TABLE C.6 (Cont'd)

Blood-Lead Level (μg/dL)	Functional Form	Parameters of Normal on LN ($\Delta_{\overline{IQ}}$) Distribution		Moments of Distribution Over $\Delta_{\overline{IQ}}$		r^2 of $\hat{F}$ on F
		Mean	SD[a]	$E[\Delta_{\overline{IQ}}]$[b]	$SD[\Delta_{\overline{IQ}}]$[c]	
Expert I, Higher SES Levels						
15	Lognormal	-1.18001	0.6191	0.3722	0.2544	0.9987
25	Lognormal	0.0637	0.6022	1.2776	0.8448	0.9903
35	Lognormal	0.8495	0.4110	2.5445	1.0915	0.9945
45	Lognormal	1.3198	0.3193	3.9383	1.2902	0.9970
55	Lognormal	1.4719	0.3839	4.6909	1.8692	0.9974
Expert J, Higher SES Levels						
15	Lognormal	0.4542	0.4502	1.7429	0.8262	0.9620
25	Lognormal	1.0069	0.3640	2.9246	1.1007	0.9410
35	Lognormal	1.2982	0.2762	3.8053	1.0713	0.9590
45	Lognormal	1.6180	0.1991	5.1442	1.0344	0.9979
55	Lognormal	1.9489	0.1466	7.0970	1.0463	0.9923
Expert K, Higher SES Levels						
25	Normal	-	-	1.2836	0.3485	0.9322
35	Normal	-	-	1.9953	0.4232	0.9993
45	Normal	-	-	3.5216	0.6461	0.9989
55	Normal	-	-	4.4858	0.7262	0.9992
65	Normal	-	-	6.0070	0.8948	0.9976

TABLE C.7 Summary and Comparison of Judgments and Fitted Functions Concerning Lead-Induced IQ Effects

Expert F, Both SES Groups							
			$\overline{\Delta IQ}$				
			Blood-Lead Level				
Index	$\overline{IQ_0}$	σ	25 μg/dL	35 μg/dL	45 μg/dL	55 μg/dL	65 μg/dL
Encoded Judgments							
Median	100.5	15	.25	.5	.9	1.5	2
50% C.I.[a]	(100.2, 101.3)	--	(--, .62)	(--, 1.2)	(--, 1.5)	(1.1, 1.9)	(1.3, 2.6)
90% C.I.	(100.1, 101.9)	--	(--, .92)	(--, 1.9)	(--, 2)	(--, 3)	(--, 4)
Fitted Representations							
Median	100.9	15	.25	.49	.97	1.4	1.7
50% C.I.	(100.6, 101.2)	--	(.12, .34)	(.36, .67)	(.71, 1.32)	(1.02, 1.91)	(1.28, 2.39)
90% C.I.	(100.1, 101.7)	--	(.11, .52)	(.23, 1.05)	(.45, 2.07)	(.65, 2.99)	(.81, 3.74)
98% C.I.	(99.8, 102)	--	(.08, .72)	(.17, 1.4)	(.33, 2.84)	(.48, 4.10)	(.59, 5.12)

[a]C.I. denotes credible interval.

TABLE C.7 (Cont'd)

Expert C, Lower SES Group								
			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ_0}$	σ	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	94	14	0.5	1.5	3.0	3.5	5.0	7.0
50% C.I.	(93, 96)	--	(.3, 1)	(.7, 2)	(1.3, 4)	(2.5, 5)	(4, 7)	(5, 10)
90% C.I.	(91, 99)	--	(.2, 3)	(.5, 5)	(1, 8.5)	(2, 10)	(3.5, 12)	(4, 15)
Fitted Representations								
Median	94.6	14	0.6	1.4	2.7	3.9	5.7	7.3
50% C.I	(93, 96)	--	(.4, .9)	(1, 2)	(1.9, 3.8)	(3, 5)	(4.6, 7)	(5.9, 9.1)
90% C.I	(91, 98)	--	(.2, 1.8)	(.6, 3.5)	(1.1, 6.3)	(2.1, 7.3)	(3.5, 9.3)	(4.3, 12.3)
98% C.I.	(90, 100)	--	(.1, 2.8)	(.4, 4.9)	(.8, 9)	(1.6, 9.4)	(2.8, 11.4)	(3.5, 15.3)

TABLE C.7 (Cont'd)

Expert G, Higher SES Groups

Index	$\overline{IQ}_0$	σ	$\overline{\Delta IQ}$ Blood-Lead Level 5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	106	14	--	0.5	1.0	2.5	3.5	5.0
50% C.I.	(105, 107)	--	--	(.4, 1)	(.7, 2)	(1.5, 4)	(2.5, 5)	(4, 7.5)
90% C.I.	(103, 109)	--	--	(.3, 3)	(.5, 4)	(1, 8)	(2, 10)	(3.5, 12)
Fitted Representations								
Median	105.6	14	--	0.7	1.2	2.6	3.9	5.8
50% C.I.	(104, 107)	--	--	(.5, 1)	(.9, 1.7)	(1.9, 3.6)	(3, 5)	(4.7, 7.1)
90% C.I.	(102, 109)	--	--	(.3, 1.8)	(.5, 2.8)	(1.2, 5.8)	(2.1, 7.3)	(3.5, 9.5)
98% C.I.	(101, 111)	--	--	(.2, 2.7)	(.4, 4)	(.8, 8)	(1.6, 9.4)	(2.8, 11.7)

