**Online Supplemental Material**

**Statistical analysis**

Analyses of specific causes of death covered all major classes of diagnoses (chapters in ICD-10), excluding those for which there were less than 100 deaths during the prolonged heat wave period (6 July to 18 August, 2010). We applied generalized linear models20 to estimate expected death counts during this period for each cause of death separately using model (1) and data from the four preceding years (2006-2009). The first three terms in model (1) reproduced the long-term trend, while the sum of four sinusoidal terms accounted for seasonal behavior:

 (1)

where the betas denote regression coefficients while *Θi* are phases estimated by

cos(*x*+*y*) = *cos(x)*× *cos(y)*  - *sin(x)*× *sin(y)*.

Model (1) was used for predictions beyond the period 2006-2009 because it assumed that the same type of seasonal regularity would have been observed during the summer of 2010 in the absence of the heat wave.

Model (2) was used to study interactions between high temperatures and air pollution in relation to mortality. In both models we assumed Gaussian errors. We applied the identity link in model (1) and the log link in model (2).

*ln[E(M)] = Const + f1(t) + f2(pollution) + f3(T) + Int + S(RH, 4)* (2)

where *ln[E(M)]* is natural logarithm of expected death counts on each day between 1 January 2006 and 31 December 2010. Model (2) is conditional on the value of the natural logarithm of mortality on the previous day *ln(M1)*.

*f1*=

*f2*=

*f3= S(T01, 8) + S(T26, 8) + S(daynum, 6)*

.

*T, PM, O* and *RH* denote mean daily temperature in °C, PM10, O3 and relative humidity in %. Subindexes 01 and 26 denote that the arguments have been averaged over days 0-1 and 2-6 prior to the day of death.

Thus, *f1* comprises all time-dependent terms in the regression equation, including long-term time trend, seasonal periodicity and day-of-week variation in deaths. *S* denotes a restricted cubic spline function of day of year (*DOY*), so that the modeled seasonal variation in deaths is exactly the same each year. {*DOW*} are day-of-week indicator variables.

Function *f2* describes variations in deaths due to variations in air pollutant levels over the preceding week; *W* is a winter half-year indicator variable; *W*=1 between 17 October and 16 April of each year, and *W*=0 during all other days. Thus, the last two terms in the equation for *f2* account for interaction between ozone and season, dichotomizing the regression coefficients into winter and summer half-year periods.

The first and the second terms in *f3* define a flexible temperature-mortality exposure-response function, where *S* denotes a restricted cubic spline function with specified number of degrees of freedom. Following terminology of Gasparrini and Armstrong,21 this function describes the “main” effect of temperature, due to independent effects of daily temperature levels, while the last term in *f3* describes the “added” effect of temperature, due to consecutive heat-wave days: *daynum* defines the day number in each heat-wave; *daynum*=0 for all non-heat-wave days. The positions of knots in *S(T01)* and *S(T26)* were selected to account for the curvature of the temperature/mortality relationship across the whole range of annual temperatures. Three knots were positioned in the ascending segment of this curve (above 18ºC), one knot near the minimum of this curve (18ºC) and five knots in the descending segment. The knots in *S(daynum, 6)* are regularly spaced every four days until day 16 and then every eight days until the end of the heat-wave. Thus, functions *f2* and *f3* account for possibly lagged effects of temperature and pollution, including short-term harvesting during one week.

The interaction is assumed to occur only on the day of the death and on the previous day (a function of *PM01* and *T01*) and at the temperatures above *T01*=18°C. This choice reflects an assumption that interaction may take place between two risk factors: high temperatures and air pollution. Only the interaction between temperature and PM10 is considered in the model.

The partial inputs from different risk factors to the total non-accidental excess mortality during the heat wave period were estimated by linearization of the multiplicative model (2), using second-order Taylor expansion of the exponential function. The partial inputs were estimated from Model (2) without the autoregression term *ln(M1)*, aiming to explain all of the excess mortality by environmental factors. Although model (2) is highly non-linear on the “worst” days, partial inputs of each contributing risk factor to total excess mortality can be calculated fairly accurately *on average* for the heat wave period because average relative increases attributable to each factor were not that high. The relative error in the calculation of the partial inputs was less than 34% on the worst day (August 8), and less than 10% on average for the heat wave period.

The risks of PM10 at different temperatures were expressed as percentage increase in mortality per 10 µg/m3, related to average weekly exposure. This was calculated as the sum of three regression coefficients *βPM01* + *βPM26* + *βInt*×max(0, *T01* - 18) estimated from Model 2. The estimated values were: *βPM01* =-0.11(-0.39; 0.17); *βPM26* =0.54(0.34; 0.74); *βInt*=0.084(0.054; 0.114). The regression coefficient *βPM01* is not significant under the specifications of Model 2, because the effect of pollution on days 0-1 is captured by the interaction term. Note that the sum of the three coefficients is statistically significant. We also checked that Model 2 *without* the interaction term produced a highly significant estimate of *βPM01* which was greater than *βPM26* . Similarly, the sum of the ozone regression coefficients *βO01*=.79(.08; 1.50) and *βO26*=.50(-.26; 1.27) produces a significant estimate of percentage increase in mortality per 10 mg/m3 of weekly average exposure: 1.29(.23; 2.35).

