# Heat and the Body: Acute Cardiometabolic Responses and the Measurement of Chronic Disease in Bangladesh

Mohammad Islamul Haque[a,*]

[a] *Yale Research Initiative on Innovation and Scale (Y-RISE), Yale University, 34 Hillhouse Avenue, Luce Hall, Suite 305, New Haven, CT 06511, USA*

**Abstract**

Household surveys diagnose chronic disease from a single measurement of each respondent, treating each reading as a measure of a chronic state. We test whether they also measure the weather using 34,879 adults with a blood-pressure reading and 33,179 with a fasting glucose result from three rounds of the Bangladesh Demographic and Health Survey, merged with ERA5 temperature at the date and hour of measurement. Identification comes from day-to-day weather within a district and survey round, net of the district's seasonal pattern. A one-degree-Celsius rise in the daily maximum lowers systolic blood pressure on the day it is read by 0.562 mmHg, and a rise the day before an early-morning fasting test raises glucose by 1.028 mg/dL; haemoglobin, measured in one round, falls 0.018 g/dL per degree, and 0.037 against the wet bulb. The response is to humid heat: with both maxima entered together, the wet bulb keeps 1.208 mg/dL on glucose and -0.778 mmHg on pressure, while the dry-bulb term falls by two-thirds to three-quarters.

*Corresponding author.
*Email address:* islamul.haque@yale.edu (Mohammad Islamul Haque)

The imprecise fall in haemoglobin points to dilution, arguing against dehydration as the source of the rise in glucose. The response moves the diagnosis: measured diabetes rises 0.626 and measured hypertension falls 0.839 percentage points per degree, while reported past diagnoses do not respond. Warmer 2022 fieldwork accounts for about a third of the 7.80-point rise in measured diabetes among adults aged 35 and over since 2017–18, a thermal component of the same order as the changes surveys are used to detect.

*Keywords:* humid heat, wet-bulb temperature, biomarkers, blood pressure, fasting glucose, chronic disease measurement

*JEL:* Q54, I15, O12, C81

---

## 1. Introduction

Household surveys are the main source of what is known about the noncommunicable disease burden of low- and middle-income countries, and they diagnose that burden from a single measurement of each respondent: one set of blood-pressure readings at a household visit and one fasting glucose test, drawn the next morning. The shares at or above 140/90 mmHg and 126 mg/dL become national prevalence rates, inputs to global burden estimates (NCD Risk Factor Collaboration (NCD-RisC), 2016, 2017, 2021, 2024), and the outcomes of a large literature on the care cascade in these countries (Geldsetzer et al., 2019; Manne-Goehler et al., 2019). Each reading is treated as a measure of a chronic state. It is also a measure of a body on a particular day, and the body responds to heat within hours. This paper asks how

strongly the biomarkers used to diagnose hypertension and diabetes respond to the weather on the day they are measured, and how much of the change in the recorded burden between survey rounds that response accounts for—which matters for how chronic disease is counted in hot, poor countries.

Bangladesh offers empirical leverage on the question. Three rounds of the Bangladesh Demographic and Health Survey (BDHS) measured adults' blood pressure at the household visit and their fasting plasma glucose at a separate early-morning visit, and one round measured haemoglobin in women aged 15–49 and children aged 6–59 months; every measurement carries its date and clock hour. Fieldwork for each round ran for five to six months and moved across the country on a cluster itinerary fixed before enumeration began, so the same district in the same calendar month supplies measurement days that differ by several degrees. Measured hypertension in these data runs between 16.6 and 23.6% across rounds and measured diabetes between 8.6 and 14.5%, so a response of a percentage point per degree is large relative to the quantity the surveys exist to estimate.

We estimate the acute response on 34,879 adults with a valid blood-pressure reading and 33,179 with a valid fasting glucose result from the 2011, 2017–18 and 2022 rounds, merged with inverse-distance-weighted daily-maximum temperature from the ERA5 reanalysis of the European Centre for Medium-Range Weather Forecasts. Blood pressure is matched to the day of the household interview at which it was read. The glucose test is drawn early in

the morning (median hour 7), before that day's afternoon peak, so glucose is matched to the previous day's maximum, and the heat that arrives after the test serves as a falsification test. The baseline specification includes district × round and district × calendar-month fixed effects, hour-of-test fixed effects for glucose, rainfall on the exposure day and the day before, and controls for age, age squared, sex, urban residence, education and wealth quintile. Identification comes from day-to-day deviations in temperature within a district's fieldwork in a round, net of that district's usual temperature in the calendar month—heat that is unusual for that district at that time of year—which we assume are orthogonal to respondents' potential biomarker outcomes given the fixed effects and controls. Standard errors are clustered at the district level (64 districts).

*The acute physiological response.* A one-degree rise in the daily maximum lowers systolic blood pressure on the day it is read by 0.562 mmHg (standard error, SE, 0.077) and diastolic pressure by 0.330 mmHg, and a one-degree rise the day before the test raises fasting plasma glucose by 1.028 mg/dL (SE 0.146), all significant at the 1% level. In 2011, the one round that measured haemoglobin, it falls by 0.018 g/dL per degree of dry-bulb maximum (SE 0.018, $p = 0.310$) and by 0.037 per degree of wet-bulb maximum (SE 0.020, $p = 0.069$). The glucose response accumulates over several days: entering the exposure day and the three preceding days together gives a cumulative 1.254 mg/dL per degree. Blood pressure loads mainly on the same day (-0.468

of a four-day cumulative -0.632), although it also tracks the temperature of the surrounding fortnight (Section 6.5). The fall in pressure is 2.4 times larger among adults aged 50 and over (interaction -0.537 mmHg, SE 0.085, on a base of -0.392 for younger adults), the group with the most cardiovascular risk, while the glucose response varies less with age (interaction 0.271, SE 0.141, $p = 0.060$).

*Dilution rather than dehydration.* The three biomarkers move as a single physiological response would move them: plasma volume expands and peripheral vessels dilate, which lowers both blood pressure and the concentration of haemoglobin, while thermal stress raises counter-regulatory hormones and, with them, circulating glucose—during exertion in the heat through greater hepatic glucose output (Hargreaves et al., 1996), although the evidence at rest is thinner (Section 2.2). The haemoglobin result bears on the obvious mechanical explanation for the rise in glucose, which is dehydration: if heat concentrated the blood, every analyte would rise together. Haemoglobin falls in every cut, though imprecisely, so the evidence argues against dehydration. It does not sign the remaining bias in the glucose estimate. Dilution would lower the glucose concentration of plasma, while point-of-care analysers tend to read glucose high when haematocrit is low (Tang et al., 2000); both effects are small beside the estimated response, since haemoglobin falls by a fraction of a per cent per degree.

*Humid heat rather than heat as such.* Evaporation, the body's main defence against heat, works against the humidity of the air rather than its temperature, and the wet-bulb temperature measures that margin. Entered alone, the wet-bulb maximum gives a larger coefficient than the dry-bulb maximum on every outcome. Entered together, which identifies the wet-bulb coefficient from days equally hot and differently humid within the same district and round, the wet bulb keeps 1.208 mg/dL on glucose (SE 0.188) and -0.778 mmHg on systolic pressure (SE 0.131), while the dry-bulb term falls to 0.364 on glucose, still significant at the 5% level, and -0.146 on systolic pressure; for the two diagnostic classifications the dry-bulb term (0.275 and -0.226) is not significant at the 5% level ($p = 0.065$ and $p = 0.267$). For haemoglobin the dry-bulb term disappears: with both entered, the wet bulb carries -0.039 g/dL per degree (SE 0.020, $p = 0.060$) and the dry bulb 0.003.

*The reading moves the diagnosis.* Because chronic disease is diagnosed from a single reading, the physiology moves the classification. Measured diabetes prevalence rises 0.626 percentage points per degree (SE 0.143) and measured hypertension falls 0.839 (SE 0.145). Across the glucose distribution the response has the shape a uniform shift of every reading would produce: the share of readings above 100 mg/dL rises 2.252 points per degree against 2.455 implied by shifting every reading by the baseline estimate, and the share above 126 mg/dL rises 0.626 against 0.665. The classification moves because readings already near the line cross it, and with measured diabetes

at 11.2% there are many such readings. Respondents' own reports of a past diagnosis of diabetes (0.008, SE 0.104) or of high blood pressure (-0.152, SE 0.152) do not respond to the weather, so hot days do not reach a different population. Among respondents already diagnosed with diabetes the glucose response is smaller, but not significantly so (interaction -0.623, SE 0.478, $p = 0.197$).

*How much of the survey record is thermal.* Among adults aged 35 and over, the age range all three rounds cover, the glucose samples were measured at mean exposure-day maxima of 29.4, 27.4 and 31.4°C, so a comparison of two rounds embeds a temperature difference of two to four degrees. Recorded diabetes prevalence in that age range rose 7.80 percentage points between 2017–18 and 2022. The 2022 round was fielded 4.0°C warmer, and on the slope re-estimated for that age range (Table 9) the warmer fieldwork accounts for 2.69 points, about a third, and for 4.51 on the wet bulb; at a common fieldwork temperature the rise is 5.11 points (95% interval from the slope, 3.85 to 6.37). Between 2011 and 2017–18 the recorded rise was 2.15 points, but the later round was fielded 2.0°C cooler, and at a common temperature the rise is 3.52 points (2.88 to 4.16). These adjustments hold the slope fixed and apply a day-to-day response to seasonal differences, so we read them as orders of magnitude rather than corrected trends (Section 9). They are of the same order as the changes the surveys are used to detect.

*Robustness.* Sections 6 and 7 and the Online Appendix test each result against the specific concern it raises; Table 7 collects the specifications. We summarise the main checks here, stating in each case the concern, the test and the outcome.

*Who is measured.* If field teams reached different households on hot days, or measured fewer of them, the measured population would differ with temperature. We regress 12 predetermined characteristics on the exposure-day maximum in the full sample of 38,813 biomarker-eligible adults: 3 are significant at the 5% level (education, wealth quintile and a prior diagnosis of high blood pressure), all are small, and the largest is 0.0212 standard deviations per degree. Consent to the blood-pressure measurement falls by 0.325 percentage points per degree ($p = 0.002$) on a base of 91%, but for selection alone to produce the systolic estimate the respondents lost per degree would need readings about 210 mmHg away from those measured. Neither the composition of the measured population nor selection into measurement can plausibly produce the estimates.

*Fieldwork and instruments.* Heat might degrade the measurement rather than the person. Fieldworker fixed effects leave the estimates at 0.861 mg/dL and -0.604 mmHg, cuff-size fixed effects in 2022 give -0.677 mmHg, and outcomes recorded on the same visit with a tape, a measuring board and a scale do not respond to heat (Section 6). The glucose analysers are the harder case. Restricting to tests at which the temperature at the test hour lies

within each round's strictest stated operating range gives 0.978 mg/dL (SE 0.198), against 1.028 in the full sample, but the test-hour temperature, which is also the analyser's, carries much of the glucose response, and we cannot exclude temperature sensitivity within the stated range without calibration records (Section 6.4). Fieldwork quality does not drive the results; temperature sensitivity of the glucose analyser within its stated range is the one instrument explanation we cannot rule out.

*Exposure timing.* If glucose responded to heat that arrived after the test, the timing would point to an artefact rather than physiology. Entering the previous day's maximum, the temperature at the hour of the test and the test day's maximum together, the previous day keeps 0.526 mg/dL ($p = 0.001$), the test hour carries 0.836, and the test-day maximum falls to 0.106 ($p = 0.590$; Section 6.6). Heat that arrives after the test does not move the reading.

*Specification.* The estimates might reflect seasonality the fixed effects leave in place, particular controls or subgroups, or the scale on which glucose is recorded. Across the specifications that vary the fixed effects or the sample, apart from the within-cluster comparison discussed below, the glucose estimate lies between 0.826 and 1.152 mg/dL and the systolic estimate between -0.677 and -0.493 mmHg, all significant at the 1% level. Replacing the level with its deviation from the district's day-of-year normal attenuates the estimates by a tenth to a fifth without changing their sign or significance; district

× round × month fixed effects give 0.935 mg/dL and -0.577 mmHg; temperature thirty days ahead does not predict glucose (0.135, SE 0.151); and leaving the 2022 glucose readings without the plasma conversion that DHS applies gives 0.997 mg/dL. Comparisons within a single survey-cluster visit, which retain a tenth to a seventh of the temperature variation the design uses, do not reproduce the baseline (0.360, SE 0.358, for glucose), and we report them as such (Section 7). Apart from that comparison, no specification changes the conclusions.

*Inference.* District-clustered standard errors might overstate precision given 64 districts and spatially correlated weather. We therefore also use Conley spatial heteroskedasticity- and autocorrelation-consistent (HAC) standard errors at 30 and 50 km (Conley, 1999), two-way clustering by district and date (Cameron et al., 2011), and a restricted wild cluster bootstrap (Cameron et al., 2008). The alternative standard errors are at most 1.22 times the district-clustered ones, and all four headline estimates remain significant at the 1% level under each. Romano–Wolf adjusted $p$-values, which account for the number of outcomes we test, are 0.001 for the four headline estimates and diastolic pressure. The precision of the blood-pressure and glucose estimates does not depend on the clustering choice or on the number of hypotheses tested.

*Related literature and contribution.* This paper speaks to three strands of work.

*Temperature and health.* The first estimates the effect of temperature on health, largely from mortality registers in rich countries and, more recently, in poor ones (Deschênes and Greenstone, 2011; Barreca et al., 2016; Carleton et al., 2022; Heutel et al., 2021; Burgess et al., 2017; Geruso and Spears, 2018; Gasparrini et al., 2015; Banerjee and Maharaj, 2020; Haque and Chopra, 2026), and increasingly from morbidity, labour supply and productivity (White, 2017; Graff Zivin and Neidell, 2014; Somanathan et al., 2021; Ebi et al., 2021); reviews of its methods stress what weather variation does and does not identify (Dell et al., 2014; Auffhammer et al., 2013; Hsiang, 2016). The physiological literature has documented thermal vasodilation and plasma-volume expansion in the laboratory for half a century, but free-living evidence on the metabolic response is thin. We provide population-representative estimates of the acute response of three biomarkers in a low-income country, an age gradient estimated on the general population, and evidence that the response is to humid rather than dry heat. Biomarkers sit upstream of mortality and are observed for the living, which makes them informative about the channel and about who is exposed to it.

*Adaptation.* The second documents that the health cost of heat falls as populations adapt, through air conditioning, electrification and other margins (Barreca et al., 2016; Deschênes, 2014; Adhvaryu et al., 2020; Mullins and White, 2020; Davis and Gertler, 2015; Jay et al., 2021; Haque and Chopra, 2026). Across our three rounds household electricity rises from 62% to 99% and electric-fan ownership from 55% to 97%. We provide a test of whether

the acute physiological response weakened over that decade: no outcome's round-specific slopes differ significantly at the 5% level, and within rounds the response does not vary with cooling assets (Section 8). Because the round-specific slopes are imprecise, we read this as an absence of evidence of adaptation rather than evidence of its absence.

*Measurement.* The third concerns how the conditions of measurement move a measured quantity. Heat on the day of a test lowers exam scores in New York, China and India (Park, 2022; Graff Zivin et al., 2020; Garg et al., 2020; Graff Zivin et al., 2018) and learning over the school year (Park et al., 2020), and DHS interviewers themselves work differently on hot, humid days (LoPalo, 2023). In health surveys, the month of measurement moves child anthropometry enough to alter trends between DHS rounds (Madan et al., 2020), and a single reading misclassifies many respondents relative to repeat measurement (Selvin et al., 2007; Plumettaz et al., 2020; Modesti et al., 2012). Closest to our question, epidemiological surveys in Ghana and rural China relate the temperature at or on the day of measurement to measured blood pressure and to the rate at which hypertension is detected, and recommend that surveys record temperature and standardise to it (Kunutsor and Powles, 2010; Su et al., 2014); and four provincial surveys in Guangdong have been used to estimate a temperature-adjusted diabetes prevalence (Luo et al., 2021). These studies compare respondents measured in different seasons or at different times of day, so their estimates combine the response to temperature with whatever else varies with the season or the hour, and none

sets the adjustment against a published national trend. We provide an estimate identified from day-to-day weather within a district and survey round, net of the district's own seasonal pattern; evidence from a tropical setting where the glucose response to heat is positive; and an accounting of how much of the change between rounds of a national DHS series, on which published trends rest (Sathi et al., 2022; Rana et al., 2024), reflects the weather at measurement.

To a literature built largely on mortality and on the level of the heat response, we therefore add three things: estimates of the acute response of the biomarkers that surveys use to diagnose chronic disease, evidence that the response is to humid heat, with suggestive evidence that it works through dilution rather than dehydration, and an accounting of the thermal component of the recorded change in measured diabetes and hypertension between survey rounds.

Two limitations bound these claims. The estimates are of an acute response; we observe no mortality, hospitalisation or clinical diagnosis, and cannot show that a hot day leaves durable harm. And BDHS fieldwork does not run through Bangladesh's April–May pre-monsoon peak, so exposure-day maxima span 19.1°C to 36.5°C and no day in the sample reaches 37°C. The estimates are therefore a moderate-range dose-response and say nothing about the extreme tail that dominates the heat-mortality literature—including in Bangladesh, where Haque and Chopra (2026) find that for recent birth co-

horts the neonatal mortality response to heat is concentrated above 40°C.

## 2. Background

Three features of the Bangladeshi setting motivate the analysis: a non-communicable disease burden that is large and measured almost entirely by survey, from a single reading per respondent; a human response to heat that is fast and well characterised for blood pressure and haemoglobin and less so for glucose; and a survey calendar that places measurement days across a wide moderate range of temperature but never at its extreme.

### *2.1. Chronic Disease Is Measured by Survey, from a Single Reading*

Bangladesh has no population disease register. What is known about the prevalence of hypertension, diabetes and anaemia comes from a handful of nationally representative surveys, of which the BDHS is the largest and the most frequently used (Akter et al., 2014; Rahman et al., 2015; Sathi et al., 2022; Hossain et al., 2022; Haider et al., 2023); the WHO STEPS survey of 2018 is the other main source (Riaz et al., 2020). In these surveys a respondent is classified as hypertensive if the average of the second and third blood-pressure readings taken at a single household visit is at least 140/90 mmHg, as diabetic if a single fasting plasma glucose result is at least 126 mg/dL, and as anaemic if one haemoglobin reading falls below the WHO threshold for the respondent's group: 12 g/dL for a non-pregnant woman aged 15–49

and 11 g/dL for a child aged 6–59 months (World Health Organization and International Diabetes Federation, 2006; World Health Organization, 2021, 2011).

Each classification is therefore a chronic construct estimated from an acute measurement. This is what a household survey can do, and clinical guidelines themselves define these conditions by thresholds that a single reading can cross, although they ask for the reading to be confirmed on another occasion (World Health Organization and International Diabetes Federation, 2006; World Health Organization, 2021; Muntner et al., 2019). Without that confirmation the classification is unstable even before the weather enters: of adults with fasting glucose at or above 126 mg/dL on one test, only about 70% were above it again two weeks later (Selvin et al., 2007), and hypertension prevalence from a single visit exceeds that from two visits by between a seventh and a half (Plumettaz et al., 2020; Modesti et al., 2012). The recorded prevalence inherits whatever moves the reading. In our estimation samples measured hypertension runs between 16.6 and 23.6% across rounds and measured diabetes between 8.6 and 14.5%, so a shift of one or two percentage points is a large share of the level being estimated and a larger share of any change between rounds.

### *2.2. Heat Moves Blood Pressure, Haemoglobin and Glucose within Hours*

The human response to ambient heat is fast, involuntary and well documented (Cramer et al., 2022). Within minutes of a rise in core or skin

temperature, cutaneous vessels dilate to move blood towards the surface for convective and evaporative heat loss; skin blood flow can reach several litres a minute, and systemic vascular resistance falls sharply (Charkoudian, 2003). Arterial pressure is then defended by a rise in cardiac output, successfully in a healthy adult at rest under mild heat and less so as the thermal load, or an added circulatory demand, grows (Crandall and Wilson, 2015). A fall in measured pressure is therefore what this physiology predicts under sufficient load, not a mechanical certainty at every temperature.