TABLE C.7 (Cont'd)

Expert H, Lower SES Group								
			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ_0}$	σ	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
			Encoded Judgments					
Median	98	13	2.6	3.0	5.2	8.2	9.5	10.5
50% C.I.	(95, 100)	(12,14)	(1.6, 3.4)	(2.6, 5)	(3.4, 7.2)	(6, 9.7)	(7.4, 11.3)	(9.7, 11.8)
90% C.I.	(92, 102)	(10, 15)	(1.1, 4.5)	(1.9, 6.1)	(2.4, 9.2)	(4.2, 12)	(6.2, 14)	(8.2, 16)
			Fitted Representations					
Median	97.2	12.4	2.3	3.5	4.9	7.6	9.4	11.1
50% C.I.	(96, 99)	(12, 13)	(1.7, 3)	(2.7, 4.6)	(3.8, 6.3)	(6.2, 9.4)	(8,11)	(9.7, 12.6)
90% C.I.	(93, 100)	(11, 14)	(1.2, 4.5)	(1.8, 6.7)	(2.6, 9)	(4.6, 12.7)	(6.4, 13.9)	(8.1, 15.1)
98% C.I.	(92, 103)	(10, 15)	(.9, 5.9)	(1.4, 8.8)	(2, 11.7)	(3.7, 15.7)	(5.4, 16.3)	(7.1, 17.2)

TABLE C.7 (Cont'd)

Expert H, Higher SES Group								
			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ_0}$	σ	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	106	14.3	0.6	2.5	4.5	5.5	6.5	8.0
50% C.I.	(104, 108)	(13.8, 14.7)	(--, .9)	(1.8, 3.5)	(3.8, 5.6)	(4.4, 6.4)	(6.1, 7.1)	(7.4, 8.7)
90% C.I.	(101, 110)	(13.2, 15.2)	(--, 1.3)	(1.2, 4.1)	(2.7, 6.7)	(3.9, 7.8)	(5.3, 8.4)	(6.4, 9.6)
Fitted Representations								
Median	105.6	14.3	0.6	2.4	4.4	5.4	6.6	7.9
50% C.I.	(104, 108)	(13.8, 14.6)	(.4, .8)	(1.8, 3.3)	(3.7, 5.2)	(4.7, 6.2)	(6, 7.2)	(7.3, 8.5)
90% C.I.	(101, 110)	(13.3, 15.2)	(.2, 1.3)	(1.2, 5.1)	(2.9, 6.7)	(3.9, 7.5)	(5.3, 8.2)	(6.6, 9.5)
98% C.I.	(99, 112)	(12.9, 15.6)	(.2, 1.9)	(.9, 7)	(2.4, 8)	(3.4, 8.5)	(4.9, 8.9)	(6.1, 10.2)

TABLE C.7 (Cont'd)

Expert I, Lower SES Groups

			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ_0}$	σ	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	--	--	--	0.7	2.3	3.6	5.3	6.3
50% C.I.	--	--	--	(0.5, 1.1)	(1.9, 3.2)	(3.1, 5.1)	(4.5, 6.9)	(5.5, 8.2)
90% C.I.	--	--	--	(0, 2)	(1, 5)	(2, 7)	(3, 10)	(4, 12)
Fitted Representations								
Median	--	--	--	.7	2.3	3.8	5.5	6.7
50% C.I.	--	--	--	(.5, 1.0)	(1.7, 3.1)	(3.0, 4.7)	(4.4, 6.7)	(5.5, 8.1)
90% C.I.	--	--	--	(.3, 1.7)	(1.2, 4.6)	(2.2, 6.5)	(3.3, 9.1)	(4.2, 10.7)
98% C.I.	--	--	--	(.2, 2.5)	(.9, 6.1)	(1.8, 8.2)	(2.7, 11.3)	(3.5, 13.1)

TABLE C.7 (Cont'd)

Expert I, Higher SES Group								
			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ}_0$	σ	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	--	--	--	0.3	1.0	2.3	3.6	4.3
50% C.I.	--	--	--	(.2, .5)	(.7, 2)	(1.9, 3.2)	(3.1, 4.7)	(3.5, 5.8)
90% C.I.	--	--	--	(0, 1)	(0, 3)	(1, 5)	(2, 7)	(2, 9)
Fitted Representations								
Median	--	--	--	0.4	1.1	2.4	3.8	4.5
50% C.I.	--	--	--	(.2, .5)	(.6, 1.7)	(1.8, 3.2)	(3.1, 4.7)	(3.4, 5.7)
90% C.I.	--	--	--	(.0, .9)	(.1, 2.8)	(1.2, 4.5)	(2.2, 6.3)	(2.3, 8.0)
98% C.I.	--	--	--	(0.7, 1.3)	(.3, 4.3)	(.9, 6.1)	(1.8, 7.9)	(1.8, 10.6)

