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D.R.S. Middleton, S.-H. Xie, L. Bouaoun, G. Byrnes, V.A. McCormack

Writing, review, and/or revision of the manuscript: D.R.S. Middleton, S.-H. Xie, L. Bouaoun, G. Byrnes, J. Schütz, W.-Q. Wei, V.A. McCormack
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D.R.S. Middleton, G.-H. Song, W.-Q. Wei
Study supervision: G.-H. Song, W.-Q. Wei, V.A. McCormack

Acknowledgments

The authors thank all study participants, and the staff and fieldworkers at the Cancer Institute/Hospital of Ci County for their valuable contributions. Illustrations used in Fig. 2 were produced by Morena Sarzo, a Visual Designer from the IARC Communications Group. The work reported was undertaken during the tenure of a Postdoctoral Fellowship awarded to D.R.S. Middleton from the IARC, partially supported by the European Commission FP7 Marie Curie Actions—People—Co-funding of regional, national and international programmes (COFUND) and the World Cancer Research Fund International (grant no. 2018/1795). Work conducted in China was funded by a grants awarded to W.-Q. Wei by the National Key Research and Development Program (Precision Medicine Research; grant no. 2016YFC0901404) supported by the Ministry of Science and Technology of the People's Republic of China; the National Natural Science Fund (grant no. 81573224) supported by the National Natural Science Foundation of China; and the Innovation Fund for Medical Sciences of Chinese Academy of Medical Sciences (grant no. 2016-I2M-3-001).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 19, 2019; revised August 28, 2019; accepted September 17, 2019; published first September 26, 2019.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Watson W. Cancer of the esophagus: some etiological considerations. *Am J Roentgenol* 1939;14:420–4.
- Loomis D, Guyton KZ, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, et al. Carcinogenicity of drinking coffee, mate, and very hot beverages. *Lancet Oncol* 2016;17:877–8.
- IARC Monographs Priorities Group. Advisory Group recommendations on priorities for the IARC Monographs. *Lancet Oncol* 2019;20:763–4.
- Yu C, Tang H, Guo Y, Bian Z, Yang L, Chen Y, et al. Effect of hot tea consumption and its interactions with alcohol and tobacco use on the risk for esophageal cancer. *Ann Intern Med* 2018;168:489–97.
- Islami F, Poustchi H, Pourshams A, Khoshnia M, Gharavi A, Kamangar F, et al. A prospective study of tea drinking temperature and risk of esophageal squamous cell carcinoma. *Int J Cancer* 2020;146:18–25.
- Munishi MO, Hanisch R, Mapunda O, Ndyetabura T, Ndaro A, Schütz J, et al. Africa's esophageal cancer corridor: do hot beverages contribute? *Cancer Causes Control* 2015;26:1477–86.
- Islami F, Pourshams A, Nasrollahzadeh D, Kamangar F, Fahimi S, Shakeri R, et al. Tea drinking habits and esophageal cancer in a high risk area in northern Iran: population based case-control study. *BMJ* 2009;338:b929.
- De Jong UW, Day NE, Mounier-Kuhn PL, Haguenaer JP. The relationship between the ingestion of hot coffee and intraoesophageal temperature. *Gut* 1972;13:24–30.
- Edwards FC, Edwards JH. Tea-drinking and gastritis. *Lancet* 1956;271(SEP15):543–5.
- Lee HS, O'Mahony M. At what temperatures do consumers like to drink coffee?: mixing methods. *J Food Sci* 2002;67:2774–7.
- Victoria CG, Muñoz N, Horta BL, Ramos EO. Patterns of mate drinking in a Brazilian city. *Cancer Res* 1990;50:7112–5.
- Mwachiro MM, Parker RK, Pritchett NR, Lando JO, Ranketi S, Murphy G, et al. Investigating tea temperature and content as risk factors for esophageal cancer in an endemic region of Western Kenya: validation of a questionnaire and analysis of polycyclic aromatic hydrocarbon content. *Cancer Epidemiol* 2019;60:60–6.
- Chen R, Ma S, Guan C, Song G, Ma Q, Xie S, et al. The National Cohort of Esophageal Cancer-Propective Cohort Study of Esophageal Cancer and Precancerous Lesions based on High-Risk Population in China (NCEC-HRP): study protocol. *BMJ Open* 2019;9:e027360.
- Middleton DR, Menya D, Kigen N, Oduor M, Maina SK, Some F, et al. Hot beverages and esophageal cancer risk in western Kenya: findings from the ESCCAPE case-control study. *Int J Cancer* 2018;144:2669–76.