BILL WHITE
MAYOR

OFFICE OF THE MAYOR
CITY OF HOUSTON
TEXAS

December 22, 2008

Administrator Stephen Johnson
National Emission Standards for Hazardous Air Pollutants from Petroleum Refineries: Residual Risk Standards Docket, Docket ID No. EPA-HQ-OAR-2003-0146
EPA, Air and Radiation Docket and Information Center
1200 Pennsylvania Ave., N.W.
Washington, DC 20460

Robert Lucas
EPA Office of Air Quality Planning
Policy and Program Division (E-143-01)
Research Triangle Park, NC 27711

Re: National Emission Standards for Hazardous Air Pollutants (NESHAPS) from Petroleum Refineries: Supplemental Notice to Proposed Rulemaking; Docket ID No.EPA-HQ-OAR-2003-0146; *Comments*

Via e-mail to a-and-r-docket@epa.gov and to lucas.bob@epa.gov
Via U.S. Mail to the addresses noted above

Dear Administrator Johnson:

The above noted Supplemental Proposal was published in the *Federal Register* on November 10, 2008 (73 *Fed. Reg.* 66694) and a public hearing was held on November 25, 2008 at EPA's Environmental Research Center Auditorium, Research Triangle Park, NC. The initial proposal addressing residual health risks presented by petroleum refineries was published on September 4, 2007. (72 *Fed. Reg.* 50716) and the City of Houston (Houston or the City) filed detailed comments on that proposal, which was inadequate to address the residual risk. Those comments are incorporated by reference. This supplemental proposal, which only includes limited additional controls on cooling water towers and storage vessels (tanks), does little to reduce the level of risk to public health from emissions from petroleum refineries and therefore EPA's proposal still does not meet the standards required by the Clean Air Act.

The comments filed today focus primarily on the methodology used by USEPA in calculating risk reduction estimates, and the data that was used in those calculations. Other comments are included regarding the way EPA has proposed to set corrective action levels for cooling water towers (heat exchange systems). The City has also raised questions regarding the quality of the data used by USEPA in these and other rulemakings in its Petition to USEPA Challenging Emission Factors filed on July 9, 2008. That petition is incorporated by reference here.

Determining the Methodology

The City is unable to determine the specific methodology used by USEPA to calculate risk reduction estimates performed as part of the residual risk assessment on which this rulemaking is based. The City is unable to determine the specific methodology because information about how the estimates were done is found in many places in the record, and in one case not found at all. Some of the information is in various documents in the docket; some of the calculations and assumptions are in other documents that are not in the docket but can be tracked down; but at least one letter does not appear to be available at all (letter from P. Murphy to J. Durham dated 1993 referred to in Lucas, B. 2007a, Memorandum from B. Lucas, EPA/SPPD, to project docket File (EPA Docket. No. EPA-HQ-OAR-2003-0146) *Average Refining Stream Composition*, August 6, 2007. Docket Item No. EPA-HQ-2003-0146-0003, ("Refining Stream Composition")). The memos outlining the risk reduction estimates, cooling water tower and storage vessel revised control options and the average refinery stream composition do not provide enough detail for validation and verification of the calculations. (Memo re Refining Stream Composition, and technical memo from B. Lucas, *Cooling Towers: Control Alternatives and Impact Estimates*, dated October 30, 2008, EPA Docket No. EPA-HQ-OAR-2003-0146 ("Cooling Towers"). The memos are vague, convoluted and key assumptions are not substantiated by USEPA. Finally, sophisticated modeling is conducted in conjunction with back of the envelope assumptions. These assumptions have orders of magnitude of uncertainty so the results based on the modeling and assumptions are unusable.

The comments provided below address selected areas of uncertainty and explain why the data used by USEPA and the calculations based on that data are fundamentally flawed.

Use of Emission Inventory Data Introduces Uncertainty of x5 to x10 in Output.

The estimates used by USEPA assume that the National Emissions Inventory (NEI) data are a refinery's actual emissions. However, USEPA has documented that there is uncertainty in this assumption (i.e., inventory data are under reported). The uncertainty of this assumption affects the calculation of cancer risks. Cancer incidence and risk are linearly related to the concentrations and concentrations are linearly tied to the emission rate. Therefore, using NEI data as a basis for the risk calculation is very likely to result in an underestimation of the risk.