*Blood pressure.* Population studies find that sign. Measured pressure is lower in summer than in winter, and on warmer days than on cooler ones, in Britain, China, France and the United States (Brennan et al., 1982; Lewington et al., 2012; Alpérovitch et al., 2009; Halonen et al., 2011), in a meta-analysis (Wang et al., 2017) and in the European Society of Hypertension's consensus statement (Stergiou et al., 2020). The fall is larger in older people (Alpérovitch et al., 2009; Kenney et al., 2014), and in coastal Bangladesh same-day heat lowers systolic pressure (Mukhopadhyay et al., 2025). In population surveys the effect is of the order of 5–7 mmHg of systolic pressure per 10°C: 5 mmHg in rural Ghana (Kunutsor and Powles, 2010) and 6.9 mmHg in rural China, where each 10°C lower outdoor temperature was also associated with 14% more newly detected hypertension (Su et al., 2014). Ambulatory monitoring adds one qualification: night-time pressure can rise with heat even as daytime and office readings fall (Modesti et al., 2006), so a daytime reading

captures one part of a daily cycle.

*Haemoglobin.* Body fluids shift quickly. Early heat exposure dilutes the blood within the first half hour, before sweating is appreciable, and the shift reverses to concentration only once sweating dehydrates (Bass and Henschel, 1956). Over hours to days, plasma volume expands, the standard early marker of heat acclimation (Senay et al., 1976; Périard et al., 2015), and dilutes whatever is dissolved in the blood: repeated heat exposure raises plasma volume without changing red-cell mass, so measured haemoglobin *concentration* falls although circulating haemoglobin does not (Sawka et al., 2000). Population data show the same pattern, with haematocrit about 3% lower in summer than in winter (Thirup, 2003) and haemoglobin lower in the hottest month (Kristal-Boneh et al., 1993).

*Glucose.* Thermal stress raises counter-regulatory hormones (catecholamines, glucagon and cortisol), and circulating glucose rises with them, principally because hepatic glucose output increases rather than because peripheral uptake fails (Hargreaves et al., 1996). For our purposes this is the weakest of the three mechanisms. The acute effect is well established during exertion in the heat and after a glucose load under passive heating, where a few hours at 31°C raise post-load glucose in healthy young men (Faure et al., 2016), but passive heating of a fasting non-diabetic adult has no consistent effect on its own (Maley et al., 2019). Heat raises hospital admissions for diabetes (Xu et al., 2019), and diabetes impairs heat

dissipation (Kenny et al., 2016). Over longer horizons the sign is unsettled: glycated haemoglobin is higher in winter (Tseng et al., 2005), while warmer places have more diabetes and glucose intolerance (Blauw et al., 2017; Valdés et al., 2019). Over days, the relation may not be monotone. In a Chinese cohort, fasting glucose follows a U-shape in the three-day mean temperature, lowest near 18°C and higher at both cold and hot temperatures (Li et al., 2016); across four surveys in Guangdong, fasting glucose falls as the survey-day temperature rises from 5 to 30°C, although among newly detected diabetics it is again higher at 30°C than at 19°C (Luo et al., 2021). Our exposure-day maxima lie above 19°C, on the side of that U where heat raises glucose. We therefore treat glucose as the margin where the physiological prior is weakest, and rely on the internal evidence rather than on the prior.

*Humidity.* Which route of heat loss can do the work depends on the air. Convection sheds heat against the dry-bulb temperature; evaporation sheds it against the humidity, and its capacity is summarised by the wet-bulb temperature, the lowest temperature a wet surface can reach by evaporating into the surrounding air. Two days at the same dry-bulb maximum therefore impose different thermal loads if their humidity differs, and the responses above should track the wet bulb more closely than the dry bulb (Section 5.3 tests this). We use the wet bulb as a tractable summary of the joint temperature–humidity constraint on heat balance rather than as an exact physiological

index. The 35°C wet-bulb survivability limit of Sherwood and Huber (2010) is an upper bound: laboratory work puts the critical value for healthy adults at rest or light activity well below it (Vecellio et al., 2022; Wolf et al., 2022), and the heat-index model implies critical wet bulbs of 20–32°C depending on humidity (Lu and Romps, 2023). The epidemiological evidence on humidity is less settled than the physiology: humidity matters for mortality in the United States (Barreca, 2012), adds little to temperature in a multi-country study of summer mortality (Armstrong et al., 2019), and the gap between the two literatures is itself debated (Baldwin et al., 2023).

Three implications of this physiology organise the analysis. First, the responses have opposite signs: pressure and haemoglobin concentration should fall while glucose rises, so a common artefact, such as sloppier fieldwork on hot days, would not generate the pattern. Second, dilution and concentration are distinguishable. If heat dehydrated respondents, every blood analyte would rise together; if plasma volume expands, concentrations fall. Haemoglobin therefore indicates which of the two is operating, and so whether dehydration could account for a rise in measured glucose. Third, the response should be larger where vasomotor regulation is impaired (Cramer et al., 2022), which points to older adults and to those already hypertensive.

### *2.3. Fieldwork Covers a Wide Moderate Range, but Not the Extreme*

BDHS fieldwork runs for five to six months and does not follow the calendar year. The three biomarker rounds were enumerated between 10 July 2011 and

10 December 2022, and none covers Bangladesh's April–May pre-monsoon peak, when heat waves—days on which both the daily minimum and the daily maximum are extreme—occur and raise mortality (Nissan et al., 2017). Outside that peak, mortality in Bangladesh is highest in the cold season, with heat effects concentrated at the top of the temperature distribution (Burkart et al., 2011, 2014). Exposure-day maxima in our samples span 19.1°C to 36.5°C, and no day in the three rounds reaches 37°C (Figure 1).

This has two consequences that we carry throughout. First, the design has a wide moderate range to work with, roughly seventeen degrees, with substantial variation across days within the same district and month, but it cannot speak to the extreme tail: we can say nothing about whether these biomarkers respond differently at 40°C and above. Second, the three rounds differ systematically in the temperature at which the population was measured. For the design this is a nuisance, absorbed by district × round fixed effects; for the survey record it is the object of interest in Section 9.

## 3. Data

### *3.1. Bangladesh Demographic and Health Survey*

The BDHS is a nationally representative household survey implemented by the National Institute of Population Research and Training with technical support from The DHS Program. Three rounds measured biomarkers in a subsample of households: 2011 (National Institute of Population Research

and Training (NIPORT), Mitra and Associates, and ICF International, 2013), 2017–18 (National Institute of Population Research and Training (NIPORT) and ICF, 2020) and 2022 (National Institute of Population Research and Training (NIPORT) and ICF, 2024). Each round is an independent, nationally representative probability sample, drawn anew; the rounds are not a panel, and no respondent is measured twice.

*Who is measured.* Eligibility differs across rounds in a way that matters for comparisons between them. BDHS 2011 collected biomarkers only from adults aged 35 and over, while 2017–18 and 2022 cover adults aged 18 and over, so mean age in the blood-pressure sample is 51.4, 39.8 and 40.8 years across the three rounds. Wherever we compare rounds we therefore restrict to adults aged 35 and over, the range all three cover; estimates within rounds use the full eligible sample. Sex is the respondent's sex as reported in the household roster, recorded as male or female; the survey does not record gender identity. We use it as a control and also estimate the response separately for women and men, for whom it is similar (Table 7). Haemoglobin was measured only in 2011, and only in women aged 15–49 and children aged 6–59 months, so Section 5.2 runs on a different population from the rest of the paper.

*Estimation samples.* The estimation samples contain every adult with a valid measurement: 34,879 adults for blood pressure and 33,179 for fasting glucose, in 64 districts. The glucose descriptives in Table 1 cover 33,421 adults; 242 of

them have no valid test hour and drop out once hour-of-test fixed effects are included, which is why every glucose regression reports 33,179. Of the 34,879 adults with a valid blood-pressure reading, the 34,861 with complete controls enter the blood-pressure regressions; the others lack education. Women who were pregnant at the time of the survey (664 in the blood-pressure sample and 628 in the glucose sample of Table 1) are kept in both samples, and Section 7 shows that excluding them changes little. Tests taken during Ramadan are excluded from the glucose sample throughout, since fasting behaviour differs; including them changes little (Table 7). For the selection tests of Section 6 we also retain all 38,813 biomarker-eligible adults, whether or not a measurement was obtained.

### *3.2. Biomarkers, Instruments and Reported Diagnoses*

*Biomarkers.* Blood pressure is taken with an automated monitor in three readings at ten-minute intervals during the household visit; following the surveys' convention we use the average of the second and third. Fasting plasma glucose is measured at a separate early-morning visit after an overnight fast, and we date that observation by the test visit rather than by the household interview. The median test follows the interview by 1 day, and 97% follow it within two days. We treat test dates more than a week from the interview as recording errors, with one exception: in 2011, households interviewed during Ramadan, who could not fast for the test, were retested when fieldwork resumed in January 2012, and we keep those deferred tests at their

recorded dates. From each reading we construct the diagnostic classification it generates: measured hypertension at 140/90 mmHg, measured diabetes at 126 mg/dL, and measured anaemia at the WHO threshold for the respondent's group (12 g/dL for a non-pregnant woman aged 15–49, 11 g/dL for a child aged 6–59 months). We call these *measured* throughout, to keep them distinct from the chronic conditions.

*Instruments.* Blood pressure was taken with A&D automatic monitors in every round: the LifeSource UA-767 Plus in 2011 and 2017–18, and in 2022 three monitors of the same family, one per cuff size. Glucose was measured with HemoCue point-of-care analysers: the Glucose 201+ in 2011, and in 2017–18 and 2022 an analyser that the reports name as the 201 RT in their methods chapters and as the plasma-converting Glucose 201 DM in their results tables (National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013; National Institute of Population Research and Training (NIPORT) and ICF, 2020, 2024). The manufacturers' operating ranges differ by model (A&D Medical (LifeSource), 2009; HemoCue AB, 2024c,b,a), and Section 6.4 uses them to test whether instrument drift can account for the results.

*The glucose scale.* The diagnostic threshold is defined on plasma glucose, which is about 11% higher than the whole-blood glucose of a capillary sample, and the three rounds reach the plasma scale in different ways. In 2011 the survey converted the analyser's whole-blood readings to plasma equiva-

lents by multiplying by 1.11 (National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013). In 2017–18 the analyser converted each reading automatically, and the recode reports those values (National Institute of Population Research and Training (NIPORT) and ICF, 2020). For 2022 the documentation conflicts: the results tables describe a plasma-converting analyser, but the recode's own plasma-glucose variable is the recorded reading multiplied by 1.11, and the report's prevalence estimates are computed on that scale. We follow DHS and apply the same factor, which puts the 2022 readings on the scale of the official statistics. Left as recorded, the 2022 readings give a glucose slope of 0.997 mg/dL rather than 1.028 and a measured-diabetes slope of 0.481 (SE 0.118) rather than 0.626 (the first is in Table 7); the second is smaller because the 126 mg/dL threshold then sits further into the upper tail of the 2022 distribution, where fewer readings lie.

*Reported diagnoses.* We also use the respondent's own report of a past diagnosis, which records the chronic condition rather than the day's physiology and so serves as a placebo. All three rounds ask whether a doctor or health worker has ever diagnosed high blood pressure or diabetes, and whether the respondent takes medication for high blood pressure. Where a filter question skipped the item, because the respondent's pressure or blood sugar had never been checked, we code the respondent as never diagnosed; answers of "don't know" and refusals are left missing. The reports are available for 34,872

respondents in the blood-pressure sample and 33,369 in the glucose sample. Table 1 describes the samples by round.

### *3.3. Heat Exposure: Inverse-Distance Weighted ERA5*

Daily-maximum temperature comes from the ERA5 reanalysis of the European Centre for Medium-Range Weather Forecasts (Hersbach et al., 2020, 2023; Soci et al., 2024). We take the hourly 2-metre temperature at each grid cell, compute the daily maximum, and aggregate it to each district centroid by inverse-distance weighting over the four nearest cells, with weights proportional to the inverse squared distance; the construction is that of Haque and Chopra (2026), and the series runs from 1976 to 2024 (the wet bulb, which needs the dewpoint, from 1980 to 2022). Every survey cluster is assigned to its district, and each respondent is matched to that district's weather on their exposure day (Section 4.2 and equation (1)): the day of the household interview for blood pressure, and the day before the early-morning test for fasting glucose.

*Local time.* ERA5 is stamped in UTC, and Bangladesh keeps UTC+6. We shift the hourly series to local time before taking the daily maximum, so every daily value, dry and wet bulb alike, is a maximum over a Bangladesh calendar day, labelled with the local date. A maximum over the UTC day would instead span 06:00 local time to 06:00 the next morning. That window still contains the afternoon peak, so the two conventions give the same

**Table 1: Summary statistics, by BDHS round**

| | 2011 | 2017–18 | 2022 |
|---|---|---|---|
| **Biomarkers** | | | |
| Systolic blood pressure (mmHg) | 118.7 | 122.3 | 118.0 |
| Diastolic blood pressure (mmHg) | 77.8 | 80.1 | 77.2 |
| Measured hypertension, ≥140/90 (%) | 20.6 | 23.6 | 16.6 |
| Fasting plasma glucose (mg/dL) | 105.9 | 102.2 | 113.9 |
| Measured diabetes, FPG ≥126 (%) | 9.3 | 8.6 | 14.5 |
| **Exposure on the measurement day** | | | |
| Daily maximum temperature (°C) | 30.0 | 27.4 | 31.4 |
| standard deviation | 3.0 | 3.0 | 2.1 |
| 10th percentile | 27.0 | 23.4 | 28.6 |
| 90th percentile | 33.3 | 31.5 | 34.0 |
| Daily maximum wet-bulb (°C) | 24.9 | 20.4 | 25.7 |
| standard deviation | 3.7 | 2.7 | 2.9 |
| 10th percentile | 20.1 | 16.6 | 21.1 |
| 90th percentile | 28.3 | 23.9 | 28.6 |
| Rainfall (mm/day) | 7.2 | 0.3 | 6.4 |
| share of days with any rain (%) | 31.6 | 4.0 | 31.7 |
| **Respondent characteristics** | | | |
| Age (years) | 51.4 | 39.8 | 40.8 |
| Female (%) | 50.5 | 56.6 | 55.5 |
| Urban (%) | 32.8 | 36.3 | 34.4 |
| Wealth quintile (1–5) | 3.15 | 3.08 | 3.10 |
| **Household cooling technology** | | | |
| Electricity (%) | 62.4 | 81.6 | 98.6 |
| Electric fan (%) | 55.2 | 80.8 | 97.3 |
| Air conditioning (%) | — | 0.4 | 1.6 |
| Refrigerator (%) | 17.7 | 32.1 | 57.5 |
| Observations, blood pressure | 7,886 | 13,000 | 13,993 |
| Observations, glucose | 7,263 | 12,209 | 13,949 |
| Districts | 64 | 64 | 64 |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 and CHIRPS. Each cell is a mean over the round's estimation sample unless the row says otherwise; the samples are unweighted. The blood-pressure sample is every adult with a valid systolic and diastolic reading, each the average of the second and third of three readings at the household visit; the glucose sample is every adult with a valid fasting plasma glucose result from an early-morning test visit, excluding tests taken during Ramadan. BDHS 2011 measured biomarkers only in adults aged 35 and over; 2017–18 and 2022 cover adults aged 18 and over. Measured hypertension is a reading of at least 140/90 mmHg and measured diabetes a fasting plasma glucose of at least 126 mg/dL; the 2022 glucose readings are multiplied by 1.11 to put them on the plasma scale, as in the DHS recode. Temperature is the ERA5 daily maximum, dry bulb or wet bulb, inverse-distance-weighted to the district centroid over Bangladesh local days; rainfall is the CHIRPS daily total, aggregated the same way, and a rainy day has at least 1 mm. The exposure block is computed on the blood-pressure sample, on the interview day. The glucose test falls on the interview day for only 2% of tests, and its exposure is the day before the test, so its round means differ: 29.4, 27.4, 31.4 °C on the dry bulb and 23.9, 20.4, 25.7 on the wet bulb, against 30.0, 27.4, 31.4 and 24.9, 20.4, 25.7 in the table. Air conditioning was not recorded in 2011.

value on 99.3% of district-days for the dry bulb and 95.0% for the wet bulb, and differ by more than a degree on 0.11% and 0.42%; Online Appendix Table A.8, Panel C, rebuilds the wet bulb on UTC days, which barely moves its estimates.

*The wet bulb.* From the same reanalysis we take the hourly 2-metre temperature and 2-metre dewpoint, compute hourly relative humidity by the August–Roche–Magnus formula (Alduchov and Eskridge, 1996) and hourly wet-bulb temperature by the approximation of Stull (2011), whose mean absolute error is below 0.3°C over the range of temperature and humidity that Bangladesh occupies, and then take the daily maximum and aggregate to districts with the same weights as the dry-bulb series. Both series are non-missing for every observation in all three samples, so comparisons between them are like for like. On the days our respondents were measured, the wet-bulb maximum averages 23.5°C and reaches 30.1°C at most, with 0.2% of days at or above 30°C: the design speaks to ordinary humid heat, not to the wet-bulb extremes that bound human survivability.

*Exposure measures.* For each district-day we also construct the deviation from the district's own day-of-year normal over the full history of each series; three- and seven-day trailing means and their anomalies; lags of one, two and three days; and leads of three, seven, fourteen and thirty days for placebo tests. Figure 1 shows the distribution of interview-day maxima by round.

*Rainfall and hourly exposure.* Rainfall comes from CHIRPS (Funk et al., 2015), a daily 0.05-degree precipitation record that blends satellite estimates with station data; we aggregate it to districts with the same four-cell inverse-distance weights and enter the exposure day's total and the previous day's as controls. To place the exposure precisely relative to the measurement, we also build an hourly district series of dry- and wet-bulb temperature from the same ERA5 fields over the three fieldwork windows. Combined with the clock hour the survey records for the glucose test and for the first blood-pressure reading, it gives the maximum over the 24 hours before the measurement and the temperature at the measurement hour itself, which Section 6.6 uses.

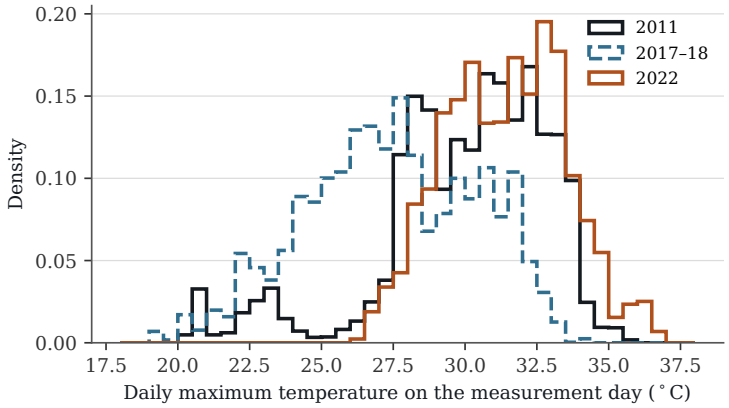



**Figure 1: Distribution of the daily maximum temperature on the interview day, by round.**

*Notes:* Histograms, normalised as densities, of the ERA5 daily-maximum temperature, inverse-distance-weighted to the district centroid over Bangladesh local days, on the date each respondent's blood pressure was measured, for the blood-pressure samples of BDHS 2011, 2017–18 and 2022.

## 4. Empirical Strategy

We estimate one linear specification for each biomarker and for the diagnostic classification it generates, and two families of variants: tighter fixed effects and alternative exposure measures, which test whether the estimate is driven by seasonality or by the choice of exposure, and the wet-bulb maximum entered alone or together with the dry-bulb maximum, which tests whether the response is to heat or to humid heat.

### *4.1. Specification*

For individual $i$ in district $d$ and BDHS round $r$, measured on calendar date $t$, we estimate

$$y_{idrt} = \alpha + \beta\, T_{d,e(i)} + \mathbf{X}_i'\boldsymbol{\gamma} + \mathbf{W}_{d,e(i)}'\boldsymbol{\delta} + \alpha_{dr} + \mu_{d,m(t)} + \theta_{h(i)} + \varepsilon_{idrt}, \quad (1)$$

where $y$ is the biomarker or the classification it generates and $T_{d,e(i)}$ is the ERA5 daily maximum at the district centroid on the exposure day $e(i)$. For blood pressure the exposure day is the day of the household interview, during which the reading was taken. For fasting glucose it is the day *before* the test: the blood is drawn at an early-morning visit (median hour 7), before that day's afternoon peak, so the last complete daily maximum the body has experienced is the previous day's. The test day's own maximum, which mostly arrives after the blood was drawn, then serves as a falsification test (Section 6.6). We refer to $T_{d,e(i)}$ as the exposure-day maximum throughout.