TABLE C.7 (Cont'd)

Expert J, Lower SES Group

Index	$\overline{IQ}_0$	σ	$\overline{\Delta IQ}$ Blood-Lead Level 5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	95	13	2	3.5	5	7	9	11
50% C.I.	(94, 96)	(12, 14)	(1, 3)	(2.3, 4.5)	(4, 5.8)	(6, 8)	(7.5, 10.5)	(9, 12.5)
90% C.I.	(91, 99)	(11, 15)	(.2, 5)	(.4, 6.4)	(1, 8)	(3, 9.6)	(5, 11.7)	(6.6, 13.7)
Fitted Representations								
Median	94.8	13	2.4	3.5	4.8	6.8	8.8	10.7
50% C.I.	(94, 96)	(13, 14)	(1.8, 3.0)	(2.7, 4.3)	(3.8, 5.8)	(5.8, 7.8)	(7.8, 9.9)	(9.7, 11.8)
90% C.I.	(92, 98	(12, 14)	(0.9, 3.9)	(1.5, 5.4)	(2.3, 7.2)	(4.4, 9.3)	(6.3, 11.4)	(8.1, 13.4)
98% C.I.	(90, 99)	(12, 14)	(0.3, 4.5)	(0.7, 6.2)	(1.3, 8.2)	(3.3, 10.3)	(5.2, 12.5)	(7.0, 14.5)

TABLE C.7 (Cont'd)

Expert J, Higher SES Group

Index	$\overline{IQ}_0$	σ	$\overline{\Delta IQ}$ Blood-Lead Level 5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL
Encoded Judgments								
Median	94	14	--	2	3	4	5	7
50% C.I	(93, 96)	--	--	(1.3, 2.3)	(2.8, 3.5)	(3.5, 5)	(4.5, 5.8)	(6.5, 7.5)
90% C.I.	(91, 99)	--	--	(.7, 3)	(1.4, 6.4)	(2, 6.7)	(3.3, 7.6)	(5.3, 9.7)
Fitted Representations								
Median	94.6	14	--	1.6	2.9	3.8	5.0	7.0
50% C.I.	(93, 96)	--	--	(1.2, 2.2)	(2.2, 3.8)	(3.1, 4.7)	(4.4, 5.8)	(6.4, 7.8)
90% C.I.	(91, 98)	--	--	(.7, 3.4)	(1.5, 5.7)	(2.2, 6.4)	(3.6, 7.0)	(5.5, 8.9)
98% C.I.	(90, 100)	--	--	(.5, 4.6)	(1.1, 7.6)	(1.8, 8.0)	(3.2, 8.0)	(5.0, 9.9)

TABLE C.7 (Cont'd)

Expert K, Lower SES Group

			$\overline{\Delta IQ}$					
			Blood-Lead Level					
Index	$\overline{IQ}_0$	σ	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL	65 µg/dL
Encoded Judgments								
Median	85	13	1.5	3.1	4	6	8.1	10
50% C.I.	(83, 88)	(12.2, 13.5)	(0.8, 1.7)	(2, 3.8)	(3.2, 4.7)	(5.4, 7.5)	(6.8, 8.6)	(8.8, 11.8)
90% C.I.	(79, 90.4)	(11, 14)	(0.3, 2)	(1, 4.8)	(2, 5.8)	(4.5, 8.8)	(5.5, 9.8)	(7.5, 13.7)
Fitted Representations								
Median	85	12.6	1.3	2.9	3.8	6.4	7.8	10.4
50% C.I.	(83, 87)	(11.9, 13.2)	(0.7, 1.6)	(2.1, 3.6)	(3, 4.5)	(5.6, 7.3)	(6.9, 8.6)	(9.3, 11.4)
90% C.I.	(80, 90)	(11, 14.1)	(0.5, 1.8)	(1.1, 4.7)	(2, 5.5)	(4.4, 8.4)	(5.6, 9.9)	(7.8, 13)
98% C.I.	(78, 92)	(10.4, 14.7)	(0.2, 2.3)	(0.3, 5.5)	(1.3, 6.3)	(3.6, 9.3)	(4.7, 10.8)	(6.7, 14)

TABLE C.7 (Cont'd)

