Supplementary materials

Urban $PM_{2.5}$ Pollution in Kazakhstan: Health Burden and Economic Costs

Aset Muratuly[a,b,c], *Ravkat Mukhtarov*[a,b], *Ivan Radelyuk*[a,b,c], *Ferhat Karaca*[d,e], *Nassiba Baimatova*[a,b,c*]

[a]Faculty of Chemistry and Chemical Technology, Center of Physical Chemical Methods of Research and Analysis, Al-Farabi Kazakh National University, Almaty, Kazakhstan

[b]Environmental and Analytical Chemistry Laboratory, Almaty, Kazakhstan

[c] Department of Chemistry and Chemical Technology, Toraighyrov University, Pavlodar, Kazakhstan

[d]School of Engineering and Digital Science, Department of Civil and Environmental Engineering, Nazarbayev University, Astana, Kazakhstan

[e] The Environment and Resource Efficiency Cluster, Nazarbayev University, Astana, Kazakhstan

*Corresponding author: Nassiba Baimatova. E-mail address: baimatova@cfhma.kz

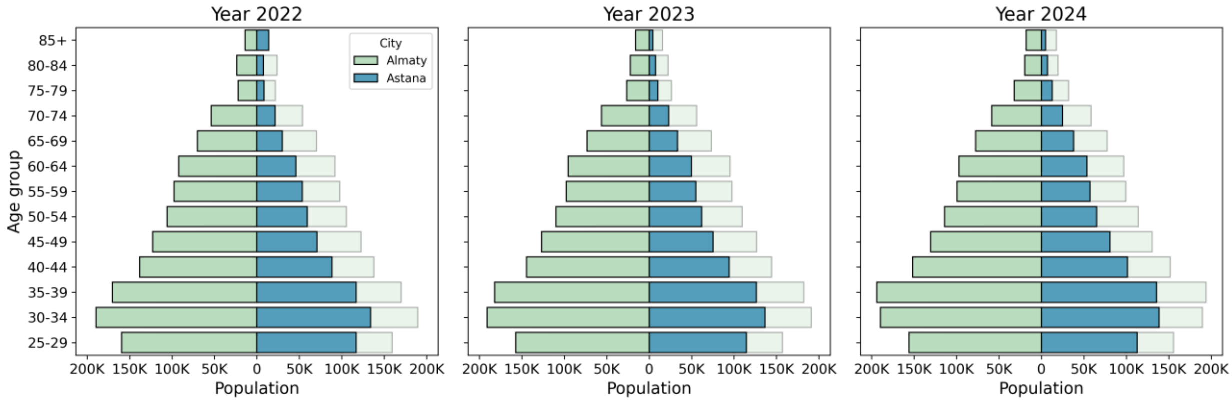

Fig. S1. Statistics of the population for Almaty and Astana for 2022-2024 by age groups.

Table S1

The detailed population distribution and mortality rates stratified by 5-year age groups for Almaty and Astana.

City	Age group	2022		2023		2024	
		population	mortality	population	mortality	population	mortality
Almaty	25-29	159478	81	157187	74	155791	66
	30-34	189482	177	190899	189	189598	146
	35-39	169977	249	182115	265	193839	253
	40-44	137981	397	144402	361	151810	361
	45-49	122851	458	126636	460	130508	492
	50-54	105615	575	109791	619	114159	613
	55-59	97546	820	97640	771	99433	730
	60-64	92014	1174	95240	1165	97045	1170
	65-69	70132	1404	73221	1404	77568	1393
	70-74	53874	1575	56118	1546	58458	1612
	75-79	21886	984	26289	1139	32016	1407
	80-84	23738	1806	22189	1448	19698	1222
	85+	13904	2051	15865	2009	17846	2242
Astana	25-29	116898	56	114536	66	112771	83
	30-34	133776	111	136610	101	138803	94
	35-39	116866	171	126251	142	135818	163
	40-44	88558	209	94408	228	101385	227
	45-49	70936	247	75252	291	80791	263
	50-54	59288	339	61992	340	65212	346
	55-59	53541	451	55102	406	57340	440
	60-64	45837	571	49724	616	53557	673
	65-69	30077	616	33615	631	37970	708
	70-74	21321	593	23021	647	24741	691
	75-79	8238	360	10121	449	12749	619
	80-84	7645	627	7606	513	7095	495
	85+	13904	587	4268	593	5056	644

Table S2

GEMM model parameters utilized for estimating mortality in the Chinese male cohort. (Burnett et al., 2018).