Three types of sensitivity analyses were performed to test the robustness of obtained unit risks to key assumptions of Model (2):

1) with respect to temporal confounding, we replaced *S(DOY)* with the set of sinusoidal functions used in Model (1): *S(DOY)* →

The two smoothing techniques used to account for seasonal confounding had the same number of degrees of freedom (8 df/year), which contributes to explaining the convergence of the results.

2) with respect to control for main effect of heat, we varied the degree of residual confounding by testing three alternative modeling choices: *S(T01)*; *S(T01)*+*S(T26)*; *S(T01)*+*S(T23)*+*S(T46)* where each spline had 8 df. Consequently, the main effect was averaged over one week after the exposure and measured relative to 18ºC.

3) with respect to control for “added” effect of prolonged heat wave, we tried three heat wave definitions, based on the 97th(22.8ºC), 98th (23.6ºC) and 99th (25.1ºC) percentiles of the distribution of daily temperatures during the years of follow-up. Under these definitions, the length of the major heat wave of 2010 varied from 45 to 44 and 38 days, respectively.

The results of sensitivity analysis are summarized in Table e1. Figure e1 shows the daily number of deaths, mean temperature and PM10 levels in Moscow 2006-2010. Finally, the Stata code used to generate the results is provided.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model description | AIC | max *RRadded*a | daynumberb | Relative change in mortality (%) |
| ≤18°C | 22°C | 30°C |
| Seasonality | Main effect of temperature | heat-wave threshold | PM10c | PM10 | Main effect of temperatured | PM10 | Main effect of temperature |
| *S(DOY)* | *S(T01, 8)*+*S(T26, 8)* | 98% | 8.9249 | 1.38 (1.27-1.50) | 14 | .43(.09-.77) | .77(.40-1.13) | 1.01(.98-1.03) | 1.44(.94-1.94) | 1.44(1.26-1.63) |
| sinusoidal | *S(T01, 8)*+*S(T26, 8)* | 98% | 8.9351 | 1.40 (1.28-1.52) | 14 | .42(.08-.76) | .76(.40-1.12) | 1.00(.97-1.02) | 1.44(.94-.94) | 1.40(1.22-1.58) |
| *S(DOY)* | *S(T01, 8)* | 98% | 8.9793 | 1.44(1.35-1.53) | 14 | .52(.21-.79) | .79(.47-1.12) | 1.02(1.01-1.04) | 1.34(.88-1.79) | 1.22(1.13-1.31) |
| *S(DOY)* | *S(T01, 8)*+*S(T23, 8)*+*S(T46, 8)* | 98% | 8.9204 | 1.38(1.27-1.49) | 14 | .47(.17-.81) | .79(.43-1,16) | 1.01(.97-1.05) | 1.44(.94-1.94) | 1.47(1.21-1.74) |
| *S(DOY)* | *S(T01, 8)*+*S(T26, 8)* | 97% | 8.9268 | 1.35(1.25-1.47) | 15 | .42(.08-.76) | .76(.39-1.12) | 1.00(.97-1.02) | 1.43(.93-1.93) | 1.45(1.26-1.64) |
| *S(DOY)* | *S(T01, 8)*+*S(T26, 8)* | 99% | 8.9418 | 1.31(1.20-1.43) | 14 | .33(-.01-.67) | .61(.25-.97) | 1.00(.97-1.02) | 1.17(.69-1.65) | 1.54(1.35-1.73) |

**Table e1**. Risks of non-accidental mortality in relation to daily mean temperature and PM10 in Moscow under different model assumptions

a max *RRadded* = maximum value of the “added” effect of the number of days into the heat wave

b day number = the day during the heat wave, on which the maximum value of *RRadded* was achieved

c per10 μg/m3 of PM10 at specified daily mean temperatures

d relative to 18°C due to independent effect of daily mean temperatures

Figure e1. Daily number of deaths, PM10concentrations, and temperature in Moscow 2006-2010.

gen daynum = \_n

order date daynum

gen daynum2=daynum^2

tsset daynum

set obs `=\_N+59'

forvalues x = -28(1)30 {

local row=\_N-30+`x'

replace T01=`x' in `row'

replace T26=`x' in `row'

replace T23=`x' in `row'

replace T46=`x' in `row'

}

mkspline t = T01 , cubic nknots(7) di

mkspline u = T26 , cubic nknots(7) di

mkspline rh = RH , cubic nknots(6) di

mkspline h = H, cubic knots(4 8 12 16 24 32 40) di

gen P = max(0, T01-18)\*(PM01)

gen NA\_1 = L1.NA

gen NA\_1log = ln(NA\_1)