Expert K, Higher SES Group

Index	$\overline{IQ_o}$	σ	$\overline{\Delta IQ}$ Blood-Lead Level 25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL	65 µg/dL
Encoded Judgments							
Median	104.9	13.5	1.5	2	3.5	4.5	6
50% C.I.	(102.6, 107.2)	(13, 14)	(.8, 1.7)	(1.6, 2.3)	(3.1, 4)	(4, 4.8)	(5.2, 6.8)
90% C.I.	(100.6, 109.4)	(12.2, 14.8)	(.6, 1.9)	(1.2, 2.9)	(2.2, 4.8)	(3.1, 5.9)	(4.3, 7.8)
Fitted Representations							
Median	104.9	13.5	1.3	2	3.5	4.5	6
50% C.I.	(103.3, 106.5)	(13, 14)	(1, 1.5)	(1.7, 2.3)	(3.1, 4)	(4, 5)	(5.4, 6.6)
90% C.I.	(101, 108.8)	(12.3, 14.6)	(.7, 1.9)	(1.3, 2.7)	(2.5, 4.6)	(3.3, 5.7)	(4.5, 7.5)
98% C.I.	(99.4, 110.4)	(11.9, 15.1)	(.5, 2.1)	(1, 3)	(2, 5)	(2.8, 6.2)	(3.9, 8.1)

Thus, Expert F judged a probability of 0.5 that the mean IQ of the unexposed group would be $\leq$ 100.5, 0.8 probability that it would be $\leq$ 101.5, and so on. These and other judgments about mean IQ of an unexposed group of children are listed in Table C.1. Expert F provided a point value of 15.25 for σ, which appears in Table C.2.

Expert F provided judgments on IQ differences at five PbB levels: 25, 35 ,45, 55, and 65 µg/dL. These judgments are listed in Table C.3. For example, at a PbB level of 65 µg/dL, Expert F judged a probability of 0.9 that IQ difference would be $\leq$ 3. Figure 4.3 shows that Expert F is increasingly uncertain about the magnitude of IQ differences as PbB level increases. Table C.1 indicates that judgments on IQ differences having cumulative probabilities less than 0.5 were not provided at the lower three PbB levels. Expert F cited difficulty in providing judgments on what would be very small IQ differences.

For the reasons presented in Sec. 1, mathematical functions were fit to the judgments of Expert F. For Expert F, the mathematical representations of the CDFs for IQ differences conditional on PbB level are guaranteed never to cross. They were obtained by fitting regression lines to transformations of the judgments. The transformation used was the log (ln) transformation. Similar to the hemoglobin case, this transformation leads to distributions that are normal distributions, with a high degree of accuracy.

Thus, the fitted distributions are lognormal distributions. They are uniquely defined by the mean and variance of the underlying normal distributions, and the transformation function. Properties of the lognormal distribution are summarized in Table B.1. A variety of distributions were fit to the judgments, and the ones reported here were determined (with input from the experts) to be best.

Tables C.4-C.6 summarize relevant information about the distributions fit to the judgments of the experts. In Table C.4, columns identify the functional form, mean and standard deviation of an underlying normal distribution (if a transformation is used),

mean ($E[\overline{IQ}_o]$) and standard deviation ($SD[\sigma_{IQ}]$) of the IQ distribution, and r^2 of the regression of fitted cumulative probability values $\hat{F}$ versus encoded values (F).

Similar column headings appear in Tables C.5 and C.6. In Table C.5, $E[\sigma_{IQ}]$ denotes the mean, and $SD[\sigma_{IQ}]$ denotes the standard deviation of the distribution for population standard deviation. L denotes PbB level, and $E[\Delta_{\overline{IQ}}]$ and $SD[\Delta_{\overline{IQ}}]$ denote the means and standard deviations of probability distributions for IQ decrements in Table C.6.

For Expert F, the fitted distribution of the uncertain mean IQ of the unexposed group has a mean value of. 100.9; it is a lognormal distribution and the r^2 value of the regression of the $\hat{F}$ values vs. the F values is 0.93. The standard deviation of the distribution is 0.47. The mean and standard deviation of the underlying normal distribution (which results after taking the natural log of the encoded mean IQ values) are 4.6 and 0.0046. These values are needed to calculate points on a cumulative probability function for the mean IQ variable. For population standard deviation, a point estimate of 15 was given.

Lognormal distributions provided the best fits of Expert F's judgments about IQ differences. Thus, all of the columns have entries. The mean values of the distributions over IQ decrement vary from about 0.3 at a PbB level of 25 μg/dL to about 1.9 at 65 μg/dL. In general, the fitted functions represent these judgments quite well, having r^2 values around 0.99 or higher.

C.3.2 Expert G

Expert G provided probabilistic judgments about mean IQ for an unexposed group and IQ decrements for populations exposed to varying levels of lead for both SES categories. A point estimate for population standard deviation (σ_{IQ}=14) was made for both SES catetories. Lognormal distributions provided the best fits for the rest of the judgments. The mean values of the IQ distributions were 94.6 and 105.6 for the lower

and higher SES cases, respectively. Averaging across SES level results in an average value of about 104 for the mean IQ for unexposed children. Mean values of the distributions for IQ decrements range from 0.8 (at 5 μg/dL) to 7.7 (at 55 μg/dL) for the lower SES category, and from 0.5 (at 15 μg/dL) to 6.0 (at 55 μg/dL) for the higher SES category. The r^2 values of the regressions of the fitted distributions versus the encoded distributions were quite good. The lowest value was 0.95.