The vector $\mathbf{X}_i$ contains age, age squared, sex, urban residence, education and wealth quintile. $\mathbf{W}_{d,e(i)}$ contains rainfall on the exposure day and the day before, because rainfall is the obvious omitted weather variable: rainy days are cooler and more humid, and rain changes who is at home, what they do and how fieldwork runs. The district $\times$ round fixed effects $\alpha_{dr}$ (192 cells) absorb everything fixed about a district and everything specific to a survey round, including that round's instruments, protocol, age eligibility rule and sample design. The district $\times$ calendar-month fixed effects $\mu_{d,m(t)}$, estimated across the three rounds, absorb each district's own seasonal cycle in diet, infectious disease, agricultural work and food prices, so the design does not assume that the seasons move together across a country that runs from coastal mangroves to north-western drylands; for glucose they refer, like the Ramadan exclusion, to the month of the test. The two sets enter separately because each alone would leave a confounder in place: district $\times$ round fixed effects alone would leave seasonality unabsorbed, and district $\times$ calendar-month fixed effects alone would let differences between rounds in protocol and eligibility load on temperature. For glucose we add hour-of-test fixed effects $\theta_{h(i)}$, since the test runs between roughly 06:00 and 09:00. Standard errors are clustered at the district level (64 districts). We also check robustness to Conley spatial HAC standard errors (Conley, 1999), two-way clustering by district and date (Cameron et al., 2011) and the restricted wild cluster bootstrap (Cameron et al., 2008) (Section 7.2).

*Variants.* In a tighter version we replace $\alpha_{dr}$ and $\mu_{d,m(t)}$ with district × round × month fixed effects, so that identification comes only from days within the same district-month of a round. We also replace the level of $T_{d,e(i)}$ with its deviation from the district's own day-of-year normal; enter the exposure day together with its three lags; replace the linear term with indicators for two-degree bins; and replace the dry-bulb maximum with the wet-bulb maximum, or enter the two together (Section 5.3).

### *4.2. Identification*

The design controls for place, round and season jointly, so that a hot day is interpreted as heat that is unusual for that district at that time of year. Identification rests on one assumption, which we label A1 and test throughout Section 6.

**Assumption A1.** *Conditional on the fixed effects and rainfall, the daily dry-bulb and wet-bulb maxima on a person's exposure day are orthogonal to their potential biomarker outcomes.*

The assumption is a statement about survey logistics. DHS fieldwork periods are fixed before enumeration begins, and field teams are assigned areas by language and to keep travel short (ICF International, 2012b); clusters are drawn from the census frame, and the date on which a particular cluster is reached depends on the team's position in its route, on travel distance and on how long the previous clusters took, not on the health of the households

waiting at the end of the road. Because a team stays in a district for several weeks, the same district in the same month supplies both hot and mild measurement days. A glucose test in Rajshahi in November 2017, for example, is compared with other tests in Rajshahi during the 2017–18 round, after removing what is typical of Rajshahi in November across the three rounds; it is never compared with a test in Sylhet, or with one in Rajshahi in another season.

The assumption need not hold in every respect. DHS interviewers complete fewer interviews per hour on the hottest and most humid days and adapt by starting earlier and working longer (LoPalo, 2023), so hot days could change who is measured, how many are measured and when. Section 6 tests each of these directly.

*What $\beta$ identifies.* $\beta$ is a physiological quantity, not a measurement error. A respondent whose systolic pressure reads lower on a hot day has lower pressure that day; the instrument is working correctly and the reading is right. Section 6.4 tests the exception to this, that the instruments themselves respond to temperature.

*What $\beta$ does not identify.* For the same reason, $\beta$ is not an effect on chronic disease. Hypertension, diabetes and anaemia are defined by states that persist, and a single reading that crosses a diagnostic line on a hot day tells us about the day, not about the condition: a person whose glucose reads higher on a hot morning has not acquired diabetes. In the language of the

measurement-error literature, heat adds a non-classical, weather-dependent error to the classification (Bound et al., 2001), and misclassification of a binary outcome biases models of it in its own right (Hausman et al., 1998). We write "measured hypertension" and "measured diabetes" when the outcome is a classification, and Section 5.4 traces how a shift in the reading becomes a shift in the classification. $\beta$ is also a short-run quantity. We observe a same-day state and, for glucose, its accumulation over a few days, but no mortality, hospitalisation or clinical diagnosis, so we cannot say what a hot day leaves behind. And the design cannot speak to extreme heat: exposure-day maxima span 19.1°C to 36.5°C, and no day reaches 37°C (Section 2.3).

## 5. Results

### *5.1. Blood Pressure Falls and Glucose Rises on Hotter Days*

Before turning to the diagnostic classifications, we estimate the physiological response itself. Table 2 reports estimates of equation (1): columns 1, 3 and 4 are the biomarkers, and columns 2 and 5 are the classifications they generate, to which we return in Section 5.4.

We find three things in the level response. First, glucose and blood pressure both respond, in opposite directions. A one-degree rise in the exposure-day maximum raises fasting plasma glucose by 1.028 mg/dL (SE 0.146) and lowers systolic and diastolic pressure by 0.562 and 0.330 mmHg, all significant

**Table 2: Ambient heat and measured cardiometabolic biomarkers**

| | (1) Fasting glucose (mg/dL) | (2) Measured diabetes FPG≥126 (pp) | (3) Systolic BP (mmHg) | (4) Diastolic BP (mmHg) | (5) Measured hypertension ≥140/90 (pp) |
|---|---|---|---|---|---|
| *Panel A. Daily maximum temperature, level* | | | | | |
| Max. temperature (°C) | $+1.028^{***}$ | $+0.626^{***}$ | $-0.562^{***}$ | $-0.330^{***}$ | $-0.839^{***}$ |
| | (0.146) | (0.143) | (0.077) | (0.038) | (0.145) |
| *Panel B. Deviation from the district's own day-of-year normal* | | | | | |
| Temperature anomaly (°C) | $+0.836^{***}$ | $+0.558^{***}$ | $-0.486^{***}$ | $-0.288^{***}$ | $-0.696^{***}$ |
| | (0.140) | (0.137) | (0.090) | (0.043) | (0.164) |
| *Panel C. Seven-day mean anomaly* | | | | | |
| Seven-day anomaly (°C) | $+0.858^{**}$ | +0.383 | −0.189 | $-0.283^{***}$ | −0.042 |
| | (0.341) | (0.290) | (0.190) | (0.101) | (0.389) |
| Outcome mean | 107.9 | 11.2 | 119.8 | 78.4 | 20.1 |
| Observations | 33,179 | 33,179 | 34,861 | 34,861 | 34,861 |
| District × round FE | Yes | Yes | Yes | Yes | Yes |
| District × calendar-month FE | Yes | Yes | Yes | Yes | Yes |
| Hour-of-test FE | Yes | Yes | – | – | – |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. Each coefficient is from a separate regression of the outcome in the column header on the exposure-day maximum temperature: fasting plasma glucose (column 1), measured diabetes (column 2), systolic and diastolic blood pressure (columns 3 and 4) and measured hypertension (column 5). Panel A uses the level of the maximum; Panel B its deviation from the district's own day-of-year normal, computed over the full history of the series; and Panel C the seven-day trailing mean of that deviation. The exposure is the daily maximum on the day of the household interview for blood pressure, when the reading was taken, and on the day before the early-morning test for fasting glucose. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before. Fixed effects are district × round and district × calendar month (the month of the test, for glucose), plus hour-of-test fixed effects for glucose; tests taken during Ramadan are excluded from the glucose sample. Coefficients are per °C; glucose is in mg/dL, blood pressure in mmHg, and measured diabetes (fasting plasma glucose ≥126 mg/dL) and measured hypertension (≥140/90 mmHg) in percentage points. Outcome mean is the mean of the outcome in the estimation sample. Standard errors clustered on district (64 clusters) in parentheses. Stars: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

at the 1% level. Moving an exposure day from the 25th to the 75th percentile of the within-district-month temperature distribution, a shift of 1.4°C in the glucose sample and 1.5°C in the blood-pressure sample, raises measured diabetes by 0.88 and lowers measured hypertension by 1.26 percentage points, against sample means of 11.2 and 20.1%.

Second, the estimates do not reflect seasonality the fixed effects leave in place. Panel B replaces the level with the deviation of the day's maximum from the district's own day-of-year normal, which removes anything common to a district's calendar. The estimates fall by a tenth to a fifth but keep their sign, their significance at the 1% level and their order of magnitude.

Third, the glucose response accumulates over several days while the blood-pressure response does not. Table 3 enters the exposure day and the three preceding days together: glucose loads on the exposure day (the day before the test) and on the day before that, with a cumulative 1.254 mg/dL per degree over four days, while for blood pressure the same-day coefficient (-0.468) carries most of the four-day cumulative -0.632 and no individual lag is significant. This is the contrast one would expect between a metabolic adjustment over several days and an acute vasomotor one, with the qualification that the pressure response also tracks the temperature of the surrounding fortnight (Section 6.5).

**Table 3: Dynamics and heterogeneity**

| | (1) $t$ | (2) $t-1$ | (3) $t-2$ | (4) $t-3$ |
|---|---|---|---|---|
| *Panel A. Distributed lags, fasting glucose (mg/dL)* | | | | |
| Max. temperature | $+0.708^{***}$ | $+0.424^{**}$ | $+0.085$ | $+0.037$ |
| | (0.143) | (0.194) | (0.172) | (0.152) |
| Cumulative $t$ to $t-3$ | | 1.254 | | |
| Observations | | 33,179 | | |
| *Panel B. Distributed lags, systolic blood pressure (mmHg)* | | | | |
| Max. temperature | $-0.468^{***}$ | $-0.116$ | $-0.049$ | $+0.001$ |
| | (0.112) | (0.145) | (0.133) | (0.095) |
| Cumulative $t$ to $t-3$ | | -0.632 | | |
| Observations | | 34,861 | | |

| *Panel C. Heterogeneity by age* | Under 50 | Interaction, age ≥50 |
|---|---|---|
| Fasting glucose | $+0.943^{***}$ | $+0.271^{*}$ |
| | (0.160) | (0.141) |
| Systolic BP | $-0.392^{***}$ | $-0.537^{***}$ |
| | (0.073) | (0.085) |
| Measured hypertension | $-0.533^{***}$ | $-0.965^{***}$ |
| | (0.143) | (0.136) |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. Panels A and B report, for fasting glucose and systolic blood pressure, one regression each in which the daily maximum on the exposure day $t$ and on each of the three preceding days enter together; the cumulative row is the sum of the four coefficients. For glucose $t$ is the day before the test, for blood pressure the day of the interview. Panel C reports, for each outcome, one regression of the outcome on the exposure-day maximum and its interaction with an indicator for age 50 and over: the first column is the slope for adults under 50 and the second the interaction, so the slope for adults aged 50 and over is their sum. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before. Fixed effects are district × round and district × calendar month (the month of the test, for glucose), plus hour-of-test fixed effects for glucose; tests taken during Ramadan are excluded from the glucose sample. Coefficients are per °C; glucose is in mg/dL, blood pressure in mmHg, and measured diabetes (fasting plasma glucose ≥126 mg/dL) and measured hypertension (≥140/90 mmHg) in percentage points. Standard errors clustered on district (64 clusters) in parentheses. Stars: $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

*Older adults.* Panel C of Table 3 reports the heterogeneity that matters clinically. The blood-pressure response is 2.4 times larger among adults aged 50 and over: the interaction is -0.537 mmHg (SE 0.085) on a base of -0.392 mmHg for younger adults, and -0.965 percentage points (SE 0.136) for measured hypertension. The glucose response varies less with age (interaction 0.271, SE 0.141, $p = 0.060$), so the age gradient is concentrated on the cardiovascular margin. The response may also be larger among respondents taking antihypertensive medication, but that interaction is significant only at the 10% level (-0.289, $p = 0.060$). We observe no hospitalisation, fall or death, so we cannot say what follows from a larger fall in pressure among older adults; it is, however, the proximate mechanism through which heat-wave morbidity is usually understood to operate (Kenney et al., 2014; Liu et al., 2022).

*The shape of the dose-response.* Figure 2 replaces the linear term with indicators for two-degree bins, relative to a 28–30°C exposure day. Both responses are monotone across the observed range, with no threshold or kink inside the support, which is consistent with a physiological response that is continuous in thermal load. The support ends at 36.5°C.

### *5.2. Haemoglobin, and What It Says about the Glucose Result*

The obvious mechanical explanation for a rise in a blood concentration on hot days is dehydration: if heat concentrated the blood, glucose would rise

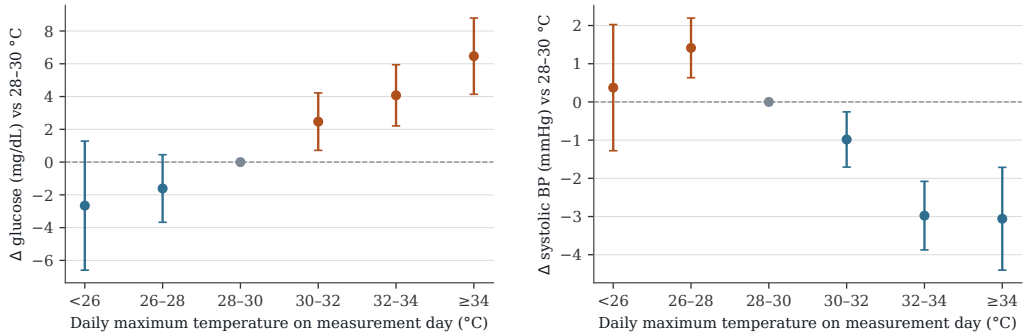



**Figure 2: Dose-response: fasting glucose (left) and systolic blood pressure (right).**

*Notes:* BDHS 2011, 2017–18 and 2022 with ERA5 temperature. Each point is the coefficient on an indicator for a two-degree bin of the exposure-day maximum (the day before the test for fasting glucose, in mg/dL; the interview day for systolic blood pressure, in mmHg), relative to a 28–30°C day, from one regression per outcome, with 95% confidence intervals from standard errors clustered on district. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before; fixed effects are district × round and district × calendar month, plus hour of test for glucose.

because the plasma in which it is dissolved had shrunk, and the finding would concern hydration rather than metabolism. Haemoglobin concentration distinguishes the two, because dehydration would raise it and plasma-volume expansion would lower it.

Table 4 reports the test on the one round that measured haemoglobin. BDHS 2011 collected it only from women aged 15–49 and children aged 6–59 months, so the test runs on a different and much smaller population from the rest of the paper. We report the two groups separately and pooled, classify anaemia at the WHO threshold for each, keep the de facto household members (those who slept in the household the night before), as DHS does, and exclude pregnant women, since pregnancy expands plasma volume through the very

mechanism under test, by up to half by the third trimester (Aguree and Gernand, 2019). Capillary haemoglobin measured with a HemoCue is analytically precise but varies from day to day within a person (Morris et al., 1999), and WHO now recommends venous samples for diagnosing anaemia (World Health Organization, 2024), so estimates from capillary readings in one round are noisy.

We find that the point estimates for haemoglobin are negative in both groups and in the pooled sample, though not significant on the dry bulb. Haemoglobin falls by 0.017 g/dL per degree among women (SE 0.024) and by 0.016 among children (SE 0.017), pooling to 0.018 (SE 0.018, $p = 0.310$), and measured anaemia rises by 0.726 percentage points (SE 0.571) in the pooled sample. A seven-day trailing anomaly strengthens the classification result, to 2.475 points per degree (SE 0.928; reported in the note to the table). Apart from that seven-day estimate, however, no dry-bulb estimate is significant once rainfall is held fixed. Against the wet-bulb maximum, the exposure the physiology points to, haemoglobin falls 0.037 g/dL per degree in the pooled sample (SE 0.020, $p = 0.069$), with a larger $t$-statistic than on the dry bulb but significant only at the 10% level, and imprecisely in each group on its own (column 2; Section 5.3 takes up the comparison).

Two things follow, with the weight of this evidence attached to them. First, the evidence argues against dehydration as the source of the rise in glucose: a concentrating artefact would raise haemoglobin, and haemoglobin falls in

every cut. This conclusion rests on the sign of imprecise coefficients estimated in a different population, and it does not sign the remaining bias in the glucose estimate. Dilution of the plasma would push measured glucose down, but point-of-care analysers tend to read glucose high when haematocrit is low (Tang et al., 2000); both effects are small beside the estimated response, since haemoglobin falls by a fraction of a per cent per degree, and we do not treat the glucose estimate as a bound in either direction. Second, the three biomarkers move as a single physiological response would move them: plasma volume expands and vessels dilate, so haemoglobin and pressure fall, while thermal stress raises counter-regulatory hormones and hepatic glucose output, so glucose rises.

### *5.3. Humid Heat Carries the Response*

The body sheds heat by convection, which works against the dry-bulb temperature, and by evaporation, which works against the humidity of the air and whose capacity the wet-bulb temperature measures. If the estimates above are physiological rather than mechanical, they should track the wet bulb more closely than the dry bulb, because evaporation is the body's main defence against heat. Table 5 enters the two exposures separately and then together. This comparison was not pre-specified: we added it after the dry-bulb estimates were in hand, and Section 7.2 reports multiple-testing-adjusted $p$-values for it.

**Table 4: Ambient heat, haemoglobin and measured anaemia, BDHS 2011**

| | Haemoglobin (g/dL) | | Measured anaemia (pp) | |
|---|---|---|---|---|
| | Dry-bulb (1) | Wet-bulb (2) | Dry-bulb (3) | Wet-bulb (4) |
| *Panel A. Women aged 15–49* | | | | |
| Daily maximum (°C) | −0.017 | −0.040 | +0.835 | +1.195 |
| | (0.024) | (0.026) | (0.663) | (0.808) |
| Outcome mean | 12.17 | 12.17 | 41.2 | 41.2 |
| Observations | 5,333 | 5,333 | 5,333 | 5,333 |
| *Panel B. Children aged 6–59 months* | | | | |
| Daily maximum (°C) | −0.016 | −0.030 | +0.361 | +1.457* |
| | (0.017) | (0.023) | (0.724) | (0.734) |
| Outcome mean | 10.78 | 10.78 | 51.5 | 51.5 |
| Observations | 2,361 | 2,361 | 2,361 | 2,361 |
| *Panel C. Both groups pooled* | | | | |
| Daily maximum (°C) | −0.018 | −0.037* | +0.726 | +1.287* |
| | (0.018) | (0.020) | (0.571) | (0.686) |
| Outcome mean | 11.74 | 11.74 | 44.4 | 44.4 |
| Observations | 7,694 | 7,694 | 7,694 | 7,694 |

*Notes:* Data: BDHS 2011, the only round that measured haemoglobin, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. The sample is women aged 15–49 (Panel A) and children aged 6–59 months (Panel B), the only groups the survey measured, pooled in Panel C; only de facto household members (those who slept in the household the night before) are kept, as in DHS, and pregnant women are excluded, because pregnancy expands plasma volume through the mechanism under test. Each coefficient is from a separate regression of capillary haemoglobin (g/dL, columns 1 and 2) or measured anaemia (percentage points, columns 3 and 4) on the daily maximum on the day of the household interview, when the sample was taken: the dry-bulb maximum in columns 1 and 3 and the wet-bulb maximum, which rises with humidity at a given dry-bulb temperature, in columns 2 and 4. Measured anaemia is haemoglobin below 12 g/dL for non-pregnant women and below 11 g/dL for children. Controls are age, age squared, sex (except in Panel A), urban residence, education, wealth quintile and CHIRPS rainfall on the interview day and the day before; with one round, the fixed effects are district × calendar month. Replacing the dry-bulb maximum with its seven-day trailing anomaly gives +2.475 (SE 0.928) on measured anaemia in the pooled sample, against the column-3 estimate. Standard errors clustered on district in parentheses. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

We find two things. First, entered alone, the wet bulb gives the larger coefficient on every outcome: 1.476 mg/dL per degree for glucose against 1.028 for the dry bulb, and 0.883 mmHg against 0.562 for systolic pressure. A degree of wet bulb is not a degree of dry bulb, since the two have different variances within a cell; scaled by its own standard deviation within district × round × month cells, a one-standard-deviation move gives 1.61 against 1.37 mg/dL of glucose and 0.94 against 0.74 mmHg of systolic pressure, so the gap narrows but does not close.