Cause of Death	Age Range	θ	St. Err θ	α	μ	ν
NCD+LRI	>25	0.143	0.01807	1.6	15.5	36.8
	27.5	0.1585	0.01477	1.6	15.5	36.8
	32.5	0.1577	0.0147	1.6	15.5	36.8
	37.5	0.157	0.01463	1.6	15.5	36.8
	42.5	0.1558	0.0145	1.6	15.5	36.8
	47.5	0.1532	0.01425	1.6	15.5	36.8
	52.5	0.1499	0.01394	1.6	15.5	36.8
	57.5	0.1462	0.01361	1.6	15.5	36.8
	62.5	0.1421	0.01325	1.6	15.5	36.8
	67.5	0.1374	0.01284	1.6	15.5	36.8
	72.5	0.1319	0.01234	1.6	15.5	36.8
	77.5	0.1253	0.01174	1.6	15.5	36.8
	85	0.1141	0.01071	1.6	15.5	36.8
IHD	>25	0.2969	0.01787	1.9	12	40.2
	27.5	0.507	0.02458	1.9	12	40.2
	32.5	0.4762	0.02309	1.9	12	40.2
	37.5	0.4455	0.0216	1.9	12	40.2
	42.5	0.4148	0.02011	1.9	12	40.2
	47.5	0.3841	0.01862	1.9	12	40.2
	52.5	0.3533	0.01713	1.9	12	40.2
	57.5	0.3226	0.01564	1.9	12	40.2
	62.5	0.2919	0.01415	1.9	12	40.2
	67.5	0.2612	0.01266	1.9	12	40.2
	72.5	0.2304	0.01117	1.9	12	40.2
	77.5	0.1997	0.00968	1.9	12	40.2
	85	0.1536	0.00745	1.9	12	40.2
Stroke	>25	0.272	0.07697	6.2	16.7	23.7
	27.5	0.4513	0.11919	6.2	16.7	23.7
	32.5	0.424	0.11197	6.2	16.7	23.7
	37.5	0.3966	0.10475	6.2	16.7	23.7
	42.5	0.3693	0.09752	6.2	16.7	23.7
	47.5	0.3419	0.0903	6.2	16.7	23.7
	52.5	0.3146	0.08307	6.2	16.7	23.7
	57.5	0.2872	0.07585	6.2	16.7	23.7
	62.5	0.2598	0.06863	6.2	16.7	23.7
	67.5	0.2325	0.0619	6.2	16.7	23.7
	72.5	0.2051	0.05418	6.2	16.7	23.7
	77.5	0.1778	0.04695	6.2	16.7	23.7
	85	0.1368	0.03611	6.2	16.7	23.7
COPD	>25	0.251	0.06762	6.5	2.5	32
LC	>25	0.2942	0.06147	6.2	9.3	29.8
LRI	>25	0.4468	0.11735	6.4	5.7	8.4

Table S3

Number of potential avoidable deaths from various counterfactual scenarios of reaching interim WHO $PM_{2.5}$ targets by 2022.

Year	City	Model cohort	Model	Data Source	Target $PM_{2.5}$ concentration (μg m^{-3})			
					25	15	10	5
2022	Almaty	w_cmc	Five causes	AirKaz	442	910	1206	1610
				AirNow	417	884	1181	1585
			NCD+LRI	AirKaz	508	1013	1335	1814
				AirNow	478	983	1304	1784
		wo_cmc	Five causes	AirKaz	397	853	1142	1507
				AirNow	375	832	1121	1486
			NCD+LRI	AirKaz	537	1073	1401	1862
				AirNow	506	1042	1369	1830
	Astana	w_cmc	Five causes	AirKaz	-*	169	300	479
				AirNow		149	280	459
			NCD+LRI	AirKaz		177	316	522
				AirNow		156	294	501
		wo_cmc	Five causes	AirKaz		166	293	456
				AirNow		146	274	436
			NCD+LRI	AirKaz		188	329	527
				AirNow		165	306	504
2023	Almaty	w_cmc	Five causes	AirNow	144	603	894	1290
			NCD+LRI		161	655	969	1439
		wo_cmc	Five causes		135	583	866	1224
			NCD+LRI		172	696	1016	1467
	Astana	w_cmc	Five causes		-*	67	201	384
			NCD+LRI			70	211	422
		wo_cmc	Five causes			66	197	362
			NCD+LRI			74	217	419
2024	Almaty	w_cmc	Five causes			440	735	1138
			NCD+LRI			474	795	1274
		wo_cmc	Five causes			430	718	1082
			NCD+LRI			503	829	1289
	Astana	w_cmc	Five causes			89	233	431
			NCD+LRI			93	246	474
		wo_cmc	Five causes			88	228	407
			NCD+LRI			98	253	472

*Note: Average concentration in the year is higher than the target concentration

Table S4

Potential economic benefits from various counterfactual scenarios of reaching interim WHO $PM_{2.5}$ targets by 2022 in constant 2022 million USD.