C.3.3 Expert H

This expert provided probabilistic judgments about all of the variables for both SES categories. Lognormal distributions again provided excellent fits to all of the encoded judgments. The mean values of the fitted IQ distributions were 97.2 and 105.6, and the mean values of the fitted standard deviation distributions were 12.3 and 14.2. The standard deviations of the fitted IQ distributions were 2.4 and 2.8. Mean values of the fitted distributions for IQ decrements ranged from 2.5 (at 5 μg/dL) to 11.3 (at 55 μg/dL), and from 0.6 (at 5 μg/dL) to 8 (at 55 μg/dL). Again, the r^2 values of the regressions are quite good (0.98 or higher). (In all cases, the values for the lower SES cases are first, and the values for the higher SES cases are second.)

C.3.4 Expert I

Expert I only provided probabilistic judgments for IQ decrement. This individual was not comfortable about providing judgments, reasonding that it should be possible to solve for the other variables, given what is currently known and published about population IQ distributions. However, we found that it was not possible to do this without making strong assumptions about Expert I's probability judgments. In particular, we would either have to assume that judgments of IQ effects at each lead level are independent of each other, or assume a particular pattern of dependencies. Since we were certain that an assumption of independence would be wrong, and had no basis for any other specific assumptions, we simply accepted the judgments as provided.

Lognormal distributions provided the best fits to these judgments. Judgments were provided at five PbB levels (15-55 μg/dL). Mean values of the fitted distributions on IQ decrement ranged from 0.8 to 7 for the lower SES group, and from 0.4 to 4.7 for the higher SES group. The r^2 values for the regressions of the fitted distributions ($\hat{F}$ values) vs. the encoded distributions (F values) were all greater than 0.99.

C.3.5 Expert J

Expert J provided probabilistic judgments on all three IQ variables. Lognormal distributions were used to represent these judgments. The mean values of the fitted IQ distributions were 94.7 and 104.1, and the mean values of the fitted standard deviation distributions were 13 and 12.9. Mean values of the distributions fitted to the judgments on IQ decrements ranged from 2.4 to 10.7 for the lower SES case, and from 1.7 to 7.1 for the higher SES case. The r^2 values for regressions of fitted values versus encoded values were generally around 0.99 for the lower SES case, and slightly less than 0.99 for the higher SES case.

C.3.6 Expert K

Expert K provided probabilistic judgments on all three IQ variables. Normal distributions provided the best fits to all of the judgments. The mean values of the fitted IQ distributions were 85.0 and 104.9, and the mean values of the fit standard deviation distributions were 12.6 and 13.5. Mean values of the distributions fit to the judgments on IQ decrements ranged from 1.3 to 10.4 for the lower SES case, and from 1.3 to 6 for the higher SES case. The r^2 values for regressions of fitted values versus encoded values were generally around 0.99 for both of the SES cases.

C.3.7 Comparison of Judgments and Fitted Functions

Table C.7 summarizes and compares encoded judgments and corresponding points on the functions fit to the judgments of each of the experts. The table can be further used to compare and contrast the judgments of the different experts. Many tables like this one (one per case per expert) were used to help the experts understand the implications of their judgments and choose mathematical functions that best represented those judgments.

The main point of the table is that the fitted functions match the judgments quite well. The fitted median values generally are quite close to the assessed median values. This is also the case for the 50% and 90% credible intervals. For example, consider the judgments of Expert G. The median judgment for the mean IQ level for unexposed children in the lower SES category is 94, and the median of the fitted distribution is 94.6. The encoded medians for IQ decrement are 0.5, 1.5, 3.0, 3.5, 5.0, and 7.0 for PbB levels of 5, 15, 25, 35, 45, and 55 μg/dL, respectively. The respective medians of the fitted distributions are 0.6, 1.4, 2.7, 3.9, 5.7, and 7.3. The lower ends of the 50% credible intervals (the 0.25 points on the cumulative probability curves) are 0.3, 0.7, 1.3, 2.5, 4, and 5 for the encoded judgments, and 0.4, 1, 1.9, 3, 4.6, and 5.9 for the fitted functions. The upper ends of the 50% credible intervals (the 0.75 points on the cumulative probability curves) are 1, 2, 4, 5, 7, and 10 for the judgments, and 0.9, 2, 3, 8, 5, 7, and 9.1 for the fitted functions.

Expert F judged that the mean IQ of children sheltered from lead should be around 101 with a high degree of certainty (indicated by a standard deviation of about 0.5). The other experts judged that the mean IQ would be one to three points higher, on the average, but with much less certainty as indicated by standard deviations for the mean IQ distribution that ranged from 1.5 to 3.0. Similarly, Expert F judged much smaller IQ decrements attributable to lead exposure than did the other experts. First judgments were at PbB levels of 25 μg/dL for Expert F, 15 μg/dL for Expert K, and 5 μg/dL for the other experts.