Second, entered together, which identifies the wet-bulb coefficient from days equally hot in dry-bulb terms but differently humid, the wet bulb keeps three-quarters or more of its stand-alone coefficient and the dry bulb gives most of it up. For glucose the split is 1.208 (SE 0.188) against 0.364 (SE 0.145); for systolic pressure -0.778 (SE 0.131) against -0.146 (SE 0.098). For the two classifications the dry-bulb term is not significant at the 5% level ($p = 0.065$ and $p = 0.267$), while the wet-bulb term is significant at the 1% level. For haemoglobin the dry-bulb term disappears: entering both exposures leaves the wet bulb carrying the whole of the effect (-0.039, SE 0.020, $p = 0.060$) with the dry bulb at 0.003, so the fall in haemoglobin, though imprecise, lines up with humid heat, which the body cannot shed by evaporation. Online Appendix Table A.9 applies the inference of Section 7.2 to the wet-bulb estimates: Conley standard errors at 30 and 50 km are close to the district-clustered ones, and the wild cluster bootstrap gives $p = 0.001$ for glucose, 0.001 for measured diabetes, 0.001 for systolic pressure and 0.002

for measured hypertension.

*Collinearity.* One concern is that the two exposures are too collinear to be entered together. Within a district, round and calendar month the two series have standard deviations of 1.34 and 1.09°C in the glucose sample and 1.32 and 1.07 in the blood-pressure sample, with correlations of 0.60 and 0.59. After the fixed effects and controls are partialled out, which is the variation the estimator uses, the correlation is 0.64 and 0.62, the variance inflation factor 1.68 and 1.62 against a conventional threshold of five, and the condition number 2.12 and 2.06 against a threshold of thirty (Online Appendix Table A.8, Panel A). Entering both inflates the wet-bulb standard errors by factors of 1.18 and 1.34, and those inflated standard errors are the ones reported. The horse race also adds no variable. Since $aT + cT_w \equiv (a + c)T - cD$ for the wet-bulb depression $D = T - T_w$, the same regression can be written in the dry-bulb maximum and the depression, and each coefficient then answers a separate question (Panel B): a degree hotter at unchanged humidity changes glucose by 1.572 mg/dL (SE 0.165) and systolic pressure by -0.924 mmHg (SE 0.097), while a degree drier at unchanged temperature changes them by -1.208 (SE 0.188) and 0.778 (SE 0.131), in the opposite direction.

*Construction of the wet bulb.* A second concern is the formula. The production series uses the Stull (2011) approximation, a regression fit at fixed sea-level pressure; Buzan et al. (2015) prefer the pseudoadiabatic method of Davies-Jones (2008), which inverts Bolton (1980), and note that the two can

differ by more than a degree in hot, dry air. Bangladesh is humid and lies near sea level, so that case is largely avoided, but we check it. Rebuilt with Davies-Jones on the same local-day window (Online Appendix Table A.8, Panel C), the two district-day series correlate 0.9996 and differ by 0.101°C on average (SD 0.122), and the estimates move from 1.208 to 1.255 for glucose and from -0.778 to -0.775 for systolic pressure. Rebuilding the production series over the UTC day in which ERA5 is stamped, rather than the Bangladesh local day, gives 1.280 and -0.794. Surface pressure over Bangladesh moves the wet bulb by less than a hundredth of a degree and is ignored.

Two cautions apply. The wet-bulb and dry-bulb maxima need not occur at the same hour, so neither is the exposure at the moment of measurement; both summarise the day. And the wet bulb reaches only 30.1°C in these data, while reanalysis smooths the most extreme humid heat, which is highly localised (Raymond et al., 2020), so this is evidence about ordinary humid heat, not about the wet-bulb thresholds that bound survivability. What the comparison establishes is narrower: the biomarker response is a response to humid heat, and a specification in dry-bulb temperature alone understates it.

### *5.4. From the Reading to the Classification*

Hypertension, diabetes and anaemia are chronic conditions, defined by states that persist, but the survey classifies each from a single reading. Table 6 and

**Table 5: Dry-bulb and wet-bulb heat, entered separately and together**

| | Entered separately | | Entered together | | |
|---|---|---|---|---|---|
| | Dry-bulb (1) | Wet-bulb (2) | Wet-bulb (3) | Dry-bulb (4) | Obs. (5) |
| Fasting glucose (mg/dL) | $+1.028^{***}$ | $+1.476^{***}$ | $+1.208^{***}$ | $+0.364^{**}$ | 33,179 |
| | (0.146) | (0.160) | (0.188) | (0.145) | |
| Measured diabetes (pp) | $+0.626^{***}$ | $+0.841^{***}$ | $+0.638^{***}$ | $+0.275^{*}$ | 33,179 |
| | (0.143) | (0.176) | (0.191) | (0.146) | |
| Systolic BP (mmHg) | $-0.562^{***}$ | $-0.883^{***}$ | $-0.778^{***}$ | $-0.146$ | 34,861 |
| | (0.077) | (0.098) | (0.131) | (0.098) | |
| Measured hypertension (pp) | $-0.839^{***}$ | $-1.310^{***}$ | $-1.148^{***}$ | $-0.226$ | 34,861 |
| | (0.145) | (0.204) | (0.286) | (0.202) | |
| Haemoglobin (g/dL) | $-0.018$ | $-0.037^{*}$ | $-0.039^{*}$ | $+0.003$ | 7,694 |
| | (0.018) | (0.020) | (0.020) | (0.016) | |
| Measured anaemia (pp) | $+0.726$ | $+1.287^{*}$ | $+1.225$ | $+0.071$ | 7,694 |
| | (0.571) | (0.686) | (0.796) | (0.595) | |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. Each row is an outcome and each pair of columns a regression on the daily dry-bulb maximum, the daily wet-bulb maximum, or both. Columns 1 and 2 enter the two exposures separately; columns 3 and 4 enter them together, so the wet-bulb coefficient is identified from days that are equally hot in dry-bulb terms and differently humid. The wet bulb is computed hour by hour from ERA5 2-metre temperature and dewpoint—relative humidity by the August-Roche-Magnus formula, wet bulb by the Stull (2011) approximation—then taken to a daily maximum and aggregated with the same inverse-distance weights as the dry-bulb series; both series are non-missing for every observation. Within a district × round × month cell the two series have standard deviations of 1.34 and 1.09°C in the glucose sample and 1.32 and 1.07 in the blood-pressure sample, with correlations of 0.60 and 0.59. The exposure is the daily maximum on the day of the household interview for blood pressure, when the reading was taken, and on the day before the early-morning test for fasting glucose. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before. Fixed effects are district × round and district × calendar month (the month of the test, for glucose), plus hour-of-test fixed effects for glucose; tests taken during Ramadan are excluded from the glucose sample. Coefficients are per °C; glucose is in mg/dL, blood pressure in mmHg, and measured diabetes (fasting plasma glucose ≥126 mg/dL) and measured hypertension (≥140/90 mmHg) in percentage points. The haemoglobin (g/dL) and measured-anaemia rows are BDHS 2011 only, on non-pregnant women aged 15–49 and children aged 6–59 months, with district × calendar-month fixed effects. Standard errors clustered on district (64 clusters) in parentheses. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 3 trace how the acute response to heat becomes a change in the classification. We find three things.

First, the response across the glucose distribution has the shape a uniform shift of every reading would produce. If heat added the same amount $\beta$ to everyone's glucose, the share of readings above a threshold $c$ would rise by approximately $\beta f(c)$, where $f(c)$ is the density of readings at $c$: most where readings are densest, and little in the sparse upper tail. Figure 3 sets the estimates against that benchmark, computed from the baseline coefficient of 1.028 mg/dL and the observed density. The share above 100 mg/dL rises 2.252 percentage points per degree against 2.455 predicted; above 110, 1.831 against 2.324; above the diabetes threshold of 126, 0.626 against 0.665; and above 140, 0.381 ($p < 0.001$) against 0.209. In the sparse upper tail the benchmark is a few hundredths of a point, and the estimates there are small in absolute terms but exceed it (0.224 above 160 and 0.157 above 200 mg/dL, $p = 0.007$ and $p = 0.011$), so the response is, if anything, not smaller among the highest readings. On its own this pattern does not distinguish a change in the reading from a change in the condition, and we do not use it that way. It identifies the mechanism behind the diagnostic effect: the respondents whose classification changes on a hot day are those whose readings already lie near the line, and with measured diabetes at 11.2% and measured hypertension at 20.1% there are many of them.

Second, the response is present among respondents not previously diagnosed.

**Table 6: From the reading to the classification**

| | Coefficient | Std. error | Obs. |
|---|---|---|---|
| *Panel A. The acute physiological response: continuous biomarkers* | | | |
| Systolic blood pressure (mmHg) | $-0.562^{***}$ | (0.077) | 34,861 |
| Diastolic blood pressure (mmHg) | $-0.330^{***}$ | (0.038) | 34,861 |
| Fasting plasma glucose (mg/dL) | $+1.028^{***}$ | (0.146) | 33,179 |
| Haemoglobin (g/dL) | $-0.018$ | (0.018) | 7,694 |
| *Panel B. Does the response concentrate among people who already have the condition?* | | | |
| Glucose, never diagnosed with diabetes | $+1.057^{***}$ | (0.168) | 33,127 |
| × previously diagnosed | $-0.623$ | (0.478) | 33,127 |
| Systolic BP, never told has high BP | $-0.480^{***}$ | (0.065) | 34,854 |
| × previously told has high BP | $-0.386^{**}$ | (0.151) | 34,854 |
| Systolic BP, not on medication | $-0.528^{***}$ | (0.069) | 34,861 |
| × on blood-pressure medication | $-0.289^{*}$ | (0.151) | 34,861 |
| *Panel C. The chronic conditions themselves* | | | |
| Ever diagnosed with diabetes (pp) | $+0.008$ | (0.104) | 33,127 |
| Ever told has high blood pressure (pp) | $-0.152$ | (0.152) | 34,854 |
| Currently on blood-pressure medication (pp) | $-0.024$ | (0.132) | 34,861 |
| *Panel D. Readings crossing a diagnostic threshold, by distance from the line* | | | |
| Fasting glucose ≥126 mg/dL (pp) [11.2% of sample] | $+0.626^{***}$ | (0.143) | 33,179 |
| Fasting glucose ≥140 mg/dL (pp) [6.1% of sample] | $+0.381^{***}$ | (0.091) | 33,179 |
| Fasting glucose ≥160 mg/dL (pp) [3.8% of sample] | $+0.224^{***}$ | (0.080) | 33,179 |
| Fasting glucose ≥200 mg/dL (pp) [2.1% of sample] | $+0.157^{**}$ | (0.060) | 33,179 |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. Each row is a separate regression of the outcome named in the row on the exposure-day maximum temperature, per °C. Panel A reports the biomarkers. Panel B asks whether the response is concentrated among people who already have the condition: each pair of rows comes from one regression with an interaction, where the first row is the slope for respondents without the diagnosis or medication and the second the interaction, so the slope for those with it is the sum. Panel C regresses the respondent's own report of a past diagnosis, or of current medication, in percentage points; these record the chronic condition and cannot depend on the day's weather, so they are placebo tests for whether hot days reach a different population. Panel D regresses indicators (in percentage points) for a fasting glucose reading at or above each threshold; the bracketed figure is the share of the sample above the threshold. The exposure is the daily maximum on the day of the household interview for blood pressure, when the reading was taken, and on the day before the early-morning test for fasting glucose. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before. Fixed effects are district × round and district × calendar month (the month of the test, for glucose), plus hour-of-test fixed effects for glucose; tests taken during Ramadan are excluded from the glucose sample. The haemoglobin row is the exception: it is BDHS 2011 only, on non-pregnant women aged 15–49 and children aged 6–59 months, with district × calendar-month fixed effects. Standard errors clustered on district (64 clusters) in parentheses. Stars: $^{*}\,p<0.10$, $^{**}\,p<0.05$, $^{***}\,p<0.01$.

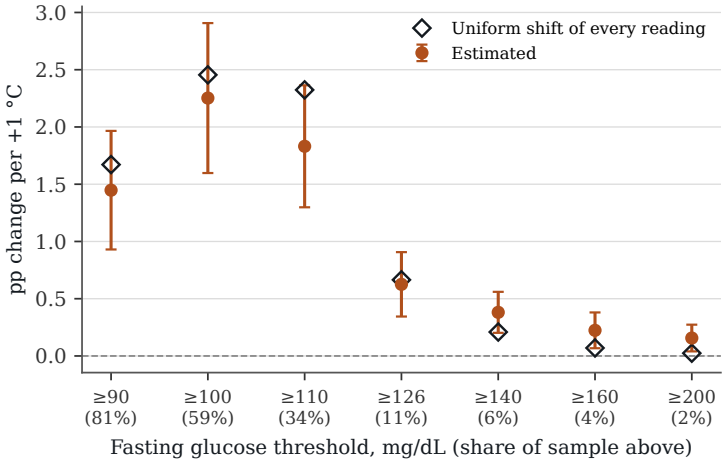



**Figure 3: The share of adults above each fasting-glucose threshold.**

*Notes:* BDHS 2011, 2017–18 and 2022 glucose sample with ERA5 temperature. Each filled point is a separate regression of an indicator for a fasting glucose reading above the threshold (in percentage points) on the maximum temperature on the day before the test, per °C, with its 95% confidence interval from standard errors clustered on district. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before; fixed effects are district × round and district × calendar month, plus hour of test for glucose. Hollow diamonds give the change a uniform shift of every reading by the baseline glucose coefficient would produce: that coefficient times the share of readings within 5 mg/dL of the threshold, per mg/dL. The share of the sample above each threshold is given under the axis.

In Panel B of Table 6, respondents never diagnosed with diabetes show a glucose response of 1.057 (SE 0.168), and the interaction with a past diagnosis is -0.623 (SE 0.478, $p = 0.197$): the response is smaller among those who report a diagnosis, but not significantly so. Treatment of the diagnosed could blunt the acute response; but the diagnosed also differ in age and in where their readings lie, and the interaction is neither precise nor a causal estimate of the effect of treatment. For blood pressure the diagnosed respond more (-0.386, SE 0.151), consistent with impaired vasomotor regulation. In both cases the response is present among the undiagnosed rather than confined to people already known to be ill.

Third, reported diagnoses do not respond to heat. Panel C regresses each respondent's own report of a past diagnosis on the exposure-day maximum. Self-reported diagnoses are an imperfect measure of the condition (Johnston et al., 2009), but a diagnosis received in the past cannot depend on this morning's weather, so these regressions are placebo tests: a nonzero coefficient would mean that hot days reach a different population. Having ever been told of diabetes moves by 0.008 percentage points per degree (SE 0.104), having been told of high blood pressure by -0.152 (SE 0.152), and currently taking blood-pressure medication by -0.024 (SE 0.132), none of them significant.

Heat therefore moves readings across the whole measured population, the classification moves because readings near the line cross it, and nothing in

these data suggests that hot days reach a different population. A prevalence estimated from a single reading inherits a weather component that is unrelated to how many people have the condition.

## 6. Threats to Identification

Assumption A1 requires that nothing that determines *which day* a respondent is measured is also correlated with what that respondent's biomarkers would otherwise have been. This section takes the assumption apart into the ways it could fail and tests each one, stating in each case the concern, the test and the outcome. Two concerns are about who and how many are measured, two about the measurement itself, and two about the timing of the exposure.

### *6.1. Does Heat Change Who Is Measured?*

If field teams responded to the weather, by postponing work on very hot days or by reaching different kinds of households on hot days than on mild ones, the measured population would differ with temperature. Online Appendix Table A.10 regresses each of 12 predetermined characteristics on the exposure-day maximum, with the fixed effects of the main specification, in the full sample of biomarker-eligible adults whether or not a measurement was obtained (38,813 people, of whom 38,787 have a complete set of controls; the 26 others are missing education).

3 of the 12 estimates are significant at the 5% level, against about 0.6 ex-

pected by chance: education, the wealth quintile and a past diagnosis of high blood pressure. All three are small; the largest is 0.0212 standard deviations per degree. Figure 4 shows the estimates in standard-deviation units. The biomarker regressions control for education and wealth, dropping every individual control moves the estimates only slightly (Table 7), and on the estimation samples the past diagnosis itself does not respond to heat (Section 5.4). We read these imbalances as too small to generate the results, but we report them among the results that are not clean.

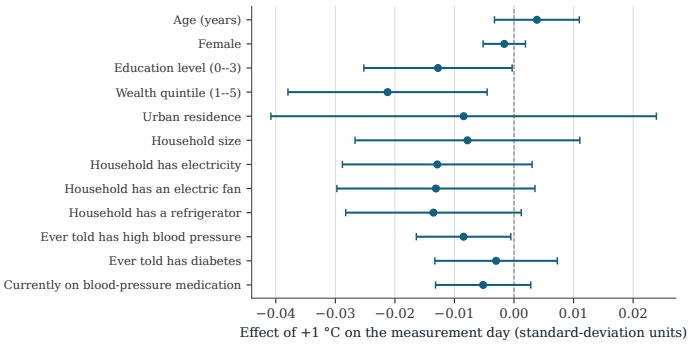



**Figure 4: Balance: predetermined characteristics and exposure-day heat**

*Notes:* BDHS 2011, 2017–18 and 2022, all biomarker-eligible adults whether or not measured, with ERA5 temperature. Each row is a separate regression of the characteristic on the daily maximum on the day the household was interviewed, with district × round and district × calendar-month fixed effects, rescaled by the standard deviation of the characteristic. Bars are 95% confidence intervals from standard errors clustered on district.

### *6.2. Does Heat Change How Many Are Measured?*

The same concern can be tested from the side of the fieldwork. Collapsing the data to district-days (Online Appendix Table A.11, Panel A), neither the

number of eligible adults visited, nor the number of blood-pressure measurements taken, nor the share of eligible adults measured responds to the day's temperature. At the individual level (Panel B), whether a glucose measurement exists does not respond to heat either (-0.113 percentage points per degree, $p = 0.470$).

Two estimates in this group are not clean. Consent to the blood-pressure measurement falls by 0.325 percentage points per degree ($p = 0.002$) on a base of 91%, and whether a blood-pressure reading exists changes by -0.268 percentage points ($p = 0.083$). A bound shows how small these margins are. Removing a share $s$ of respondents whose readings differ from everyone else's by $D$ moves the mean by about $sD$, so for selection alone to produce the systolic estimate the respondents lost per degree of heat would need readings about 210 mmHg away from those measured; for glucose, on its own selection margin, about 906 mg/dL. In addition, women's employment, which only the women's questionnaire records, falls with heat in 2017–18; this is an imbalance across survey clusters that leaves the biomarker estimates unchanged when it is controlled for (Online Appendix A.4). We take these margins as too small to threaten A1, but a reader may weigh them differently.

### *6.3. Is It the Fieldwork, or the Physiology?*

A further concern is that heat degrades the measurement rather than the person: a tired team on a hot afternoon, a rushed protocol, or an instrument used outside its operating range. We test the first two here and the third in

Section 6.4.

*Fieldworkers.* We add fixed effects for the fieldworker code recorded with the measurement: the biomarker specialist who took the blood pressure in 2022, and the measurer in the 2017–18 biomarker questionnaire, who also took height and weight. BDHS 2011 records no such code, so the test runs on 78% of the glucose sample and 77% of the blood-pressure sample, and the comparison with the baseline is not a change of specification alone. The estimates are 0.861 mg/dL for glucose and -0.604 mmHg for systolic pressure (Table 7).

*Cuffs and monitors.* BDHS 2022 records the cuff size used for each respondent. Adding cuff-size fixed effects in that round gives -0.677 mmHg, so the gradient is not an artefact of which cuff a respondent's arm required. Because each cuff size has its own monitor in 2022—a UA-767PVS for the small cuff, a UA-767F/FAC for the medium and a UA-789AC for the extra-large (National Institute of Population Research and Training (NIPORT) and ICF, 2024)—these are also monitor fixed effects for that round.