Year	City	Model cohort	Model	Data Source	Target $PM_{2.5}$ concentration ($\mu g\ m^{-3}$)			
					25	15	10	5
2022	Almaty	w_cmc	Five causes	AirKaz	875	1,801	2,388	3,187
				AirNow	825	1,751	2,338	3,137
			NCD+LRI	AirKaz	1,006	2,005	2,642	3,592
				AirNow	946	1,945	2,582	3,532
		wo_cmc	Five causes	AirKaz	785	1,689	2,261	2,983
				AirNow	743	1,647	2,219	2,941
			NCD+LRI	AirKaz	1,064	2,124	2,773	3,685
				AirNow	1,002	2,062	2,711	3,623
	Astana	w_cmc	Five causes	AirKaz	-*	334	593	949
				AirNow		294	553	908
			NCD+LRI	AirKaz		351	625	1,034
				AirNow		308	582	991
		wo_cmc	Five causes	AirKaz		328	581	902
				AirNow		289	542	863
			NCD+LRI	AirKaz		372	651	1,043
				AirNow		326	605	998
2023	Almaty	w_cmc	Five causes	AirNow	283	1,180	1,749	2,524
			NCD+LRI		315	1,281	1,896	2,814
		wo_cmc	Five causes		264	1,140	1,695	2,395
			NCD+LRI		337	1,362	1,989	2,870
	Astana	w_cmc	Five causes		-*	132	393	751
			NCD+LRI			137	413	825
		wo_cmc	Five causes			130	385	709
			NCD+LRI			144	425	821
2024	Almaty	w_cmc	Five causes			921	1,540	2,383
			NCD+LRI			992	1,664	2,666
		wo_cmc	Five causes			900	1,503	2,264
			NCD+LRI			1,052	1,736	2,699
	Astana	w_cmc	Five causes			187	488	902
			NCD+LRI			195	514	991
		wo_cmc	Five causes			184	478	852
			NCD+LRI			205	530	988

*Note: Average concentration in the year is higher than the target concentration

Text S1 Data processing and quality control measures

$PM_{2.5}$ levels were measured in real-time by each instrument; however, hourly data was included in the datasets. Hourly $PM_{2.5}$ data were quality-checked prior to analysis, with observations of $PM_{2.5}$ concentrations ≤0 µg m^{-3} removed. Although outlier filtering using 0.1st and 99th percentiles was tested, it was not applied. Filtering out these percentiles decreased the annual mean in Almaty from 37.1 µg m^{-3} to 35.0 µg m^{-3}, a reduction of approximately 5.6%.

However, since such concentrations were observed during peaks of high pollution episodes rather than isolated incidents, the decision was made to keep these values in the final analysis, as they represented actual exposure conditions rather than measurement errors. However, data from extremely high outliers (>800 µg m^{-3}) were still excluded, as they indicated sensor malfunctions, data transmission errors, or exceptional pollution events rather than typical ambient conditions.

For each city, data from individual sensors were evaluated for completeness, and only those with at least 80% hourly data availability during the study period were retained to ensure temporal representativeness. Then, data from individual sensors in the Airkaz network were combined sensor by sensor and used in the GEMM model.

Hourly data were converted into daily averages for modeling, while mortality data were available on an annual basis. Therefore, annual mean $PM_{2.5}$ concentrations calculated from the daily data were used to align the temporal resolution between the exposure and outcome datasets.

To estimate the annual population exposure for the GEMM model, hourly $PM_{2.5}$ data were aggregated over the study year. A comparison between the hourly-based and daily-averaged concentration datasets for 2022 revealed a negligible difference in the resulting annual means (37.1 µg m^{-3} vs. 36.8 µg m^{-3}, respectively). This indicates that hourly averaging has an insignificant impact on the long-term exposure estimates, since both approaches tend to the same annual averages when sufficient data coverage is available.

Using hourly data captures more detailed temporal variability and prevents potential information loss due to early aggregation. It also ensures that the annual exposure levels remain consistent with the GEMM framework. Therefore, hourly data were employed in the final mortality estimation to preserve data accuracy and maintain adherence to the original measurement resolution.

Text S2

The VSL estimates for a sample of OECD countries were 3.83 million (2011 USD, PPP-adjusted), corresponding to a GDP per capita of 37 thousand (2011 USD, PPP-adjusted).

Kazakhstan's GDP per capita in local currency (KZT) was 5,284.7 thousand, 6,002.0 thousand, and 6,659.8 thousand KZT in 2022, 2023, and 2024, respectively. The consumer price index (CPI, 2010 = 100) increased from 108.5 in 2011 to 251.1, 288.0, and 313.5 in 2022-2024, respectively. The PPP conversion factor for KZT in 2011 was 82.09 KZT per international dollar.

The U.S. consumer price index (CPI, 2010 = 100) rose from 103.2 in 2011 to 134.2 in 2022.