While these judgments are interesting in and of themselves, they take on added meaning when included in a health risk assessment. Thus, the results of Sec. 5 provide further insight into the significance of the differences in judgments among the experts.

C.4 SUMMARY DISCUSSIONS

C.4.1 Summary of Discussions with Expert F

1. Low lead levels (below 30 μg/dL or 40 μg/dL) do not have a discernable, deleterious effect on IQ. Studies that purport to show such an effect have methodological and bias problems.
2. Although low lead levels do not have independent effects on IQ, one cannot currently rule out the possibility that lead at higher levels interacts with other variables to affect IQ. Nor can we rule out completely the possibility that IQ influences lead level. These variables could relate to general health status, mental health status, or other environmental factors.
3. Numerous covariates are associated with lead exposure, many of which are known to be negatively related to IQ. It is difficult to use regression techniques to determine which of the variables are "causative."
4. Behavioral, social, and cognitive measures have some degrees of unreliability that add uncertainty to any conclusions.
5. Cord blood lead may be related to a few negative birth outcomes, but one cannot infer direction of causality from such data. Perhaps a distressed fetus accumulates lead.
6. There is a fairly strong correlation (around 0.8) between maternal blood lead and cord blood lead.

7. Children are robust and can recover from minor disruptions in their cognitive development. It is not clear to what degree they can recover from larger disruptions. Middle-class children, in particular, might be able to compensate.
8. The available evidence does not indicate what is the threshold level for IQ effects of lead.
9. The assumption of normal IQ distributions within exposure groups in our hypothetical experiment is reasonable, except that at high lead levels, the IQ distribution may be skewed negatively, so that the percentage of children below a specified IQ level would be underestimated.

C.4.2 Summary of Discussions with Expert G

1. IQ effects of lead on lower SES children may be less in European countries than in the U.S., because the educational and health disparities between the SES levels are probably less in these countries than in the U.S.
2. There are no data to suggest that physiological response to lead varies as a function of SES level.
3. The effects of lead on the developing central nervous system are probably long lasting and not reversible.
4. If there is a threshold for IQ effects of lead, it is very low.
5. Animal models for risk assessment are limited in usefulness, because of the difficulty in extrapolating to humans in a quantitative manner. Their importance is in studying mechanisms of lead effects and other important topics (e.g., reversibility of effects and vulnerable periods of development), not in

establishing dose-response relationships directly applicable to humans.

6. Specific neurobehavioral effects of lead, such as attentional deficits, should not be overlooked.
7. The assumption of normally distributed IQ scores within exposure groups in our hypothetical experiment is reasonable.

C.4.3 Summary of Discussions with Expert H

1. Lead interacts with other factors in its effect on IQ. These factors include SES, maternal stimulation, and nutrition.
2. The concept of critical developmental periods is crucial in understanding the effects of lead. When opportunities for intellectual development are lost, it is very hard to compensate later. This is especially true in lower SES children. Furthermore, the effect can snowball, in that skills generally build on each other.
3. Children who are exposed to lead also tend to get more cadmium, pesticides, etc. This is especially true among lower SES children.
4. There are no data on the minimum exposure time necessary to produce an effect.
5. At a given dose level, the cumulative are worse than the immediate effects. Exposures earlier in life have more impact than do exposures later in life.
6. Some lead effects are reversible and others are not.
7. The half-life of lead in the brain is very long, based on rat studies by Gary Goldstein at UCLA.

8. Some recent research suggests that lead may be involved in senility. Specifically, the amount of bone lead goes down with aging. The question is, where does the lead go? Is it eliminated, or is it recirculated (in which case it may have neurobehavioral effects)?
9. It is reasonable to assume that IQ scores are normally distributed within exposure groups in our hypothetical experiment.
10. There is important and relevant work by Bernard Weiss and by Ketz that involves behavioral toxicology.

C.4.4 Summary of Discussions with Expert I

1. The meta-analyses occasionally done with the studies on IQ effects of lead are inappropriate because the studies differ in so many ways, including different ages of the subjects, different sorts of control groups, different methods, and different exposure conditions.
2. It is reasonably certain that there is an IQ effect of lead above 30 µg/dL, but less clear about whether there is an effect at lower levels.
3. There is an exposure-sensitivity interaction, in that children who suffer the greatest exposures tend to be those who are most susceptible to the effects of lead.
4. Except for high acute exposures, neurobehavioral effects are more likely to result from chronic than from acute exposures.
5. One should distinguish between biological and functional effects of lead, but in either case there are no data suggesting that the effects are necessarily irreversible.

6. The data are not clear either way with regard to the question of whether there is a threshold for the IQ effects of lead.