*Placebo outcomes from the same visit.* Panel C of Online Appendix Table A.11 uses outcomes recorded on the same visit by the same team: arm circumference with a tape, height with a measuring board, and weight, and the body-mass index it implies, with a digital SECA scale, the instrument used for weight in every round (National Institute of Population Research

and Training (NIPORT), Mitra and Associates, and ICF International, 2013; National Institute of Population Research and Training (NIPORT) and ICF, 2020, 2024). If hot days produced sloppier fieldwork or a different mix of respondents, these would move. Arm circumference moves by 0.010 cm per degree (SE 0.018), height by 0.013 (SE 0.033), weight by -0.012 kg (SE 0.057) and body-mass index by -0.003 (SE 0.018), none of them significant and all precisely estimated. The tape and the board are the cleaner test, since they involve no electronics; the scale shows that one electronic instrument used on the same visit does not drift with the day's heat, but it is a different device from the analyser and the monitor. A past diagnosis of diabetes, which cannot depend on the day's weather, is likewise a precise zero (0.008, SE 0.104).

These tests leave little room for fieldwork quality or respondent composition to generate the results. They do not rule out temperature sensitivity of the glucose analyser or the blood-pressure monitor, which we take up next.

### *6.4. The Instruments and Their Operating Ranges*

The BDHS final reports name the instruments. Blood pressure was taken with a LifeSource (A&D) UA-767 Plus in 2011 and 2017–18 (National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013; National Institute of Population Research and Training (NIPORT) and ICF, 2020) and in 2022 with A&D monitors of the same family, one per cuff size (National Institute of Population Research

and Training (NIPORT) and ICF, 2024). The manufacturer states an operating range of 10–40°C and relative humidity below 85% (A&D Medical (LifeSource), 2009), and the UA-767 Plus meets the British Hypertension Society validation standard (Verdecchia et al., 2004). Glucose was measured with a HemoCue Glucose 201+ in 2011 (National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013), operating range 15–30°C (HemoCue AB, 2024c). For 2017–18 and 2022 the reports are internally inconsistent: their methods chapters name the HemoCue 201 RT, operating range 15–27°C (HemoCue AB, 2024b), and the notes to their results tables the Glucose 201 DM, 18–30°C (National Institute of Population Research and Training (NIPORT) and ICF, 2020, 2024; HemoCue AB, 2024a), so we apply the stricter range, 18–27°C, to those two rounds. HemoCue glucose analysers agree closely with laboratory reference methods (Stork et al., 2005), and the 201 RT with the 201+ (Segerhag et al., 2015), but those validations were run under controlled conditions.

What matters for an instrument is the temperature around it when it is used. The hourly reanalysis gives the outdoor temperature at the clock hour of each measurement; the instruments are used indoors, so this is a proxy, but a far better one than the day's maximum. Online Appendix Table A.12 restricts the sample on it, and the answer differs for the two biomarkers.

*Blood pressure.* Only 0.05% of readings were taken at an hour whose outdoor temperature lay outside the monitor's 10–40°C range, and the gradient is

present at cool reading hours as well as hot ones (-0.419 mmHg, SE 0.099, at or below 30°C). The monitor's humidity limit we cannot check, since ERA5 gives the outdoor humidity and not that of the room in which the cuff was applied.

*Glucose.* The analysers were often used near the edges of their stated range. At the hour of the test the outdoor temperature was below the 18°C floor of the stricter range for 48% of tests in the cool 2017–18 round, and above the 27°C ceiling for 42% in 2022. But the gradient does not come from those tests. Restricting to tests taken when the test-hour temperature lay inside each round's strictest stated range—15–30°C in 2011 and 18–27°C in 2017–18 and 2022, 58% of the sample—gives 0.978 mg/dL (SE 0.198, $p < 0.001$), close to the baseline of 1.028. Within the wider 15–30°C range of the 201+ (87% of tests) the estimate is 1.038, and among tests taken outside the strict range it is smaller, not larger (0.626, SE 0.234). An analyser that drifts outside its specification cannot produce a gradient that is as strong among the tests taken inside it.

These restrictions cannot exclude temperature sensitivity *within* the stated range, or heat damage to reagent cuvettes carried by the teams between tests. Consumer glucose meters drift with temperature and humidity even inside their stated ranges (Haller et al., 2007; Ginsberg, 2009), and HemoCue's haemoglobin microcuvettes are sensitive to heat and humidity (ICF International, 2012a; Whitehead et al., 2019). In a tropical hospital, glucose-

oxidase meters read lower at higher operating temperature and humidity (Pratumvinit et al., 2016); the HemoCue analysers use glucose dehydrogenase, a different chemistry (HemoCue AB, 2024c,b,a), and an artefact of that sign would work against the positive gradient we find rather than produce it. That evidence concerns other devices and reagents, but it is why we do not treat the in-range result as decisive. Neither question can be settled without the analysers' calibration and storage records, which we do not have.

### *6.5. Future Weather as a Placebo*

The natural falsification test replaces the exposure-day temperature with the temperature at a future date, which cannot have affected the reading. Panel B of Online Appendix Table A.13 shows that several short leads are not clean nulls: the three-day lead for both outcomes, the seven-day lead for glucose and the fourteen-day lead for systolic pressure. This is what one would expect, because a district's daily temperature anomaly is autocorrelated (Panel A): $\rho(T_t, T_{t+3}) = 0.346$ and $\rho(T_t, T_{t+7}) = 0.146$, falling to 0.063 at thirty days. A short-horizon lead is mechanically loaded with the exposure day's own weather, so a nonzero coefficient there is evidence of autocorrelation, not of confounding.

Two valid versions of the test are available. In a horse race that enters the exposure day and the lead together, the three-day lead shrinks, from 0.268 to 0.195 for glucose and from -0.189 to -0.150 for systolic pressure, while the exposure-day coefficients are essentially unchanged. The seven-day glucose

lead does not shrink (0.256, $p = 0.061$), and we report it among the results that are not fully clean; the seven-day systolic lead is zero alone and in the horse race (0.007, $p = 0.916$). At thirty days, where the autocorrelation is 0.063, the lead entered alone is a precise zero for both outcomes (glucose 0.135, $p = 0.374$; systolic pressure -0.033, $p = 0.764$).

One lead result is not clean. The fourteen-day lead for systolic pressure keeps a coefficient of -0.251 ($p = 0.015$) even in the horse race, larger than an autocorrelation of 0.090 can produce mechanically. We read this as blood pressure responding to the thermal environment over roughly a fortnight, since plasma-volume expansion is an acclimation process that takes several days, so that a lead falling inside the same warm episode captures genuine exposure. The thirty-day placebo is clean (-0.033, SE 0.110), and the estimate is -0.577 under district × round × month fixed effects against -0.562 in the baseline (Table 7), so the result is not evidence against A1; it is a reason to describe $\beta$ for blood pressure as the response to the recent thermal environment rather than to a single day.

### *6.6. Heat Before and After the Test*

The fasting glucose test is drawn early in the morning (median hour 7), and the day's afternoon peak comes after it. This is why the baseline glucose exposure is the previous day's maximum (Section 4.2), and it gives a further falsification test: heat that arrives after the blood was drawn cannot have changed the reading. Online Appendix Table A.14 reports it.

Panel A starts from the baseline of 1.028 mg/dL per degree (SE 0.146). The test day's own maximum, entered alone, gives 0.829, as it must for a series correlated 0.92 with the day before. Entered together, the day before keeps 0.790 and the test day retains 0.381 (SE 0.185, $p = 0.043$), which is not the zero the test calls for. The hourly data show why. The test-day maximum is correlated with conditions at the moment of the test, and the temperature at the hour of the test is itself a strong predictor (1.062 entered alone). Entering all three, the test-day maximum falls to 0.106 (SE 0.195, $p = 0.590$), while the day before keeps 0.526 and the test hour carries 0.836. Heat after the test carries no effect once the heat before and at the test is held fixed.

The horse race settles one question but not the other. The previous afternoon's peak keeps an effect of its own once the temperature at the test is held fixed (0.581, SE 0.114), and no property of the analyser at the moment of testing can produce that. The test-hour coefficient, however, cannot be attributed by timing alone: the temperature at seven in the morning is both the warmth of the night the respondent has just slept through and the ambient temperature of the analyser. Section 6.4 takes up the second reading and finds the gradient as strong among tests taken within the analysers' stated operating range as in the whole sample. The maximum over the 24 hours before the test is in effect the same series as the previous day's maximum (correlation 1.00) and gives the same estimate, 1.031, and the wet-bulb versions behave alike.

Panel B repeats the exercise for blood pressure, which is read during the household visit (median hour 12, with 52% of readings taken after noon). Every timing gives a similar estimate: -0.562 on the interview day, -0.436 on the day before, -0.484 over the 24 hours before the reading and -0.468 at the hour of the reading.

## 7. Additional Robustness and Inference

Table 7 collects the specification checks, and Figure 5 plots them. Each row re-estimates the baseline of Table 2 with one change. Across the 17 glucose rows that keep the baseline exposure and vary the fixed effects, the controls or the sample, the glucose estimate lies between 0.826 and 1.152 mg/dL per degree, against 1.028 in the baseline; across the 18 such rows for systolic pressure the estimate lies between -0.677 and -0.493 mmHg, against -0.562; and every one of these estimates is significant at the 1% level. The single exception is the comparison within a survey-cluster visit, which we discuss separately below. We group the checks by the concern each addresses and, in each case, state the concern, the test and the outcome.

*Seasonality and the fixed effects.* The baseline might leave seasonal variation in place that correlates with both temperature and the biomarkers, or the fixed effects might be so tight that they leave only noise. We test this by loosening and tightening them in turn. Replacing the district × calendar-month fixed effects with national calendar-month fixed effects, which assumes

that the seasons move together across the country, gives 1.139 mg/dL for glucose and -0.562 mmHg for systolic pressure, and replacing the interacted fixed effects with separate district, round and month effects changes little. Tightening them to district × round × month, so that identification comes only from days within the same district-month of a round, gives 0.935 and -0.577. Dropping the first and last week of each round's fieldwork, when the teams' composition and routing are least typical, gives 1.027 and -0.546, and adding day-of-week fixed effects also leaves the estimates close to the baseline. Neither the choice of seasonal controls nor the edges of the fieldwork period drives the results.

*Controls, outliers and subgroups.* The estimates might depend on the individual controls, on a few extreme readings, or on a particular group of respondents. Dropping all individual controls gives 0.945 and -0.566; trimming the outer 1% of the outcome gives 0.826 and -0.519; and restricting to women, to men, to rural respondents (0.940 and -0.493) or to adults aged 35 and over leaves each estimate within the 95% confidence interval of the baseline. For blood pressure, excluding respondents on antihypertensive medication also leaves the estimate unchanged. The results are not carried by the controls, by outliers or by any one group.

*Other weather and Ramadan.* Rainfall is correlated with temperature and changes both behaviour and fieldwork, and Ramadan changes eating and fasting. Dropping the rainfall controls gives 1.033 mg/dL and -0.544 mmHg,

so the temperature coefficient is not picking up rain. Including tests taken during Ramadan, which the glucose baseline excludes, gives 1.021, and excluding Ramadan days from the blood-pressure sample, which the baseline keeps, gives -0.595.

*Pregnancy and survey weights.* Pregnancy expands plasma volume, the mechanism behind the fall in pressure and haemoglobin, and the baseline samples keep the women who were pregnant at the time of the survey (664 in the blood-pressure sample and 628 in the glucose sample). The estimates are also unweighted, while the DHS sample is drawn with unequal probabilities. We test both by excluding pregnant women and, separately, by weighting each respondent by the DHS household sample weight. Excluding pregnant women gives 1.036 mg/dL for glucose and -0.560 mmHg for systolic pressure (0.643 and -0.858 percentage points for the two classifications); weighting gives 0.931 and -0.528 (0.543 and -0.812). Neither the inclusion of pregnant women nor the use of unweighted estimates drives the results.

*The glucose scale.* The 2022 glucose readings are put on the plasma scale by the factor of 1.11 that DHS applies, and the documentation for that round conflicts (Section 3). Leaving them as recorded gives a glucose slope of 0.997 mg/dL, against 1.028 in the baseline, so the scale of the 2022 readings does not drive the glucose result. The measured-diabetes slope is more sensitive, at 0.481 (SE 0.118) against 0.626, because the 126 mg/dL threshold then falls further into the upper tail of the 2022 distribution.

*Comparisons within a single cluster visit.* The most demanding comparison is between respondents in the same survey cluster measured on different days of a single team visit, which holds fixed everything about the place and the visit. We test it by replacing the district × round fixed effects with survey-cluster fixed effects. The glucose estimate falls to 0.360 (SE 0.358) and the systolic estimate changes sign, to 0.122 (SE 0.196), neither significant and both far from the baseline. Two features limit what this comparison can show. First, teams stay in a cluster for a day or two, so the within-cluster variance of temperature is 14% of the within-district-month variance the baseline uses in the glucose sample and 10% in the blood-pressure sample (standard deviations of 0.51 and 0.42°C, against 1.34 and 1.32); with so little signal, error in a district-level reanalysis exposure attenuates the estimate far more than in the baseline. Second, the days of one visit do not separate respondents at random, since those measured after the first day are the ones the team had to return for. The comparison nonetheless does not reproduce the baseline, and we report it as a result that is not clean. Our estimates rest on comparisons across clusters within a district and round, net of the district's seasonal pattern, and in 2017–18 the clusters reached on hotter days differed in at least one predetermined characteristic, women's employment, although controlling for it leaves the biomarker estimates unchanged (Online Appendix A.4).

**Table 7: Specification robustness**

| | Coefficient | Std. error | Obs. |
|---|---|---|---|
| *Panel A. Fasting glucose (mg/dL)* | | | |
| Baseline | $+1.028^{***}$ | (0.146) | 33,179 |
| National calendar-month FE | $+1.139^{***}$ | (0.120) | 33,179 |
| District + round + month FE | $+1.134^{***}$ | (0.114) | 33,179 |
| Add day-of-week FE | $+1.012^{***}$ | (0.146) | 33,179 |
| District × round × month FE | $+0.935^{***}$ | (0.136) | 33,179 |
| Add measurer FE | $+0.861^{***}$ | (0.184) | 25,949 |
| Survey-cluster FE | +0.360 | (0.358) | 33,179 |
| Include Ramadan days | $+1.021^{***}$ | (0.144) | 33,376 |
| Trim outer 1% of the outcome | $+0.826^{***}$ | (0.123) | 32,523 |
| No individual controls | $+0.945^{***}$ | (0.142) | 33,194 |
| No rainfall controls | $+1.033^{***}$ | (0.128) | 33,179 |
| Exclude pregnant women | $+1.036^{***}$ | (0.148) | 32,555 |
| Household sampling weights | $+0.931^{***}$ | (0.176) | 33,179 |
| Women only | $+1.001^{***}$ | (0.195) | 18,216 |
| Men only | $+1.045^{***}$ | (0.176) | 14,963 |
| Rural only | $+0.940^{***}$ | (0.161) | 21,777 |
| Age 35 and over only | $+1.152^{***}$ | (0.177) | 22,162 |
| Drop first and last fieldwork week | $+1.027^{***}$ | (0.148) | 30,471 |
| 2022 readings without the ×1.11 conversion | $+0.997^{***}$ | (0.141) | 33,179 |
| Day-of-year anomaly | $+0.836^{***}$ | (0.140) | 33,179 |
| Wet-bulb maximum | $+1.476^{***}$ | (0.160) | 33,179 |
| Test-day maximum (mostly after the test) | $+0.829^{***}$ | (0.155) | 33,179 |
| Placebo: temperature at $t + 30$ | +0.135 | (0.151) | 33,179 |
| *Panel B. Systolic blood pressure (mmHg)* | | | |
| Baseline | $-0.562^{***}$ | (0.077) | 34,861 |
| National calendar-month FE | $-0.562^{***}$ | (0.082) | 34,861 |
| District + round + month FE | $-0.608^{***}$ | (0.072) | 34,861 |
| Add day-of-week FE | $-0.550^{***}$ | (0.079) | 34,861 |
| District × round × month FE | $-0.577^{***}$ | (0.079) | 34,861 |
| Add measurer FE | $-0.604^{***}$ | (0.094) | 26,975 |
| Survey-cluster FE | +0.122 | (0.196) | 34,861 |
| Add BP cuff-size FE (2022) | $-0.677^{***}$ | (0.148) | 13,980 |
| Drop Ramadan days | $-0.595^{***}$ | (0.078) | 33,856 |
| Trim outer 1% of the outcome | $-0.519^{***}$ | (0.069) | 34,202 |
| No individual controls | $-0.566^{***}$ | (0.083) | 34,879 |
| No rainfall controls | $-0.544^{***}$ | (0.074) | 34,861 |
| Exclude respondents on BP medication | $-0.567^{***}$ | (0.068) | 31,171 |
| Exclude pregnant women | $-0.560^{***}$ | (0.077) | 34,197 |
| Household sampling weights | $-0.528^{***}$ | (0.090) | 34,861 |
| Women only | $-0.524^{***}$ | (0.095) | 19,102 |
| Men only | $-0.602^{***}$ | (0.116) | 15,759 |
| Rural only | $-0.493^{***}$ | (0.108) | 22,745 |
| Age 35 and over only | $-0.659^{***}$ | (0.100) | 23,393 |
| Drop first and last fieldwork week | $-0.546^{***}$ | (0.085) | 32,575 |
| Day-of-year anomaly | $-0.486^{***}$ | (0.090) | 34,861 |
| Wet-bulb maximum | $-0.883^{***}$ | (0.098) | 34,861 |
| Placebo: temperature at $t + 30$ | −0.033 | (0.110) | 34,861 |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with

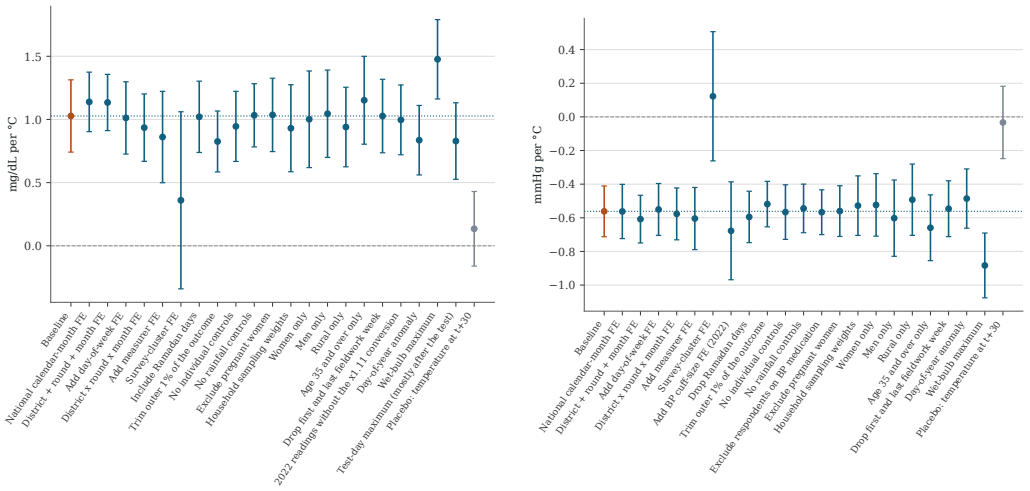



**Figure 5: Specification curves: fasting glucose (left) and systolic blood pressure (right)**

*Notes:* Each point is one row of Table 7: the coefficient on the exposure-day maximum (the day before the test for glucose, the interview day for blood pressure) from a re-estimate of the baseline with one change, with its 95% confidence interval from standard errors clustered on district. BDHS 2011, 2017–18 and 2022 with ERA5 temperature. The dotted horizontal line marks the baseline estimate, and the final entry in each panel is the placebo that replaces the exposure with the temperature thirty days later.