The following is an excerpt from an EPA technical memorandum from Brenda Shine (EPA Docket No. EPA-HQ-OAR-2003-0146, 7/27/2007), "Potential Low Bias of Reported VOC Emissions from the Petroleum Refining Industry") that supports the City's comment:

> "The 2002 national emissions inventory (NEI) indicates that there are 128,000 tons per year of VOC reported, and approximately 1000 tones per year of benzene reported emitted form approximately 150 refineries operating in the US. The total crude capacity of these refineries is approximately 16MM bbls/day; assuming

> actual throughput is close to capacity, and using a crude specific gravity of .85, the average VOC emission factor is approximately .015 wt % of crude throughput, or about 4 times lower than the lowest emissions factor measured for VOC b and NPL DIAL system and an order of magnitude lower than the average NPL factor and the Alberta Research Council factor of 0.2 wt % VOC. The average benzene emission factor is 0.0001 wt %, approximately one-fifth of the 0.0005 wt % factor measured by the Alberta Research Council."

Based on its own assessment of the validity of the NEI, USEPA should not rely on that data to evaluate cancer risk because the NEI data underestimates the cancer risk.

<u>Model Scenario Results are Statistically Indistinguishable and are Contradicted by Empirical Data</u>

The information provided about the annual cancer incidence does not explain the basis of the range of incidence. Table 1 of the Technical Memorandum-Risk Reduction Estimates (Technical Memorandum from B. Lucas, *Risk Reduction Estimates,* dated November 4, 2008, to project docket file, (EPA Docket No. EPA –HQ-OAR-2003-0146) shows a range of annual cancer incidence between the baseline, control level 2 and control level 3 scenarios for storage vessels and cooling towers. Although not indicated, the range of incidence may be assumed to represent the two model plant scenarios or the use of the range of slope factors for benzene. It is not clear. More information about what this range represents could have been provided in footnote b; but it is not.

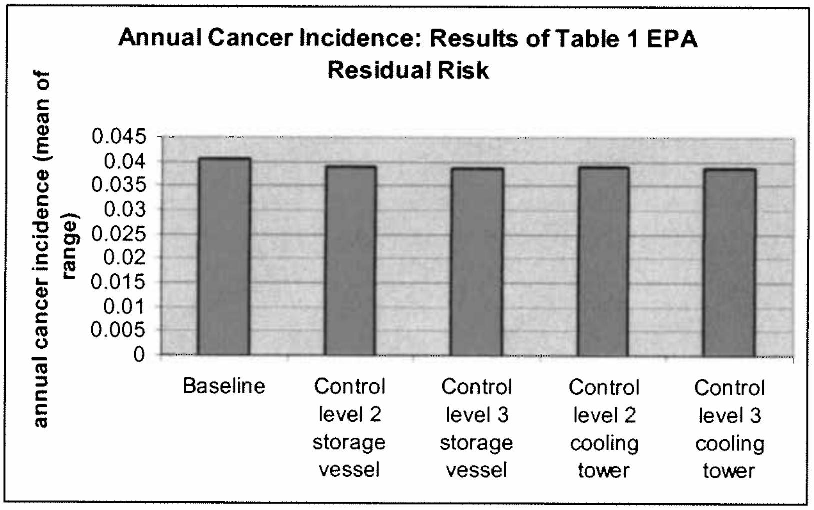
A look at the ranges of annual cancer incidence and annual cancer incidence reduction indicates very little improvement with the different control scenarios (see bar chart below). Statistically, there is no difference between the three groups: baseline, control level 2 and control level 3 for either cooling towers or storage vessels (See Analysis of Variance (ANOVA) below). The confidence intervals of annual cancer incidence and reduction include zero (i.e., it is possible that there is 0 difference between groups).

While the modeling data show no change in risk when different controls are implemented, empirical data from actual individual cooling towers where controls have been employed show substantial improvement in risk reduction. (Galveston-Houston Association for Smog Prevention (GHASP), "Cooling Off: State Investigations Show Reductions in Cooling Tower Emissions," 2006. (GHASP Report)) The GHASP Report indicates substantial reductions in emissions.