\* spline(DOY) 6 df/year

mkspline spline = DOY, cubic nknots(7) di

\* spline(DOY) 8 df/year

drop spline\*

mkspline spline = DOY, cubic knots(7 51 95 139 183 227 271 315 359) di

gen cos1=cos(2\*\_pi\*daynum/365)

gen sin1=sin(2\*\_pi\*daynum/365)

gen cos3=cos(6\*\_pi\*daynum/365)

gen sin3=sin(6\*\_pi\*daynum/365)

gen cos4=cos(8\*\_pi\*daynum/365)

gen sin4=sin(8\*\_pi\*daynum/365)

gen cos2=cos(4\*\_pi\*daynum/365)

gen sin2=sin(4\*\_pi\*daynum/365)

gen cos6=cos(12\*\_pi\*daynum/365)

gen sin6=sin(12\*\_pi\*daynum/365)

\*Model 2

drop t\*

mkspline t = T01, cubic knots(-15.2 -2.8 2.02859 7.1 13.1 17.6 25.4325) di

drop u\*

mkspline u = T26, cubic knots(-14.5 -2.8 2 7.3 12.9 17.41109 24.9 ) di

\*Model 3

drop t\*

gen t1 = T01

gen t2 = max(0, T01+15.2)^3

gen t3 = max(0, T01+2.8)^3

gen t4 = max(0, T01-2)^3

gen t5 = max(0, T01-7.1)^3

gen t6 = max(0, T01-13.1)^3

gen t7 = max(0, T01-17.6)^3

gen t8 = max(0, T01-25.4)^3

drop u\*

gen u1 = T26

gen u2 = max(0, T26+14.5)^3

gen u3 = max(0, T26+2.8)^3

gen u4 = max(0, T26-2)^3

gen u5 = max(0, T26-7.3)^3

gen u6 = max(0, T26-12.9)^3

gen u7 = max(0, T26-17.4)^3

gen u8 = max(0, T26-24.9)^3

\*Model 4 preferred

drop t\*

mkspline t = T01, cubic knots(-15.2 -2.8 2.02859 7.1 13.1 17.6 19.4 22.5 27.8) di

drop u\*

mkspline u = T26, cubic knots(-14.5 -2.8 2 7.3 12.9 17.4 19.1 21.7 27.7) di

\* trigonometric spline

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 cos1 sin1 cos2 sin2 cos3 sin3 cos4 sin4 i.DOW t\* u\* rh\* h\*, link(log) fam(Gauss)

\* periodic cubic spline

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* u\* rh\* h\*, link(log) fam(Gauss)

\*without interaction

glm NA NA\_1log PM01 PM26 O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* u\* rh\* h\*, link(log) fam(Gauss)

\* Sensitivity S(T01, 8)

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* rh\* h\*, link(log) fam(Gauss)

xblc t\* , covname(T01) at(-28(1)30) reference(18) eform

xblc u\* , covname(T26) at(-28(1)30) reference(18) eform

\* Sensitivity S(T01, 8)+S(T23, 8)+S(T46, 8)

mkspline v = T23 , cubic nknots(7) di

mkspline w = T46 , cubic nknots(7) di

\*leave first 6 knots and place the rest at 88% 94% 99% then run 'replace'

drop v\*

mkspline v=T23, cubic knots( -15.2 -2.8 2.02859 7.1 13.1 17.6 19.4 22.6 27.7) di

drop w\*

mkspline w=T46, cubic knots(-15.2325 -2.7 2 7.2 13 17.5 19.3 22.2 27.6) di

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* v\* w\* rh\* h\*, link(log) fam(Gauss)

xblc t\* , covname(T01) at(-28(1)30) reference(18) eform

xblc v\* , covname(T23) at(-28(1)30) reference(18) eform

xblc w\* , covname(T46) at(-28(1)30) reference(18) eform

\*Model 97%

mkspline k = K, cubic knots(4 8 12 16 24 32 40) di

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* u\* rh\* k\*, link(log) fam(Gauss)

xblc k\* , covname(K) at(0(1)45) reference(0) gen(var1 var2) eform

\*Model 99%

mkspline l = L, cubic knots(4 8 12 16 24 32) di

glm NA NA\_1log PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* u\* rh\* l\*, link(log) fam(Gauss)

xblc l\* , covname(L) at(0(1)38) reference(0) gen(var1 var2) eform

\*Model for Fig 4

glm NA PM01 PM26 P O01 W01 O26 W26 daynum daynum2 spline\* i.DOW t\* u\* rh\* h\*, link(log) fam(Gauss) vce(hac nw 10)

drop (var1 var2)

xblc h\* , covname(H) at(0(1)44) reference(0) gen(var1 var2) eform

predict Y, mu