### *7.1. Occupation*

The response might be occupational rather than physiological: carried by respondents who work outdoors and take the day's heat directly, while everyone else shelters from it. If so, the response should be concentrated in agricultural and manual work. We test this by merging the women's questionnaire, which records employment and occupation, onto the measured sample, which gives 10,151 women aged 15–49 in the 2017–18 and 2022 rounds, 27.5% of them in agricultural or unskilled manual work, and interacting the exposure-day maximum with an indicator for that work. The interaction is small and not significant for any outcome (Online Appendix Table A.6, Panel A): 0.033 (SE 0.205) for glucose, -0.052 (SE 0.180) for measured diabetes, 0.077 (SE 0.113) for systolic pressure and 0.148 (SE 0.252) for measured hypertension, while the main effects keep their sign and their significance in this smaller sample of women. The acute response is therefore not an occupational exposure.

A broader split, between any current work and none, cannot be read the same way. Pooled, the interaction is positive for the two glucose outcomes (0.170, SE 0.194; 0.243, SE 0.177), but it is absent in 2017–18 (-0.037 and -0.139) and positive but imprecise in 2022 (0.449 and 0.696; Panel C). The employment question changed between those rounds—49% of these women report current work in 2017–18 against 32% in 2022—so the split does not measure the same thing twice, and we draw no conclusion from a heterogeneity that appears only where the measure changes. The null on the kind of work, which is the

comparison the occupational story implies, holds in the pooled sample.

### *7.2. Inference*

District-clustered standard errors might overstate precision. There are only 64 districts, fieldwork is spread over a few calendar years, and because the regressor is weather, a hot day is hot over much of the country, so errors may be correlated across districts measured on the same day, which district clustering does not allow (Cameron and Miller, 2015; MacKinnon et al., 2023). Table 8 reports four checks on the baseline estimates.

*Correlation across districts on the same day.* We cluster two ways, on district and on the calendar date of the exposure (Cameron et al., 2011), which gives 373 date clusters in the glucose sample and 372 in the blood-pressure sample. The standard errors are 0.144 for glucose and 0.094 for systolic pressure, against 0.146 and 0.077 under district clustering, and every headline estimate remains significant at the 1% level.

*Spatial correlation.* Conley spatial HAC standard errors (Conley, 1999) allow unrestricted correlation within a district and add correlation between districts whose centroids lie within a band. At 30 km they are 0.147 for glucose and 0.076 for systolic pressure, and at 50 km 0.145 and 0.079, within a few per cent of the district-clustered values for both outcomes (Table 8 shows the classifications, where the 50 km values differ by up to a seventh). This is partly by construction: the average district has 1.41 district centroids

within 30 km, including its own, and 3.31 within 50 km. Districts average roughly 2,300 km$^2$, so a much wider band would span most of a division and cease to be a local spatial correction.

*Few clusters.* With 64 clusters the $t$ approximation may matter, so we also report a restricted wild cluster bootstrap-$t$ with Rademacher weights (Cameron et al., 2008; MacKinnon and Webb, 2017; Roodman et al., 2019). It gives $p = 0.001$ for glucose, 0.001 for measured diabetes, 0.001 for systolic pressure and 0.001 for measured hypertension; values of 0.001 are the floor the $(r + 1)/(B + 1)$ convention allows with 999 replications.

*Influential districts.* Re-estimating the baseline 64 times, leaving out one district at a time, the glucose estimate ranges from 0.944 to 1.065, against 1.028, and the systolic estimate from -0.580 to -0.536, against -0.562. No district drives either result.

Online Appendix Table A.3 applies the same checks, apart from leaving out districts, to every other estimate the paper relies on: the haemoglobin panel, the age heterogeneity, the wet-bulb estimates, the tightest specification, the distributional results and the placebos. The results hold with one exception. The haemoglobin panel rests on one round, a different population and 7,694 observations, and on the dry bulb its wild-bootstrap $p$-values are 0.340 for measured anaemia and 0.490 for haemoglobin, neither significant at conventional levels. Against the wet bulb, the exposure the response tracks (Section 5.3), the haemoglobin estimate with the dry bulb held fixed is sig-

nificant at the 10% level under the wild bootstrap ($p = 0.062$) and at the 5% level under two-way clustering ($p = 0.044$), although it does not survive the multiple-testing adjustment below. Measured anaemia, with district-clustered standard errors, is significant at the 10% level on its own ($p = 0.065$) but not once the dry bulb is held fixed ($p = 0.129$). The haemoglobin evidence is therefore enough to indicate the direction of the response, and not enough to support any claim about its magnitude.

*Multiple testing.* The paper reports many outcomes, exposures and subgroups, and none of the analyses was pre-registered; the wet-bulb comparison in particular was added after the dry-bulb estimates were in hand. Some of the significant results could therefore be chance. We test this by computing Romano–Wolf step-down $p$-values (Romano and Wolf, 2005), which control the family-wise error rate, for two families of hypotheses: the four headline estimates and diastolic pressure (Table 2), and the wet-bulb coefficients of the horse race, with the dry-bulb maximum held fixed, for glucose, measured diabetes, systolic pressure, measured hypertension, haemoglobin and measured anaemia. The bootstrap uses the same draw for the same district across every outcome in a family, which preserves the dependence between them (Online Appendix Table A.4). All five keep an adjusted $p$-value of 0.001. Among the wet-bulb coefficients, glucose, measured diabetes, systolic pressure and measured hypertension keep adjusted $p$-values of at most 0.008, while haemoglobin (0.102) and measured anaemia (0.148) do not survive the

adjustment. The blood-pressure and glucose results, dry and wet, are not an artefact of testing many hypotheses; the haemoglobin result is the one we treat as suggestive.

## 8. Adaptation across Rounds

Between our first and last round, household electricity in the estimation sample rose from 62% to 99% and electric-fan ownership from 55% to 97%. If cooling protects the body, the biomarker response to a hot day should have weakened as protection spread. We test this in two ways: across rounds, by letting the slope differ by round, and within rounds, by letting it vary with access to cooling.

*Across rounds.* Estimating equation (1) with round-specific slopes (Online Appendix Table A.1, Panel A) gives glucose slopes of 1.470, 0.741 and 0.988 mg/dL per degree in 2011, 2017–18 and 2022, and equality across the three rounds is not rejected ($F = 2.12$, $p = 0.129$); nor is it for systolic pressure ($p = 0.733$), measured diabetes ($p = 0.503$) or measured hypertension ($p = 0.649$). Because BDHS 2011 measured only adults aged 35 and over, and mean age is 51.4 years in 2011 against 39.8 and 40.8 in the later rounds, Panel B holds the age range fixed at 35 and over. The glucose slopes there are 1.452, 0.829 and 1.092 ($p = 0.348$), and those for measured diabetes 0.650, 0.509 and 0.852 ($p = 0.763$). For systolic pressure in this age range equality

## Table 8: Inference

| | $\hat{\beta}$ | Standard error | | | | Wild bootstrap $p$ | Leave-one-district-out range |
|---|---|---|---|---|---|---|---|
| | | District (64) | District × date | Conley (30 km) | Conley (50 km) | | |
| Fasting glucose | 1.028 | (0.146) | (0.144) | (0.147) | (0.145) | 0.001 | [0.944, 1.065] |
| Measured diabetes (pp) | 0.626 | (0.143) | (0.129) | (0.143) | (0.148) | 0.001 | [0.542, 0.664] |
| Systolic BP | -0.562 | (0.077) | (0.094) | (0.076) | (0.079) | 0.001 | [-0.580, -0.536] |
| Measured hypertension (pp) | -0.839 | (0.145) | (0.161) | (0.142) | (0.125) | 0.001 | [-0.893, -0.796] |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. Each row is one baseline estimate: the coefficient on the exposure-day maximum (column 1) for fasting glucose (mg/dL), measured diabetes (fasting plasma glucose ≥126 mg/dL, percentage points), systolic blood pressure (mmHg) and measured hypertension (≥140/90 mmHg, percentage points), per °C. The exposure is the daily maximum on the day of the household interview for blood pressure, when the reading was taken, and on the day before the early-morning test for fasting glucose. Controls are age, age squared, sex, urban residence, education, wealth quintile and CHIRPS rainfall on the exposure day and the day before. Fixed effects are district × round and district × calendar month (the month of the test, for glucose), plus hour-of-test fixed effects for glucose; tests taken during Ramadan are excluded from the glucose sample. Columns 2–5 report alternative standard errors for that estimate: clustering on district (64 clusters, the baseline); two-way clustering on district and on the calendar date of the exposure (Cameron et al., 2011), which allows errors to be correlated across districts measured on the same day; and Conley (Conley, 1999) spatial HAC standard errors with a uniform kernel out to 30 and 50 km, which allow unrestricted correlation within a district and add correlation between districts whose centroids lie within the band. Column 6 is the $p$-value from a restricted wild cluster bootstrap-$t$ on the district clusters, with Rademacher weights, the null imposed and 999 replications (Cameron et al., 2008); under the $(r+1)/(B+1)$ convention the smallest attainable value is 0.001. Column 7 gives the range of the point estimate across the 64 re-estimations that each leave out one district. The average district has 1.41 district centroids within 30 km, including its own, and 3.31 within 50 km, so at these bandwidths the Conley estimator stays close to district clustering by construction; districts average roughly 2,300 km$^2$, so a much wider band would span most of a division.

is rejected only at the 10% level ($p = 0.088$), and the largest slope is the 2022 one (-0.962 mmHg, against -0.537 in 2011), the opposite of what adaptation would produce. There is no monotone decline in the slopes, and no difference between rounds is significant at the 5% level. Online Appendix Figure A.1 plots the round-specific slopes with and without the age restriction.

The comparison across rounds is not a clean test of adaptation, for two reasons. First, the glucose measurement changed between rounds. BDHS 2011 used a HemoCue Glucose 201+ and converted its whole-blood readings to plasma equivalents afterwards by multiplying by 1.11 (National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013); the 2017–18 analyser converted its readings automatically; and for 2022 we follow the recode's plasma variable, which applies the factor of 1.11 to the recorded reading, while the reports name the HemoCue 201 RT in their methods and the Glucose 201 DM in their results tables (National Institute of Population Research and Training (NIPORT) and ICF, 2020, 2024). Any difference in slopes between rounds would therefore mix the response with the instrument. Second, the measured-diabetes slope is the slope of a classification, and how far a uniform shift in readings moves it depends on how much of each round's distribution lies near 126 mg/dL (Section 5.4).

*Within rounds.* The second test asks whether, within a round, respondents with more access to cooling respond less (Online Appendix Table A.2). In-

teracting the exposure-day maximum with district-level electricity or fan coverage, computed leaving out the respondent's own household, yields coefficients that are not significant. Household-level interactions are significant for glucose before, but not after, temperature is also interacted with wealth (glucose × fan: -0.252, $p = 0.028$, becoming -0.141, $p = 0.269$), so the raw interaction cannot be separated from wealth.

We therefore find no evidence that the acute physiological response to heat weakened over a decade in which cooling technology went from about half of households to nearly all of them. We read this as an absence of evidence of adaptation rather than evidence of its absence, for three reasons. Three rounds give the tests little power to detect a gradual attenuation, and the round-specific slopes are imprecise. Owning a fan or an electricity connection measures access to cooling, not its use on the morning of measurement. And a fan is not protective in every condition: fans reduce heat strain across most hot weather but can add to it in very hot, dry air (Jay et al., 2015; Morris et al., 2019, 2021). Air conditioning, which cools in any condition and whose adoption follows income (Davis and Gertler, 2015), was owned by almost no household in these rounds.

The result is worth setting beside the mortality record. In the same country, with the same surveys and the same temperature data, Haque and Chopra (2026) find that the neonatal mortality response to days of 30–40°C flattened across birth cohorts from the 1980s onward, in step with rural electrification,

while no decline is detected above 40°C. The two results concern different populations, newborns and adults, and different margins, so they cannot be combined into one estimate. They are consistent, however, with a reading that we cannot test here: over the range of heat our measurement days cover, the body's acute response stayed the same while what follows from it changed.

## 9. How Much of the Recorded Burden Is Thermal?

Section 5.4 shows that heat moves the readings of the whole measured population, and with them the diagnostic classification. This section asks what that implies for the survey record itself: how much of the change in measured prevalence between BDHS rounds reflects the temperature at which each round was fielded.

The three rounds are the basis of the published trends in hypertension and diabetes in Bangladesh (Sathi et al., 2022; Rana et al., 2024), and they were fielded in different parts of the year. Among adults aged 35 and over in the glucose sample, the age range all three rounds cover, mean exposure-day maxima were 29.4, 27.4 and 31.4°C, so a comparison of two rounds embeds a temperature difference of two to four degrees (the blood-pressure sample is dated by the interview rather than the test visit, so its means differ slightly). We re-estimate the slopes on adults aged 35 and over, multiply each by the difference in mean exposure-day maxima between two rounds to obtain the

thermal component of the recorded change, and subtract it to obtain the change at a common fieldwork temperature. Table 9 reports the result.

We find that the thermal components are of the same order as the recorded changes. Recorded diabetes prevalence among adults aged 35 and over rose 2.15 percentage points between 2011 and 2017–18. The later round was fielded 2.0°C cooler, so the slope attributes -1.37 points to temperature, and the rise at a common fieldwork temperature is 3.52 points (95% interval from the uncertainty in the slope, 2.88 to 4.16), larger than the recorded change. Between 2017–18 and 2022 recorded diabetes rose 7.80 points. The 2022 round was fielded 4.0°C warmer, which accounts for 2.69 of them, about a third, and the rise at a common temperature is 5.11 points (3.85 to 6.37). Most of the recorded rise in that window survives the adjustment, but about a third of it reflects the weather in which the two rounds were fielded. For measured hypertension over the same window, -3.94 of the recorded change of -10.36 points is attributed to temperature, leaving -6.41 (-7.88 to -4.95).

Panels C and D repeat the exercise with the wet-bulb maximum, the exposure the response tracks most closely (Section 5.3), and in the consecutive-round windows the thermal component grows: between 2017–18 and 2022, where it is 4.51 points rather than 2.69, the recorded rise of 7.80 points in measured diabetes becomes 3.29 at a common wet-bulb temperature, against 5.11 on the dry bulb.

**Table 9: How much of the change in recorded prevalence between rounds is thermal?**

| | Round | | | Change between rounds (pp) | | |
|---|---|---|---|---|---|---|
| | 2011 | 2017–18 | 2022 | Observed | Thermal | Adjusted |
| *Panel A. Measured diabetes, fasting glucose ≥126 mg/dL* | | | | | | |
| Mean measurement-day $T_{max}$ (°C) | 29.4 | 27.4 | 31.4 | | | |
| Recorded prevalence (%) | 9.3 | 11.5 | 19.3 | | | |
| 2011 → 2017–18 | | -2.0 | | +2.15 | -1.37 [-2.01, -0.73] | +3.52 [+2.88, +4.16] |
| 2017–18 → 2022 | | +4.0 | | +7.80 | +2.69 [+1.43, +3.95] | +5.11 [+3.85, +6.37] |
| 2011 → 2022 | | +2.0 | | +9.95 | +1.32 [+0.70, +1.94] | +8.63 [+8.01, +9.24] |
| *Panel B. Measured hypertension, ≥140/90* | | | | | | |
| Mean measurement-day $T_{max}$ (°C) | 30.0 | 27.4 | 31.4 | | | |
| Recorded prevalence (%) | 20.6 | 34.0 | 23.6 | | | |
| 2011 → 2017–18 | | -2.7 | | +13.42 | +2.64 [+1.66, +3.61] | +10.79 [+9.81, +11.77] |
| 2017–18 → 2022 | | +4.0 | | -10.36 | -3.94 [-5.41, -2.48] | -6.41 [-7.88, -4.95] |
| 2011 → 2022 | | +1.3 | | +3.07 | -1.31 [-1.79, -0.82] | +4.37 [+3.89, +4.86] |
| *Panel C. Measured diabetes, wet-bulb exposure* | | | | | | |
| Mean measurement-day wet-bulb (°C) | 23.9 | 20.4 | 25.6 | | | |
| Recorded prevalence (%) | 9.3 | 11.5 | 19.3 | | | |
| 2011 → 2017–18 | | -3.5 | | +2.15 | -3.06 [-4.35, -1.77] | +5.21 [+3.92, +6.50] |
| 2017–18 → 2022 | | +5.2 | | +7.80 | +4.51 [+2.61, +6.42] | +3.29 [+1.38, +5.19] |
| 2011 → 2022 | | +1.7 | | +9.95 | +1.45 [+0.84, +2.07] | +8.49 [+7.88, +9.11] |
| *Panel D. Measured hypertension, wet-bulb exposure* | | | | | | |
| Mean measurement-day wet-bulb (°C) | 24.9 | 20.4 | 25.6 | | | |
| Recorded prevalence (%) | 20.6 | 34.0 | 23.6 | | | |
| 2011 → 2017–18 | | -4.5 | | +13.42 | +6.68 [+4.34, +9.02] | +6.74 [+4.40, +9.08] |
| 2017–18 → 2022 | | +5.2 | | -10.36 | -7.68 [-10.37, -4.99] | -2.68 [-5.37, +0.01] |
| 2011 → 2022 | | +0.7 | | +3.07 | -1.00 [-1.34, -0.65] | +4.06 [+3.71, +4.41] |

*Notes:* Data: BDHS 2011, 2017–18 and 2022 biomarker samples, merged with ERA5 daily-maximum temperature inverse-distance-weighted to the district centroid over Bangladesh local days. The sample is adults aged 35 and over, the age range common to all three rounds, so recorded prevalence differs from the whole-sample figures in the summary statistics; prevalence is the unweighted share of the sample classified as having the condition. Measured diabetes is a fasting plasma glucose of at least 126 mg/dL and measured hypertension a reading of at least 140/90 mmHg. Panels A and B use the dry-bulb daily maximum and Panels C and D the wet-bulb daily maximum; the first row of each panel is the mean exposure-day maximum in each round. In the three round-pair rows, the cell spanning the round columns is the change in mean exposure-day maximum between the two rounds (in °C, not a prevalence). "Observed" is the change in recorded prevalence, in percentage points. "Thermal" is the slope of the classification on the exposure-day maximum, estimated on this sample with the baseline specification, multiplied by the change in mean temperature. "Adjusted" is Observed minus Thermal, the change that would have been recorded had both rounds been fielded at the same mean temperature. The slopes are 0.673 (SE 0.160) percentage points per °C for measured diabetes and -0.985 (SE 0.186) for measured hypertension on the dry bulb, and 0.869 (SE 0.187) and -1.474 (SE 0.263) on the wet bulb. The exposure is the daily maximum on the day of the household interview for blood pressure, when the reading was taken, and on the day before the early-morning

*How to read the adjustments.* We read these adjustments as orders of magnitude rather than corrected trends, for four reasons. First, the intervals reflect only the uncertainty in the slope, not sampling error in each round's prevalence. Second, the slope is identified from day-to-day variation within a district and month, while the differences between rounds are seasonal; a population measured in the cool season is acclimatised to it, and if the acute response to a hotter day exceeds the response to a hotter season, applying the acute slope overstates the thermal component. Third, the slope is pooled across rounds; the round-specific slopes for measured diabetes among adults aged 35 and over (0.650, 0.509 and 0.852) do not differ significantly, but they are imprecise (Section 8). Fourth, recorded prevalence here is the unweighted mean of the measured sample, not the survey-weighted national estimate. Other differences between rounds are larger still: recorded hypertension among adults aged 35 and over moves by more than ten points in each of the two consecutive-round windows, far more than temperature can account for. The exercise therefore does not say what happened to Bangladesh's non-communicable disease burden. It says that a comparison of rounds fielded at different times of year carries a thermal component of the same order as the changes the surveys are used to detect.

### *9.1. Implications for Measurement*

Three uses of the survey record are affected. First, comparisons *across countries* in the global burden estimates (NCD Risk Factor Collaboration (NCD-

RisC), 2021, 2024) pool surveys with different fieldwork calendars, and the calendar is not usually reported with the prevalence estimate; DHS fieldwork is scheduled with the seasons in mind, though with only a slight tilt towards the dry season across countries (Wright et al., 2012). Second, comparisons *over time* within a country compare rounds that were not fielded in the same months, which shifts measured child anthropometry enough to change trends between Indian DHS rounds (Madan et al., 2020), and which Table 9 illustrates for Bangladesh. Third, projections of the burden under warming will move mechanically: warming raises measured diabetes and lowers measured hypertension before anything happens to the population's underlying disease, so the errors do not cancel in an aggregate index. Measured anaemia moves in the direction dilution implies, but we observe it in only one round and one population and do not decompose it here.