The reason for the discrepancy between the modeled lack of reductions and the measured reductions may be a result of the over-averaged model scenarios and inherent assumptions used by USEPA. Regardless of the cause, *the discrepancy points to the underlying lack of confidence in these results*. The model cannot be validated or verified.

<u>Model Scenario Reductions Differ Significantly from Empirical Data; therefore, EPA has not Substantiated its Conclusions about the Effect of Controls</u>

Reductions in emissions of HAPS from cooling water towers are reported by USEPA in Table 1 to range from 4 to 10% depending upon control level.(See *Risk Reduction Memo*) However, the GHASP Report finds the range to be more on the order of 90% reductions as noted between 2003 and 2004. *This large discrepancy supports the conclusion that the reductions from these controls are unknown and that EPA's conclusions regarding the data are unreliable.*



Storage Vessel Options
Results of one-way ANOVA

Summary stats for samples

	Baseline	Control level 2	Control level 3
Sample sizes	2	2	2
Sample means	0.041	0.039	0.039
Sample standard deviations	0.012	0.011	0.012
Sample variances	0.000	0.000	0.000
Weights for pooled variance	0.333	0.333	0.333

Number of samples	3
Total sample size	6
Grand mean	0.039
Pooled variance	0.000
Pooled standard deviation	0.012

OneWay ANOVA table

Source	SS	df	MS	F	p-value
Between variation	0.000	2	0.000	0.016	0.9846
Within variation	0.000	3	0.000		
Total variation	0.000	5			

Confidence intervals for mean differences

Confidence level 95.0%

Tukey method

Difference	Mean diff	Lower	Upper	Signif?
Baseline - Control level 2	0.002	-0.048	0.051	No
Baseline - Control level 3	0.002	-0.047	0.051	No
Control level 2 - Control level 3	0.001	-0.049	0.050	No

Cooling Tower Options
Results of one-way ANOVA

Summary stats for samples

	Baseline	Control level 2	Control level 3
Sample sizes	2	2	2
Sample means	0.041	0.039	0.039
Sample standard deviations	0.012	0.011	0.012
Sample variances	0.000	0.000	0.000
Weights for pooled variance	0.333	0.333	0.333

Number of samples	3
Total sample size	6
Grand mean	0.039
Pooled variance	0.000
Pooled standard deviation	0.012

OneWay ANOVA table

Source	SS	df	MS	F	p-value
Between variation	0.000	2	0.000	0.016	0.9846
Within variation	0.000	3	0.000		
Total variation	0.000	5			