C.4.5 Summary of Discussions with Expert J

1. Lead does have a deleterious effect on neurobehavioral development. In this regard, lead interacts with diet and SES factors.
2. Middle-class children are somewhat buffered against the negative effects of lead by virtue of their better diet, general health status, and richer intellectual resources.
3. Lead crosses the blood-brain barrier, and once in the cells stays there. This lead cannot be chelated out and affects the neurotransmitters. The effects are irreversible.
4. Research with animals suggests that lead causes demyelination. It is not clear whether this effect would occur in children.
5. The implications of the EEG data are not clear because there is uncertainty as to what the measures refer and because findings are somewhat inconsistent.
6. Although some functional effects of lead may be reversible, morphological effects are not.
7. There is probably no threshold for lead effects; if there is one, it is very low. The German studies on primates suggest the absence of a threshold.
8. There may be racial differences in terms of the effects of lead. However, this is a very complicated issue that remains to be untangled.
9. The assumption that IQ scores are normally distributed within

exposure groups in our hypothetical experiment is reasonable with the possible exception of the 55-μg/dL group.

C.4.6 Summary of Discussions with Expert K

1. Children of all genotypes are vulnerable to the effects of lead.
2. Some neurobehavioral effects of lead, such as the effect on attention span, are reversible. These are generally associated with low levels of exposure. Even some perinatal effects can be reversed given proper treatment.
3. Other neurobehavioral effects are irreversible. These include morphological changes, as well as effects on some behaviors that depend on critical developmental periods. (Is there evidence or theory that allows specificity as to which behaviors might be irreversibly affected, or is this to be taken as a generalized statement that some behaviors might be irreversibly altered?)
4. Lead in the range of 5-55 μg/dL has a deleterious effect in the presence of other factors. In that sense, lead interacts with other factors. However, the interaction is only in one direction, so that a usual MANOVA (multivariate analysis of variance) would show both an interaction and a main effect.
5. Some of the factors that lead interacts with are physical and psychological hygiene, physiological status, parental IQ, nutritional and health status, organophosphates, and pesticides.
6. Early, long-term exposure to lead can lead to long-term effects, while eposodic exposure is more likely to lead to reversible effects, if the exposure is not too great.
7. There is a threshold for IQ effects of lead.

8. The assumption that IQ scores are normally distributed within exposure groups in the hypothetical experiment is reasonable.

APPENDIX D

RISK ESTIMATES

This appendix consists mainly of tables that list the functions and the parameters that describe the risk estimates for lead-induced EP, hemoglobin, and IQ health effects. Some discussion is also provided to explain the formats of the tables. The tabular material can be used to generate data for the figures presented and discussed in Sec. 5, and to perform additional analyses.

PbB distributions are assumed to be lognormal with geometric standard deviation (GSD) 1.42 and the geometric mean (GM) values listed in Table D.1. GM values are tabulated for six air-lead scenarios at each of 5 point sources of lead emissions. The six air-lead scenarios result from 3 time periods (baseline, precontrol, and postcontrol periods), each of which have upper- and lower-bound estimates for the GM.

D.1 ESTIMATES FOR THE RISKS OF LEAD-INDUCED ERYTHROCYTE PROTOPORPHYRIN EFFECTS AMONG U.S. CHILDREN AGED 0-6 YEARS

Probabilistic dose-response functions for 2 EP levels considered to be adverse (33 μλÅφ↓ αvφ ^{53} μg/dL) were considered in Sec. 2 and App. A. We calculated the overall response rates for populations of children by combining the dose-response functions (which are specified at particular PbB levels) with PbB distributions. In the simplest of cases, the probabilistic dose-response functions were normal probability distributions and the overall risk result was also normal with mean

$$\mu = \sum P_i \mu_i$$

and standard deviation

$$\sigma = \left(\sum P_i^2 \sigma_i^2\right)^{1/2}$$

TABLE D.1 Upper- and Lower-Bound Estimates of the Geometric Mean Blood-Lead Levels Children Living around Selected U.S. Sources of Lead Emissions[a]

	Source 1		Source 2		Source 3		Source 4		Source 5	
Age	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Baseline Period										
0-3 yr	19.1	26.6	19.9	27.0	18.7	26.2	18.7	26.2	8.6	13.9
4-6 yr	14.0	20.8	14.3	21.2	13.7	20.4	13.7	20.4	6.4	11.5
0-6 yr	16.9	24.1	17.2	24.5	16.6	23.7	16.6	23.7	7.7	12.9
36th month	18.8	26.7	19.2	27.1	18.4	26.2	18.5	26.3	7.7	13.4
Precontrol Period										
0-3 yr	4.8	9.5	5.2	9.9	4.7	9.1	4.6	8.6	3.3	6.7
4-6 yr	2.2	5.4	2.4	5.9	2.1	5.0	2.1	4.8	2.0	5.5
0-6 yr	3.7	7.7	4.0	8.2	3.6	7.3	3.5	7.0	2.7	6.2
36th month	3.8	9.1	4.4	9.7	3.6	8.6	3.8	8.2	2.9	6.5
Postcontrol Period										
0-3 yr	1.4	3.4	1.4	3.6	1.4	3.4	1.4	3.1	1.4	3.7
4-6 yr	1.0	3.2	1.1	3.4	1.0	3.1	1.0	2.8	1.1	3.6
0-6 yr	1.2	3.3	1.3	3.7	1.3	3.4	1.2	3.0	1.3	3.7
36th month	1.3	3.4	1.3	3.7	1.3	3.3	1.2	3.1	1.3	3.7

[a]The PbB distribution is assumed to be lognormal with GSD = 1.42 and the tabulated GM values.

where i denotes a PbB level, P_i is the probability of PbB level i, and μ_i and σ_i are the mean and variance of the EP distribution at PbB level i.