Earlier work has proposed the same correction in other settings, standardising blood pressure to a fixed ambient temperature in Ghana and China and estimating temperature-adjusted diabetes prevalence in Guangdong (Kunutsor and Powles, 2010; Su et al., 2014; Luo et al., 2021); our estimates show that the correction matters for the trends a national DHS series reports.

The remedy costs little. Surveys already record the date and hour of every biomarker measurement, and the reanalysis needed to attach temperature and humidity to that time stamp is public. Reporting the distribution of fieldwork conditions alongside the headline prevalence—on the evidence of

Section 5.3, the wet bulb rather than the dry bulb—or publishing an estimate adjusted to a common thermal reference would make comparisons across surveys easier to interpret at no cost in data collection.

## 10. Conclusion

Ambient heat moves the biomarkers from which household surveys diagnose chronic disease. In 34,879 adults with a blood-pressure reading and 33,179 with a fasting glucose result from three BDHS rounds, a one-degree rise in the daily maximum lowers systolic pressure by 0.562 mmHg on the day it is read and a rise the day before the test raises fasting glucose by 1.028 mg/dL, both significant at the 1% level under district, two-way and Conley standard errors and the wild cluster bootstrap. The fall in pressure is 2.4 times larger among adults aged 50 and over, while the glucose response varies much less with age (interaction 0.271, $p = 0.060$). Haemoglobin, collected in 2011 for women aged 15–49 and children aged 6–59 months, falls by 0.018 g/dL per degree of dry-bulb maximum ($p = 0.310$) and by 0.037 per degree of wet-bulb maximum ($p = 0.069$); the fall is imprecise, but its sign is the one dilution predicts and the opposite of the one dehydration predicts. With the wet-bulb and dry-bulb maxima entered together, the wet bulb keeps 1.208 mg/dL on glucose and -0.778 mmHg on systolic pressure, and the dry-bulb term is not significant at the 5% level for either diagnostic classification, so the response is to humid heat rather than to heat as such.

Because chronic disease is diagnosed from a single reading, the response moves the diagnosis. Measured diabetes rises 0.626 percentage points per degree (SE 0.143), close to what a uniform shift of every glucose reading would produce, and measured hypertension falls 0.839 (SE 0.145), while respondents' reports of a past diagnosis do not respond. Applied to the survey record among adults aged 35 and over, the 2022 round was fielded 4.0°C warmer than 2017–18, which accounts for 2.69 of the 7.80-point rise in measured diabetes between the two rounds, about a third; the cooler 2017–18 fieldwork hid part of the rise since 2011 (a thermal component of -1.37 points against a recorded 2.15). These adjustments hold the slope fixed and apply a day-to-day response to seasonal differences, so we read them as orders of magnitude rather than corrected trends. They are nonetheless of the same order as the changes the surveys are used to detect.

The design rests on day-to-day weather variation within a district and survey round, net of the district's own seasonal pattern, and it is tested for balance, selection, fieldwork intensity, measurers and cuff sizes, placebo outcomes recorded on the same visit with a tape, a board and a scale, rainfall controls, and standard errors clustered on district and date together. Eight results are not clean, and we report them (Sections 6 and 7, Online Appendices A.2 and A.4). First, comparisons within a single survey-cluster visit, which retain a tenth to a seventh of the temperature variation the design uses, do not reproduce the baseline. Second, 3 of 12 respondent characteristics are significantly related to the day's temperature, by at most 0.0212 standard

deviations per degree. Third, in 2017–18 women interviewed on hotter days were less likely to be employed, an imbalance across clusters that leaves the estimates unchanged when controlled for. Fourth, consent to and completion of the blood-pressure measurement fall slightly on hotter days, by far too little to generate the estimates through selection. Fifth, the systolic response also loads on the temperature of the following fortnight. Sixth, the seven-day glucose lead does not shrink in the horse race with the exposure day. Seventh, the haemoglobin estimates are imprecise: significant only at the 10% level on the wet bulb, and not after the multiple-testing adjustment. Eighth, the temperature at the hour of the glucose test, which is also the analyser's temperature, carries much of the glucose response. The gradient is as strong among tests taken within each round's strictest stated operating range as in the whole sample, and the previous afternoon's heat keeps an effect that no analyser can explain, but we cannot exclude temperature sensitivity within the stated range without calibration records.

We find evidence on the physiological response but not on what follows from it. The design identifies a same-day state; we observe no mortality, hospitalisation or clinical diagnosis, and the sample contains no day above 36.5°C. Two questions therefore remain open, and both require following people rather than measuring them once: what a season of hot days does to metabolic control, as opposed to what a single hot morning does to a reading; and whether the larger fall in pressure among older adults translates into the cardiovascular events that heat waves are known to produce (Liu

et al., 2022; Ebi et al., 2021). Two features of the data would sharpen the estimates themselves. BDHS 2011 records the clock time at which each respondent last ate, so realised fasting duration can be conditioned on in that round, although 2017–18 records only whether the fast reached eight hours and 2022 records neither. BDHS 2022 records the clock time of each of the three blood-pressure readings, which supports a within-person design against the diurnal temperature cycle.

For measurement the implication is direct. A household survey that diagnoses chronic disease from one reading also records the weather on the day of the visit, and two rounds fielded in different seasons, or in years of different warmth, differ in their recorded burden for that reason alone. Recording the date and hour of each measurement, as the BDHS does, makes the thermal component estimable; reporting prevalence at a common fieldwork temperature, alongside the raw figure, would make it visible to the users of the series.

## Declaration of competing interest

The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## CRediT authorship contribution statement

**Mohammad Islamul Haque:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review and editing.

## Ethics statement

The study analyses de-identified secondary data from the Bangladesh Demographic and Health Survey. The surveys were conducted under protocols approved by the ICF Institutional Review Board and the Bangladesh Medical Research Council, and informed consent was obtained from every respondent before the interview and before each biomarker measurement. The data are distributed by the DHS Program to registered users without personal identifiers. No new data were collected from human participants, and the analysis required no further ethical approval.

## Data availability

The Bangladesh Demographic and Health Survey microdata (2011, 2017–18 and 2022) are available from the DHS Program on registration (https://dhsprogram.com); its terms of use do not permit redistribution. ERA5 reanalysis data are available from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu) and CHIRPS daily rainfall from the Climate Hazards Center (https://www.chc.ucsb.edu/data/chirps). The replication code, which rebuilds every table, figure and number in the paper from these sources, and the constructed district-level exposure data will be deposited in a public repository on acceptance and cited and linked in the published article.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used Claude (Anthropic) to assist with writing and debugging analysis code, typesetting tables and figures, and editing prose for clarity. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

## References


A&D Medical (LifeSource), 2009. Trilingual Instruction Guide: Model UA-767 Plus. URL: https://medical.andonline.com/wp-content/uploads/2021/08/UA-767-Plus-Instruction-Manual_0.pdf. operating environment 10–40 °C, less than 85% relative humidity.

Adhvaryu, A., Kala, N., Nyshadham, A., 2020. The light and the heat: Productivity co-benefits of energy-saving technology. Review of Economics and Statistics 102, 779–792. doi:10.1162/rest_a_00886.

Aguree, S., Gernand, A.D., 2019. Plasma volume expansion across healthy pregnancy: a systematic review and meta-analysis of longitudinal studies. BMC Pregnancy and Childbirth 19, 508. doi:10.1186/s12884-019-2619-6.

Akter, S., Rahman, M.M., Abe, S.K., Sultana, P., 2014. Prevalence of diabetes and prediabetes and their risk factors among Bangladeshi adults: a nationwide survey. Bulletin of the World Health Organization 92, 204–213A. doi:10.2471/BLT.13.128371.

Alduchov, O.A., Eskridge, R.E., 1996. Improved Magnus form approximation of saturation vapor pressure. Journal of Applied Meteorology 35, 601–609. doi:10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Alpérovitch, A., Lacombe, J.M., Hanon, O., Dartigues, J.F., Ritchie, K., Ducimetière, P., Tzourio, C., 2009. Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: The Three-City study. Archives of Internal Medicine 169, 75–80. doi:10.1001/archinternmed.2008.512.

Armstrong, B., Sera, F., Vicedo-Cabrera, A.M., Abrutzky, R., Åström, D.O., Bell, M.L., Chen, B.Y., de Sousa Zanotti Stagliorio Coelho, M., Correa, P.M., Dang, T.N., et al., 2019. The role of humidity in associations of high temperature with mortality: A multicountry, multicity study. Environmental Health Perspectives 127, 097007. doi:10.1289/EHP5430.

Auffhammer, M., Hsiang, S.M., Schlenker, W., Sobel, A., 2013. Using weather data and climate model output in economic analyses of climate change. Review of Environmental Economics and Policy 7, 181–198. doi:10.1093/reep/ret016.

Baldwin, J.W., Benmarhnia, T., Ebi, K.L., Jay, O., Lutsko, N.J., Vanos, J.K., 2023. Humidity's role in heat-related health outcomes: A heated debate. Environmental Health Perspectives 131, 055001. doi:10.1289/EHP11807.

Banerjee, R., Maharaj, R., 2020. Heat, infant mortality, and adaptation: Evidence from India. Journal of Development Economics 143, 102378. doi:10.1016/j.jdeveco.2019.102378.

Barreca, A., Clay, K., Deschênes, O., Greenstone, M., Shapiro, J.S., 2016. Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century. Journal of Political Economy 124, 105–159. doi:10.1086/684582.

Barreca, A.I., 2012. Climate change, humidity, and mortality in the United States. Journal of Environmental Economics and Management 63, 19–34. doi:10.1016/j.jeem.2011.07.004.

Bass, D.E., Henschel, A., 1956. Responses of body fluid compartments to heat and cold. Physiological Reviews 36, 128–144. doi:10.1152/physrev.1956.36.1.128.

Blauw, L.L., Aziz, N.A., Tannemaat, M.R., Blauw, C.A., de Craen, A.J., Pijl, H., Rensen, P.C.N., 2017. Diabetes incidence and glucose intolerance prevalence increase with higher outdoor temperature. BMJ Open Diabetes Research & Care 5, e000317. doi:10.1136/bmjdrc-2016-000317.

Bolton, D., 1980. The computation of equivalent potential temperature. Monthly Weather Review 108, 1046–1053. doi:10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Bound, J., Brown, C., Mathiowetz, N., 2001. Measurement error in survey data, in: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics. Elsevier, Amsterdam. volume 5, pp. 3705–3843. doi:10.1016/S1573-4412(01)05012-7.

Brennan, P.J., Greenberg, G., Miall, W.E., Thompson, S.G., 1982. Seasonal variation in arterial blood pressure. British Medical Journal (Clinical Research Edition) 285, 919–923. doi:10.1136/bmj.285.6346.919.

Burgess, R., Deschênes, O., Donaldson, D., Greenstone, M., 2017. Weather, climate change and death in India. Unpublished manuscript.

Burkart, K., Breitner, S., Schneider, A., Khan, M.M.H., Krämer, A., Endlicher, W., 2014. An analysis of heat effects in different subpopulations of Bangladesh. International Journal of Biometeorology 58, 227–237. doi:10.1007/s00484-013-0668-5.

Burkart, K., Khan, M.H., Krämer, A., Breitner, S., Schneider, A., Endlicher, W.R., 2011. Seasonal variations of all-cause and cause-specific mortality by age, gender, and socioeconomic condition in urban and rural areas of Bangladesh. International Journal for Equity in Health 10, 32. doi:10.1186/1475-9276-10-32.

Buzan, J.R., Oleson, K., Huber, M., 2015. Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5. Geoscientific Model Development 8, 151–170. doi:10.5194/gmd-8-151-2015.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Review of Economics and Statistics 90, 414–427. doi:10.1162/rest.90.3.414.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2011. Robust inference with multiway clustering. Journal of Business & Economic Statistics 29, 238–249. doi:10.1198/jbes.2010.07136.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. Journal of Human Resources 50, 317–372. doi:10.3368/jhr.50.2.317.

Carleton, T., Jina, A., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Kopp, R.E., McCusker, K.E., Nath, I., Rising, J., Rode, A., Seo, H.K., Viaene, A., Yuan, J., Zhang, A.T., 2022. Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. Quarterly Journal of Economics 137, 2037–2105. doi:10.1093/qje/qjac020.

Charkoudian, N., 2003. Skin blood flow in adult human thermoregulation: How it works, when it does not, and why. Mayo Clinic Proceedings 78, 603–612. doi:10.4065/78.5.603.

Conley, T.G., 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1–45. doi:10.1016/S0304-4076(98)00084-0.

Cramer, M.N., Gagnon, D., Laitano, O., Crandall, C.G., 2022. Human temperature regulation under heat stress in health, disease, and injury. Physiological Reviews 102, 1907–1989. doi:10.1152/physrev.00047.2021.

Crandall, C.G., Wilson, T.E., 2015. Human cardiovascular responses to passive heat stress. Comprehensive Physiology 5, 17–43. doi:10.1002/cphy.c140015.

Davies-Jones, R., 2008. An efficient and accurate method for computing the wet-bulb temperature along pseudoadiabats. Monthly Weather Review 136, 2764–2785. doi:10.1175/2007MWR2224.1.

Davis, L.W., Gertler, P.J., 2015. Contribution of air conditioning adoption to future energy use under global warming. Proceedings of the National Academy of Sciences 112, 5962–5967. doi:10.1073/pnas.1423558112.

Dell, M., Jones, B.F., Olken, B.A., 2014. What do we learn from the weather? the new climate-economy literature. Journal of Economic Literature 52, 740–798. doi:10.1257/jel.52.3.740.

Deschênes, O., 2014. Temperature, human health, and adaptation: A review of the empirical literature. Energy Economics 46, 606–619. doi:10.1016/j.eneco.2013.10.013.

Deschênes, O., Greenstone, M., 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. American Economic Journal: Applied Economics 3, 152–185. doi:10.1257/app.3.4.152.

Ebi, K.L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R.S., Ma, W., Malik, A., et al., 2021. Hot weather and heat extremes: health risks. The Lancet 398, 698–708. doi:10.1016/S0140-6736(21)01208-3.

Faure, C., Charlot, K., Henri, S., Hardy-Dessources, M.D., Hue, O., Antoine-Jonville, S., 2016. Impaired glucose tolerance after brief heat exposure: a randomized crossover study in healthy young men. Clinical Science 130, 1017–1025. doi:10.1042/CS20150461.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The Climate Hazards Infrared Precipitation with Stations—a new environmental record for monitoring extremes. Scientific Data 2, 150066. doi:10.1038/sdata.2015.66.

Garg, T., Jagnani, M., Taraz, V., 2020. Temperature and human capital in India. Journal of the Association of Environmental and Resource Economists 7, 1113–1150. doi:10.1086/710066.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., et al., 2015. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. The Lancet 386, 369–375. doi:10.1016/S0140-6736(14)62114-0.

Geldsetzer, P., Manne-Goehler, J., Marcus, M.E., Ebert, C., Zhumadilov, Z., Wesseh, C.S., Tsabedze, L., Supiyev, A., Sturua, L., Bahendeka, S.K., et al., 2019. The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from $1 \cdot 1$ million adults. The Lancet 394, 652–662. doi:10.1016/S0140-6736(19)30955-9.

Geruso, M., Spears, D., 2018. Heat, Humidity, and Infant Mortality in the Developing World. NBER Working Paper 24870. National Bureau of Economic Research. doi:10.3386/w24870.

Ginsberg, B.H., 2009. Factors affecting blood glucose monitoring: Sources of errors in measurement. Journal of Diabetes Science and Technology 3, 903–913. doi:10.1177/193229680900300438.

Graff Zivin, J., Hsiang, S.M., Neidell, M., 2018. Temperature and human capital in the short and long run. Journal of the Association of Environmental and Resource Economists 5, 77–105. doi:10.1086/694177.

Graff Zivin, J., Neidell, M., 2014. Temperature and the allocation of time: Implications for climate change. Journal of Labor Economics 32, 1–26. doi:10.1086/671766.

Graff Zivin, J., Song, Y., Tang, Q., Zhang, P., 2020. Temperature and high-stakes cognitive performance: Evidence from the national college entrance

examination in China. Journal of Environmental Economics and Management 104, 102365. doi:10.1016/j.jeem.2020.102365.

Haider, M.M., Kamal, N., Bashar, M.I., Rahman, M.M., Khan, S.H., Alam, N., 2023. Religious disparities in health in Bangladesh–the case of hypertension and diabetes: evidence from two nationally representative cross-sectional surveys. BMJ Open 13, e067960. doi:10.1136/bmjopen-2022-067960.

Haller, M.J., Shuster, J.J., Schatz, D., Melker, R.J., 2007. Adverse impact of temperature and humidity on blood glucose monitoring reliability: A pilot study. Diabetes Technology & Therapeutics 9, 1–9. doi:10.1089/dia.2006.0051.

Halonen, J.I., Zanobetti, A., Sparrow, D., Vokonas, P.S., Schwartz, J., 2011. Relationship between outdoor temperature and blood pressure. Occupational and Environmental Medicine 68, 296–301. doi:10.1136/oem.2010.056507.

Haque, M.I., Chopra, A., 2026. Adapting to high heat, failing at the extreme: Newborn heat mortality over four decades in Bangladesh. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=7346240, doi:10.2139/ssrn.7346240. preprint, SSRN Working Paper 7346240.

Hargreaves, M., Angus, D., Howlett, K., Conus, N.M., Febbraio, M., 1996. Effect of heat stress on glucose kinetics during exercise. Journal of Applied Physiology 81, 1594–1597. doi:10.1152/jappl.1996.81.4.1594.

Hausman, J.A., Abrevaya, J., Scott-Morton, F.M., 1998. Misclassification of the dependent variable in a discrete-response setting. Journal of Econometrics 87, 239–269. doi:10.1016/S0304-4076(98)00015-3.

HemoCue AB, 2024a. HemoCue Glucose 201 DM System: Product Sheet. URL: https://hemocue.com/wp-content/uploads/2024/05/GPM323INT-Glucose-201-DM-Product-Sheet-240712.pdf. gPM323INT; operating temperature 18–30 °C.

HemoCue AB, 2024b. HemoCue Glucose 201 RT System: Product Sheet. URL: https://hemocue.com/wp-content/uploads/2024/05/GPM319INT-Glucose-201-RT-Product-Sheet-240712.pdf. gPM319INT; operating temperature 15–27 °C.

HemoCue AB, 2024c. HemoCue Glucose 201+ System: Product Sheet. URL: https://hemocue.com/wp-content/uploads/2024/05/GPM322INT-Glucose-201-Product-Sheet-240712.pdf. gPM322INT; operating temperature 15–30 °C.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al., 2023. ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi:10.24381/cds.adbb2d47.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146, 1999–2049. doi:10.1002/qj.3803.

Heutel, G., Miller, N.H., Molitor, D., 2021. Adaptation and the mortality effects of temperature across US climate regions. Review of Economics and Statistics 103, 740–753. doi:10.1162/rest_a_00936.

Hossain, M.B., Khan, M.N., Oldroyd, J.C., Rana, J., Magliago, D.J., Chowdhury, E.K., Karim, M.N., Islam, R.M., 2022. Prevalence of, and risk factors for, diabetes and prediabetes in Bangladesh: Evidence from the national survey using a multilevel Poisson regression model with a robust variance. PLOS Global Public Health 2, e0000461. doi:10.1371/journal.pgph.0000461.

Hsiang, S., 2016. Climate econometrics. Annual Review of Resource Economics 8, 43–75. doi:10.1146/annurev-resource-100815-095343.

ICF International, 2012a. Biomarker Field Manual: Demographic and Health Surveys Methodology. MEASURE DHS Manual DHSM7. ICF International. Calverton, Maryland. URL: https://dhsprogram.com/pubs/pdf/DHSM7/DHS6_Biomarker_Manual_9Jan2012.pdf.

ICF International, 2012b. Survey Organization Manual for Demographic and Health Surveys. MEASURE DHS Manual DHSM10. ICF International. Calverton, Maryland. URL: https://dhsprogram.com/pubs/pdf/DHSM10/DHS6_Survey_Org_Manual_7Dec2012_DHSM10.pdf.

Jay, O., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R.S., Ma, W., Malik, A., et al., 2021. Reducing the health effects of hot weather and heat extremes: from personal cooling strategies to green cities. The Lancet 398, 709–724. doi:10.1016/S0140-6736(21)01209-5.