Confidence intervals for mean differences

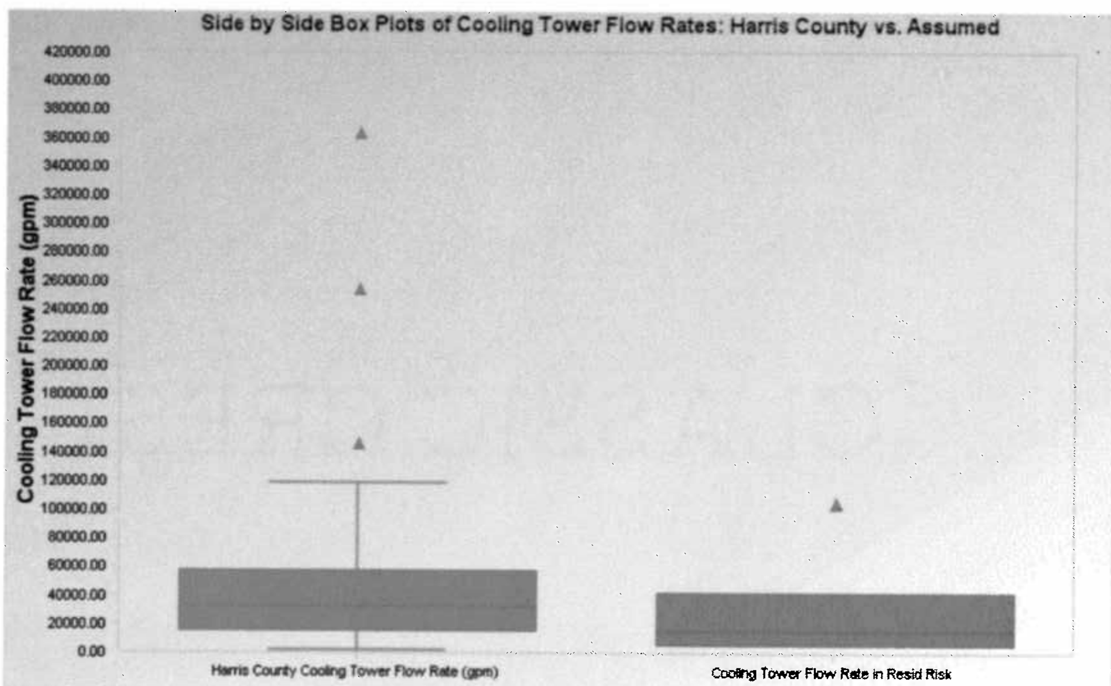
Confidence level 95.0%

Tukey method

Difference	Mean diff	Lower	Upper	Signif?
Baseline - Control level 2	0.002	-0.048	0.051	No
Baseline - Control level 3	0.002	-0.047	0.051	No
Control level 2 - Control level 3	0.001	-0.049	0.050	No

<u>Assumed Distribution of the Size of Cooling Towers in the U.S. Under-Represents Emissions (x2)</u>

The assumed distribution used in the model is statistically different (lower emitting) than the documented distribution in the Houston area. The distribution of 54 cooling water towers by gpm from the Houston area, (from GHASP Report), indicates statistically larger emission rates than those used in the model. The distribution used by USEPA was reconstructed as outlined in Table 2 of the cooling water tower memorandum (Risk Reduction Estimates). Half that count, proportionally weighted, was used to reconstruct the distribution shown below. Although USEPA's distribution was based on data from 149 petroleum refinery cooling water towers nationwide, the GHASP data is from cooling water towers in both chemical plants and refineries. The GHASP data is the only data available for verification of the assumed cooling water tower size used in the model. The data are not normally distributed, therefore, nonparametric inference was used to compare the distribution in Houston/Harris County area to the model distribution (p-value=0.001, the medians are different). The USEPA model data is not representative of Harris County. *The Harris County area mean flow rate is 2x the mean flow rate used in the model.*



Parameters: Nonparametric Test of Significance

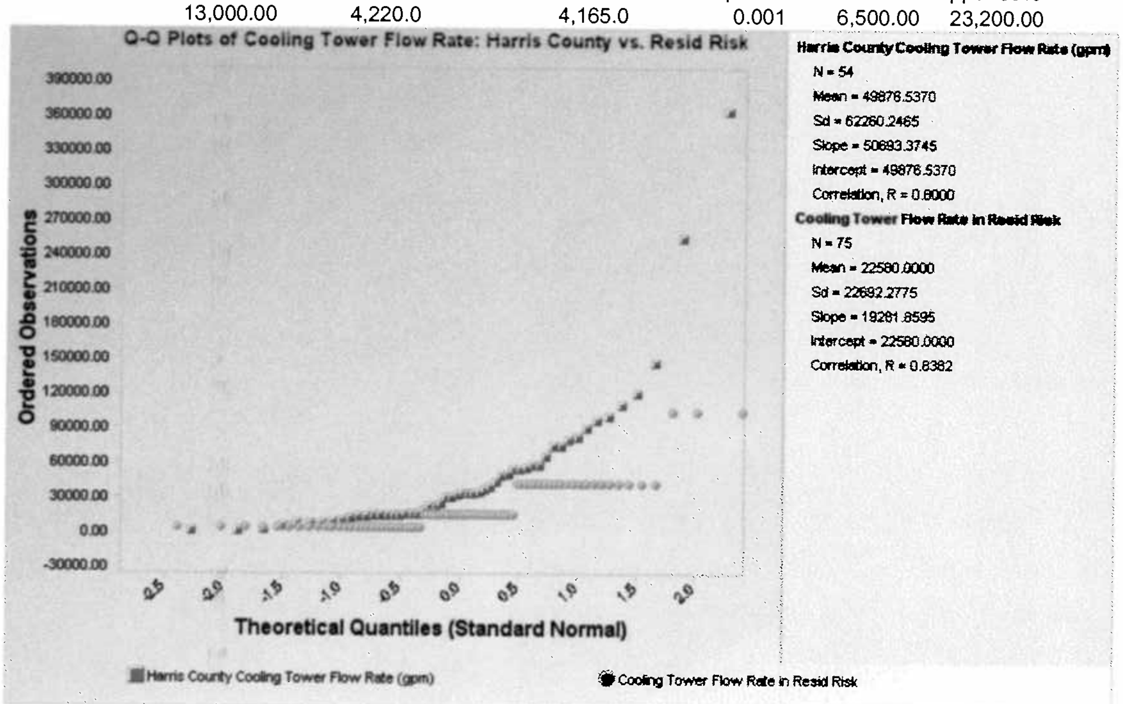
Analysis	2 Sample Rank	Ho: Median Diff. = 0	0
Input Column 1	Harris County Cooling Tower Flow Rate (gpm)	Ha: Not equal to 0	0
Input Column 2	Cooling Tower Flow Rate in Resid Risk	Confidence	0.95

Descriptive Statistics

	N	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Harris County Cooling Tower Flow Rate (gpm)	54	1,000	15,000	32,500	58,000	364,000
Cooling Tower Flow Rate in Resid Risk (gpm)	75	5,000	5,000	15,500	42,000	105,000

Mann-Whitney Rank Analysis

Median Diff.	Rank Sum1	Rank Sum2	p-value	lower 95%	upper 95%
13,000.00	4,220.0	4,165.0	0.001	6,500.00	23,200.00



<u>4Assumed Average Refinery Stream Composition used in the Model Under-Represents Emissions (x3)</u>

The model used by USEPA assumes the HAP concentration to be 15.45% of the strippable total VOC concentrations. The method for developing this percentage is presented in a one paragraph, one table memo regarding average refinery stream composition (Lucas, 2007a, Refining Stream Composition). The details of the HAP weight percent in the refinery stream are excluded from the memo with a reference to a letter from P. Murphy to J. Durham (1993) cited in the Refining Stream Composition memorandum. This document could not be found. Presumably this letter includes more than a simple arithmetic average of, for example, benzene in gasoline, but instead an appropriate treatment of the sample data to reflect confidence in the parameter presented. The content of benzene in a refinery stream is not constant.

The handling of the refinery stream composition data epitomizes the theme of the entire assessment of using a mixture of back of the envelope calculations with sophisticated modeling. On the one hand, the HAP concentration was assumed to be a specific unique number, the simple average of liquid and vapor concentration with no margin of error, while stochastic (Monte Carlo) modeling was conducted to evaluate model scenarios.

The weighting of the average stream composition based on capacity is reasonable but the lack of incorporation of statistical confidence is unreasonable. With the minimal information provided, one unique number for the calculated weighted average concentration for liquid and one for vapor, *it is assumed, without explanation that 50% of the stream will be liquid and 50% vapor.* The simple average of these two numbers is used in the model: 20.56 and 10.34 averaging to 15.45%. There is no confidence associated with the 15.45%. It is the sample mean around which the interval containing the true mean would be constructed. *With only two numbers available (n=2), the 95th upper confidence limit of the mean weighted average is 47%, not 15.45%.*