Jay, O., Cramer, M.N., Ravanelli, N.M., Hodder, S.G., 2015. Should electric fans be used during a heat wave? Applied Ergonomics 46, 137–143. doi:10.1016/j.apergo.2014.07.013.

Johnston, D.W., Propper, C., Shields, M.A., 2009. Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. Journal of Health Economics 28, 540–552. doi:10.1016/j.jhealeco.2009.02.010.

Kenney, W.L., Craighead, D.H., Alexander, L.M., 2014. Heat waves, aging, and human cardiovascular health. Medicine & Science in Sports & Exercise 46, 1891–1899. doi:10.1249/MSS.0000000000000325.

Kenny, G.P., Sigal, R.J., McGinn, R., 2016. Body temperature regulation in diabetes. Temperature 3, 119–145. doi:10.1080/23328940.2015.1131506.

Kristal-Boneh, E., Froom, P., Harari, G., Shapiro, Y., Green, M.S., 1993. Seasonal changes in red blood cell parameters. British Journal of Haematology 85, 603–607. doi:10.1111/j.1365-2141.1993.tb03354.x.

Kunutsor, S.K., Powles, J.W., 2010. The effect of ambient temperature on blood pressure in a rural West African adult population: a cross-sectional study. Cardiovascular Journal of Africa 21, 17–20.

Lewington, S., Li, L., Sherliker, P., Guo, Y., Millwood, I., Bian, Z., Whitlock, G., Yang, L., Collins, R., Chen, J., et al., 2012. Seasonal variation in blood pressure and its relationship with outdoor temperature in 10 diverse regions of China: the China Kadoorie Biobank. Journal of Hypertension 30, 1383–1391. doi:10.1097/HJH.0b013e32835465b5.

Li, S., Zhou, Y., Williams, G., Jaakkola, J.J.K., Ou, C., Chen, S., Yao, T., Qin, T., Wu, S., Guo, Y., 2016. Seasonality and temperature effects on fasting plasma glucose: a population-based longitudinal study in China. Diabetes & Metabolism 42, 267–275. doi:10.1016/j.diabet.2016.01.002.

Liu, J., Varghese, B.M., Hansen, A., Zhang, Y., Driscoll, T., Morgan, G., Dear, K., Gourley, M., Capon, A., Bi, P., 2022. Heat exposure and cardiovascular health outcomes: a systematic review and meta-analysis. The Lancet Planetary Health 6, e484–e495. doi:10.1016/S2542-5196(22)00117-6.

LoPalo, M., 2023. Temperature, worker productivity, and adaptation: Evidence from survey data production. American Economic Journal: Applied Economics 15, 192–229. doi:10.1257/app.20200547.

Lu, Y.C., Romps, D.M., 2023. Predicting fatal heat and humidity using the heat index model. Journal of Applied Physiology 134, 649–656. doi:10.1152/japplphysiol.00417.2022.

Luo, J., He, G., Xu, Y., Chen, Z., Xu, X., Peng, J., Chen, S., Hu, J., Ji, G., Liu, T., Zeng, W., Li, X., Xiao, J., Guo, L., He, Q., Ma, W., 2021. The relationship between ambient temperature and fasting plasma glucose, temperature-adjusted type 2 diabetes prevalence and control rate: a series of cross-sectional studies in Guangdong Province, China. BMC Public Health 21, 1534. doi:10.1186/s12889-021-11563-5.

MacKinnon, J.G., Nielsen, M.Ø., Webb, M.D., 2023. Cluster-robust inference: A guide to empirical practice. Journal of Econometrics 232, 272–299. doi:10.1016/j.jeconom.2022.04.001.

MacKinnon, J.G., Webb, M.D., 2017. Wild bootstrap inference for wildly different cluster sizes. Journal of Applied Econometrics 32, 233–254. doi:10.1002/jae.2508.

Madan, E.M., Frongillo, E.A., Unisa, S., Dwivedi, L., Johnston, R., Daniel, A., Agrawal, P.K., Deb, S., Khera, A., Menon, P., et al., 2020. Effect of differences in month and location of measurement in estimating prevalence and trend of wasting and stunting in India in 2005–2006 and 2015–2016. Current Developments in Nutrition 4, nzaa092. doi:10.1093/cdn/nzaa092.

Maley, M.J., Hunt, A.P., Stewart, I.B., Faulkner, S.H., Minett, G.M., 2019. Passive heating and glycaemic control in non-diabetic and diabetic individuals: A systematic review and meta-analysis. PLOS ONE 14, e0214223. doi:10.1371/journal.pone.0214223.

Manne-Goehler, J., Geldsetzer, P., Agoudavi, K., Andall-Brereton, G., Aryal, K.K., Bicaba, B.W., Bovet, P., Brian, G., Dorobantu, M., Gathecha, G., et al., 2019. Health system performance for people with diabetes in 28 low- and middle-income countries: A cross-sectional study of nationally representative surveys. PLOS Medicine 16, e1002751. doi:10.1371/journal.pmed.1002751.

Modesti, P.A., Morabito, M., Bertolozzi, I., Massetti, L., Panci, G., Lumachi, C., Giglio, A., Bilo, G., Caldara, G., Lonati, L., et al., 2006. Weather-related changes in 24-hour blood pressure profile: Effects of age and implications for hypertension management. Hypertension 47, 155–161. doi:10.1161/01.HYP.0000199192.17126.d4.

Modesti, P.A., Rapi, S., Bamoshmoosh, M., Baldereschi, M., Massetti, L., Padeletti, L., Gensini, G.F., Zhao, D., Al-Hidabi, D., Al Goshae, H., 2012. Impact of one or two visits strategy on hypertension burden estimation in HYDY, a population-based cross-sectional study: implications for healthcare resource allocation decision making. BMJ Open 2, e001062. doi:10.1136/bmjopen-2012-001062.

Morris, N.B., Chaseling, G.K., English, T., Gruss, F., Maideen, M.F.B., Capon, A., Jay, O., 2021. Electric fan use for cooling during hot weather: a biophysical modelling study. The Lancet Planetary Health 5, e368–e377. doi:10.1016/S2542-5196(21)00136-4.

Morris, N.B., English, T., Hospers, L., Capon, A., Jay, O., 2019. The effects of electric fan use under differing resting heat index conditions: A clinical trial. Annals of Internal Medicine 171, 675–677. doi:10.7326/M19-0512.

Morris, S.S., Ruel, M.T., Cohen, R.J., Dewey, K.G., de la Brière, B., Hassan, M.N., 1999. Precision, accuracy, and reliability of hemoglobin assessment with use of capillary blood. The American Journal of Clinical Nutrition 69, 1243–1248. doi:10.1093/ajcn/69.6.1243.

Mukhopadhyay, A., Haque Mondol, M., Rahman, M., Unicomb, L., Khan, R., Mazumder, H., Nahian Ferdous, M., Pickering, E.V., Makris, K.C., Caban-Martinez, A.J., et al., 2025. The direct and urinary electrolyte-mediated effects of ambient temperature on population blood pressure: A causal mediation analysis. Environment International 195, 109208. doi:10.1016/j.envint.2024.109208.

Mullins, J.T., White, C., 2020. Can access to health care mitigate the effects of temperature on mortality? Journal of Public Economics 191, 104259. doi:10.1016/j.jpubeco.2020.104259.

Muntner, P., Shimbo, D., Carey, R.M., Charleston, J.B., Gaillard, T., Misra, S., Myers, M.G., Ogedegbe, G., Schwartz, J.E., Townsend, R.R., et al., 2019. Measurement of blood pressure in humans: A scientific statement from the American Heart Association. Hypertension 73, e35–e66. doi:10.1161/HYP.0000000000000087.

National Institute of Population Research and Training (NIPORT) and ICF, 2020. Bangladesh Demographic and Health Survey 2017-18. Technical Report. NIPORT and ICF. Dhaka, Bangladesh and Rockville, Maryland, USA. URL: https://dhsprogram.com/pubs/pdf/FR344/FR344.pdf. dHS Final Report FR344.

National Institute of Population Research and Training (NIPORT) and ICF, 2024. Bangladesh Demographic and Health Survey 2022: Final Report. Technical Report. NIPORT and ICF. Dhaka, Bangladesh and Rockville, Maryland, USA. URL: https://dhsprogram.com/pubs/pdf/FR386/FR386.pdf. dHS Final Report FR386.

National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, 2013. Bangladesh Demographic and Health Survey 2011. Technical Report. NIPORT, Mitra and Associates, and ICF International. Dhaka, Bangladesh and Calverton, Maryland, USA. URL: https://dhsprogram.com/pubs/pdf/FR265/FR265.pdf. dHS Final Report FR265.

NCD Risk Factor Collaboration (NCD-RisC), 2016. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with $4 \cdot 4$ million participants. The Lancet 387, 1513–1530. doi:10.1016/S0140-6736(16)00618-8.

NCD Risk Factor Collaboration (NCD-RisC), 2017. Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with $19 \cdot 1$ million participants. The Lancet 389, 37–55. doi:10.1016/S0140-6736(16)31919-5.

NCD Risk Factor Collaboration (NCD-RisC), 2021. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. The Lancet 398, 957–980. doi:10.1016/S0140-6736(21)01330-1.

NCD Risk Factor Collaboration (NCD-RisC), 2024. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. The Lancet 404, 2077–2093. doi:10.1016/S0140-6736(24)02317-1.

Nissan, H., Burkart, K., Coughlan de Perez, E., Van Aalst, M., Mason, S., 2017. Defining and predicting heat waves in Bangladesh. Journal of Applied Meteorology and Climatology 56, 2653–2670. doi:10.1175/JAMC-D-17-0035.1.

Park, R.J., 2022. Hot temperature and high-stakes performance. Journal of Human Resources 57, 400–434. doi:10.3368/jhr.57.2.0618-9535R3.

Park, R.J., Goodman, J., Hurwitz, M., Smith, J., 2020. Heat and learning. American Economic Journal: Economic Policy 12, 306–339. doi:10.1257/pol.20180612.

Périard, J.D., Racinais, S., Sawka, M.N., 2015. Adaptations and mechanisms of human heat acclimation: Applications for competitive athletes and sports. Scandinavian Journal of Medicine & Science in Sports 25, 20–38. doi:10.1111/sms.12408.

Plumettaz, C., Viswanathan, B., Bovet, P., 2020. Hypertension prevalence based on blood pressure measurements on two vs. one visits: A community-based screening programme and a narrative review. International Journal of Environmental Research and Public Health 17, 9395. doi:10.3390/ijerph17249395.

Pratumvinit, B., Charoenkoop, N., Niwattisaiwong, S., Kost, G.J., Tientadakul, P., 2016. The effects of temperature and relative humidity on point-of-care glucose measurements in hospital practice in a tropical clinical setting. Journal of Diabetes Science and Technology 10, 1094–1100. doi:10.1177/1932296816633485.

Rahman, M.M., Gilmour, S., Akter, S., Abe, S.K., Saito, E., Shibuya, K., 2015. Prevalence and control of hypertension in Bangladesh: a multilevel analysis of a nationwide population-based survey. Journal of Hypertension 33, 465–472. doi:10.1097/HJH.0000000000000421.

Rana, J., Oldroyd, J.C., Hossain, M.B., Katha, S., Khan, M.N., Karim, M.N., Renzetti, S., Islam, R.M., 2024. Rising trends in diabetes and fasting plasma glucose among Bangladeshi adults: Insights from national surveys (2011–2022). medRxiv preprint. doi:10.1101/2024.10.30.24316405. preprint, not peer reviewed; posted 30 October 2024.

Raymond, C., Matthews, T., Horton, R.M., 2020. The emergence of heat and humidity too severe for human tolerance. Science Advances 6, eaaw1838. doi:10.1126/sciadv.aaw1838.

Riaz, B.K., Islam, M.Z., Islam, A.N.M.S., Zaman, M.M., Hossain, M.A., Rahman, M.M., Khanam, F., Amin, K.M.B., Noor, I.N., 2020. Risk factors for non-communicable diseases in Bangladesh: findings of the population-based cross-sectional national survey 2018. BMJ Open 10, e041334. doi:10.1136/bmjopen-2020-041334.

Romano, J.P., Wolf, M., 2005. Stepwise multiple testing as formalized data snooping. Econometrica 73, 1237–1282. doi:10.1111/j.1468-0262.2005.00615.x.

Roodman, D., Nielsen, M.Ø., MacKinnon, J.G., Webb, M.D., 2019. Fast and wild: Bootstrap inference in Stata using boottest. The Stata Journal 19, 4–60. doi:10.1177/1536867X19830877.

Sathi, N.J., Islam, M.A., Ahmed, M.S., Islam, S.M.S., 2022. Prevalence, trends and associated factors of hypertension and diabetes mellitus in Bangladesh: Evidence from BHDS 2011 and 2017–18. PLOS ONE 17, e0267243. doi:10.1371/journal.pone.0267243.

Sawka, M.N., Convertino, V.A., Eichner, E.R., Schnieder, S.M., Young, A.J., 2000. Blood volume: Importance and adaptations to exercise training, environmental stresses, and trauma/sickness. Medicine & Science in Sports & Exercise 32, 332–348. doi:10.1097/00005768-200002000-00012.

Segerhag, E., Gyberg, V., Ioannides, K., Jennings, C., Rydén, L., Stagmo, M., Hage, C., 2015. Accuracy of a simplified glucose measurement device—the HemoCue Glucose 201RT. Diabetes Technology & Therapeutics 17, 755–758. doi:10.1089/dia.2014.0354.

Selvin, E., Crainiceanu, C.M., Brancati, F.L., Coresh, J., 2007. Short-term variability in measures of glycemia and implications for the classification of diabetes. Archives of Internal Medicine 167, 1545–1551. doi:10.1001/archinte.167.14.1545.

Senay, L.C., Mitchell, D., Wyndham, C.H., 1976. Acclimatization in a hot, humid environment: Body fluid adjustments. Journal of Applied Physiology 40, 786–796. doi:10.1152/jappl.1976.40.5.786.

Sherwood, S.C., Huber, M., 2010. An adaptability limit to climate change due to heat stress. Proceedings of the National Academy of Sciences 107, 9552–9555. doi:10.1073/pnas.0913352107.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., et al., 2024. The ERA5 global reanalysis from 1940 to 2022. Quarterly Journal of the Royal Meteorological Society 150, 4014–4048. doi:10.1002/qj.4803.

Somanathan, E., Somanathan, R., Sudarshan, A., Tewari, M., 2021. The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing. Journal of Political Economy 129, 1797–1827. doi:10.1086/713733.

Stergiou, G.S., Palatini, P., Modesti, P.A., Asayama, K., Asmar, R., Bilo, G., de la Sierra, A., Dolan, E., Head, G., Kario, K., et al., 2020. Seasonal variation in blood pressure: Evidence, consensus and recommendations for clinical practice. consensus statement by the European Society of Hypertension Working Group on Blood Pressure Monitoring and Cardiovascular Variability. Journal of Hypertension 38, 1235–1243. doi:10.1097/HJH.0000000000002341.

Stork, A.D.M., Kemperman, H., Erkelens, D.W., Veneman, T.F., 2005. Comparison of the accuracy of the hemocue glucose analyzer with the Yellow Springs Instrument glucose oxidase analyzer, particularly in hypoglycemia.

European Journal of Endocrinology 153, 275–281. doi:10.1530/eje.1.01952.

Stull, R., 2011. Wet-bulb temperature from relative humidity and air temperature. Journal of Applied Meteorology and Climatology 50, 2267–2269. doi:10.1175/JAMC-D-11-0143.1.

Su, D., Du, H., Zhang, X., Qian, Y., Chen, L., Chen, Y., Guo, Y., Bian, Z., Chen, Z., Li, L., Yu, M., 2014. Season and outdoor temperature in relation to detection and control of hypertension in a large rural Chinese population. International Journal of Epidemiology 43, 1835–1845. doi:10.1093/ije/dyu158.

Tang, Z., Lee, J.H., Louie, R.F., Kost, G.J., 2000. Effects of different hematocrit levels on glucose measurements with handheld meters for point-of-care testing. Archives of Pathology & Laboratory Medicine 124, 1135–1140. doi:10.5858/2000-124-1135-EODHLO.

Thirup, P., 2003. Haematocrit: Within-subject and seasonal variation. Sports Medicine 33, 231–243. doi:10.2165/00007256-200333030-00005.

Tseng, C.L., Brimacombe, M., Xie, M., Rajan, M., Wang, H., Kolassa, J., Crystal, S., Chen, T.C., Pogach, L., Safford, M., 2005. Seasonal patterns in monthly Hemoglobin A1c values. American Journal of Epidemiology 161, 565–574. doi:10.1093/aje/kwi071.

Valdés, S., Doulatram-Gamgaram, V., Lago, A., García Torres, F., Badía-Guillén, R., Olveira, G., Goday, A., Calle-Pascual, A., Castaño, L., Castell, C., et al., 2019. Ambient temperature and prevalence of diabetes and insulin resistance in the Spanish population: Di@bet.es study. European Journal of Endocrinology 180, 273–280. doi:10.1530/EJE-18-0818.

Vecellio, D.J., Wolf, S.T., Cottle, R.M., Kenney, W.L., 2022. Evaluating the 35 °c wet-bulb temperature adaptability threshold for young, healthy subjects (PSU HEAT project). Journal of Applied Physiology 132, 340–345. doi:10.1152/japplphysiol.00738.2021.

Verdecchia, P., Angeli, F., Poeta, F., Reboldi, G.P., Borgioni, C., Pittavini, L., Porcellati, C., 2004. Validation of the A&D UA-774 (UA-767Plus) device for self-measurement of blood pressure. Blood Pressure Monitoring 9, 225–229. doi:10.1097/00126097-200408000-00008.

Wang, Q., Li, C., Guo, Y., Barnett, A.G., Tong, S., Phung, D., Chu, C., Dear, K., Wang, X., Huang, C., 2017. Environmental ambient temperature and blood pressure in adults: A systematic review and meta-analysis. Science of The Total Environment 575, 276–286. doi:10.1016/j.scitotenv.2016.10.019.

White, C., 2017. The dynamic relationship between temperature and morbidity. Journal of the Association of Environmental and Resource Economists 4, 1155–1198. doi:10.1086/692098.

Whitehead, R.D., Mei, Z., Mapango, C., Jefferds, M.E.D., 2019. Methods and analyzers for hemoglobin measurement in clinical laboratories and field settings. Annals of the New York Academy of Sciences 1450, 147–171. doi:10.1111/nyas.14124.

Wolf, S.T., Cottle, R.M., Vecellio, D.J., Kenney, W.L., 2022. Critical environmental limits for young, healthy adults (PSU HEAT project). Journal of Applied Physiology 132, 327–333. doi:10.1152/japplphysiol.00737.2021.

World Health Organization, 2011. Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity. Vitamin and Mineral Nutrition Information System WHO/NMH/NHD/MNM/11.1. World Health Organization. Geneva. URL: https://www.who.int/publications/i/item/WHO-NMH-NHD-MNM-11.1.

World Health Organization, 2021. Guideline for the Pharmacological Treatment of Hypertension in Adults. Technical Report. World Health Organization. Geneva. URL: https://www.who.int/publications/i/item/9789240033986. iSBN 978-92-4-003398-6 (electronic version).

World Health Organization, 2024. Guideline on Haemoglobin Cutoffs to Define Anaemia in Individuals and Populations. Technical Report. World Health Organization. Geneva. URL: https://www.who.int/publications/i/item/9789240088542. iSBN 978-92-4-008854-2 (electronic version).

World Health Organization, International Diabetes Federation, 2006. Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia:

Report of a WHO/IDF Consultation. Technical Report. World Health Organization. Geneva. URL: https://iris.who.int/handle/10665/43588. iSBN 978-92-4-159493-6.

Wright, J.A., Yang, H., Walker, K., 2012. Do international surveys and censuses exhibit 'dry season' bias? Population, Space and Place 18, 116–126. doi:10.1002/psp.681.

Xu, R., Zhao, Q., Coelho, M.S., Saldiva, P.H., Zoungas, S., Huxley, R.R., Abramson, M.J., Guo, Y., Li, S., 2019. Association between heat exposure and hospitalization for diabetes in Brazil during 2000–2015: A nationwide case-crossover study. Environmental Health Perspectives 127, 117005. doi:10.1289/EHP5688.