	gasoline		alkylate		crude oil		reformate		Naptha		calc weighted average Conc. (wt%)	
HAP	liquid	vapor	liquid	vapor	liquid	vapor	liquid	vapor	liquid	vapor	liquid	vapor
2,2, 4-tri	4.72	0.953	35.4	9.4	0.263	0.247	1.36	0.509	0.799	0.355	2.457	0.707
benzene	1.61	0.626	0.118	0.061	0.446	0.807	4.61	3.32	1.244	1.06	1.212	1.024
biphenyl	0.01	0	0	0	0.061	0.00002	0.025	0	0		0.037	0.000
cresols	0.789	0.00065	0	0	0.218	0.0008	0.141	0.0002	0.0191	0.00003	0.315	0.001
cumene	0.849	0.1565	0.004	0.0001	0.126	0.0108	0.968	0.033	0.91	0.0368	0.460	0.049
ethylbenzene	1.605	0.0631	0.011	0.0006	0.346	0.0633	3.73	0.272	1.366	0.118	1.079	0.088
hexane	7.138	4.428	3.22	2.63	2.46	7.11	3.9	4.48	7.53	10.27	4.260	6.440
mtbe	3.54	3.62	0	0	0	0	0	0	0	0	0.821	0.839
napthalene	0.444	0.0006	0.012	0.00002	0.219	0.00125	0.798	0.0018	0.399	0.0011	0.343	0.001
phenol	0.0551	0.0001	0	0	0.323	0.0031	0.0276	0.0001	0.066	0.0003	0.194	0.002
styrene	3.528	0.0882	0	0	0	0	0.0907	0.0042	0	0	0.827	0.021
toluene	7.21	0.842	0.503	0.077	0.878	0.477	14.54	3.14	5.05	1.295	4.161	0.907
xylene	7.17	0.235	0.0212	0.0009	1.42	0.216	13.79	0.8367	5.49	0.395	4.398	0.296
	38.6681	11.0132	39.2892	12.1696	6.76	8.93627	43.9803	12.597	22.8731	13.5312	20.563	10.375
density	0.74		0.692		0.88		0.79		0.79		average=	15.469
EIA 2006 bbl/cd	7901893				1.5E+07						stdev=	7.204
bbl/sd			1238479		1.8E+07		3859070				95th ucl=	47.632
stream day to cal day					0.83136						t-stat=	6.314
calc production bbl/cd			1029619				3208269		3461132			total weights
weighting factor	0.43658		0.0532		1		0.18923		0.20415			1.883

The following are examples of assumptions in the model without empirical support:

• API crude of 30 degrees was assumed- API of crude oils varies from 5 to 50, while the average is 25 to 35, the light crude API 35 to 45 are rich in gasoline and would have higher emissions. West Texas Intermediate is 38 to 40 degree API.

• Time to find and repair leaks were assumed- (e.g., leaks were assumed to exist for ½ the monitoring period, 45 days to identify and fix leaking heat exchanger and 120 days to fix under delay scenario).

• Cooling water tower repairs- assumed 50% would repair as soon as possible and 50% delay (See additional comments below.)

Additional Comments on Cooling Water Tower Repair

Use of a modified El Paso Method to monitor and set corrective action thresholds for cooling water towers is a positive step. However, basing the corrective action threshold of the concentration of strippable volatile organic compounds (VOCs), without specifying the size of the cooling water stream where monitoring occurs is not sufficient. The emission rate from a heat exchange system is proportional to the flow rate of water through the heat exchange system. Therefore, monitoring a heat exchange system with HAP leaks at the proposed corrective action level where the water recirculation rate is 80,000 gpm will have 10 times the emissions compared to a monitoring program where monitoring targets a heat exchange system with a water recirculation rate of 8,000 gpm. Therefore, further detail regarding the locations where monthly monitoring will measure against the corrective action thresholds is necessary so that monitoring will be based on the potential overall emissions from the heat exchange system, not just the concentrations of strippable VOCs.

Conclusion

For all the reasons stated above, the supplemental proposal suffers from the same flaws as the original proposal, namely, the underlying data and methodology are flawed, often contradicting actual, measured empirical data, and the proposed controls do not achieve the standards in the Clean Air Act to reduce risk. The proposed rules, if promulgated without additional review, would be arbitrary and capricious and would not meet the standards of the Administrative Procedure Act, 42, U.S.C. Sec. 7607 (d) (9) (A), the Clean Air Act, or Executive Order 12898. Under these circumstances, Houston urges EPA to (1) adopt the additional controls proposed in the initial proposal and in this supplemental proposal on an interim final basis while it collects the data it needs to perform an acceptable risk assessment, (2) conduct a risk assessment in accordance with the statute, and (3) propose additional controls as required under the Clean Air Act, as soon as practicable.

Thank you for the opportunity to comment on this important supplemental proposal.

Sincerely,

Loren Raun, Ph.D
Mayor's Office of Health and Environmental Policy

cc: Richard Greene, EPA Administrator Region 6