These calculations implicitly assume statistical independence, a necessary assumption that we were unable to verify. In some cases, it was necessary to use beta distributions in the calculations. Fortunately, in those cases the final result was nearly identical to the dose-response function at a PbB of 5 μg/dL. This occurred in all of the postcontrol scenarios and most of the precontrol scenarios because modeled PbB levels were quite low, (<10 μg/dL).

Risk results for EP are tabulated in Table D.2. In the table, the form (either normal or beta), the expected value, and the standard deviation of each risk distribution are listed for the two critical EP levels specified above. When the form of the distribution is beta, the values of the parameters are always a = 3 and b = 128, which result in a mean of 2.3% and a standard deviation of 1.3%.

D.3 ESTIMATES OF THE RISKS OF LEAD-INDUCED HEMOGLOBIN EFFECTS AMONG U.S. CHILDREN AGED 0-3 YEARS

Calculation of risk estimates for lead-induced hemoglobin effects is in principle identical to that for EP risks. The only difference is that normal-on-log-odds (NOLO) distributions were found to best represent the judgments of the experts we consulted, so a computation analogous to a convolution integral had to be employed. Numerical methods were used to perform the calculations, and it was necessary to assume statistical independence.

Results are listed in Table D.3. Included are the form (normal or NOLO), parameters, mean, and standard deviation of the risk distributions. All of these items are provided for NOLO distributions for the reasons discussed in App. B. The parameters of the normal distribution are the mean and standard deviation, so only the latter are listed for normal distributions. The functions were determined applying linear regression

TABLE D.2 Specification of Distributions for the Risks of Elevated EP Levels among Children Ages 0-6 Years

	EP* = 33 µg/dL			EP* = 53 µg/dL		
Scenario	Form	Mean	S.D.[a]	Form	Mean	S.D.
			Source 1			
Lower Base	Normal	20.10	0.82	Normal	7.52	0.52
Upper Base	Normal	38.69	1.24	Normal	21.14	1.15
Lower Pre	Normal	10.70	2.69	Beta[b]	2.34	1.30
Upper Pre	Normal	10.78	2.09	Normal	2.43	1.04
Lower Post	Normal	10.70	2.70	Beta	2.34	1.30
Upper Post	Normal	10.70	2.70	Beta	2.34	1.30
			Source 2			
Lower Base	Normal	20.75	0.83	Normal	7.92	0.54
Upper Base	Normal	39.78	1.26	Normal	22.07	1.19
Lower Pre	Normal	10.70	2.69	Beta	2.34	1.30
Upper Pre	Normal	10.83	1.95	Normal	2.46	0.96
Lower Post	Normal	10.70	2.70	Beta	2.34	1.30
Upper Post	Normal	10.70	2.69	Beta	2.34	1.30
			Source 3			
Lower Base	Normal	19.46	0.81	Normal	7.13	0.50
Upper Base	Normal	37.60	1.22	Normal	20.22	1.11
Lower Pre	Normal	10.70	2.70	Beta	2.34	1.30
Upper Pre	Normal	10.75	2.21	Beta	2.34	1.30
Lower Post	Normal	10.70	2.70	Beta	2.34	1.30
Upper Post	Normal	10.70	2.70	Beta	2.34	1.30
			Source 4			
Lower Base	Normal	19.46	0.81	Normal	7.13	0.50
Upper Base	Normal	37.60	1.22	Normal	20.22	1.11
Lower Pre	Normal	10.70	2.70	Beta	2.34	1.30
Upper Pre	Normal	10.73	2.29	Beta	2.34	1.30
Lower Post	Normal	10.70	2.70	Beta	2.34	1.30
Upper Post	Normal	10.70	2.70	Beta	2.34	1.30

TABLE D.2 (Cont'd)

Scenario	EP* = 33 µg/dL Form	Mean	S.D.[a]	EP* = 53 µg/dL Form	Mean	S.D.
			Source 5			
Lower Base	Normal	10.78	2.09	Normal	2.43	1.04
Upper Base	Normal	13.45	0.89	Normal	3.74	0.45
Lower Pre	Normal	10.70	2.70	Beta	2.34	1.30
Upper Pre	Normal	10.71	2.47	Beta	2.34	1.30
Lower Post	Normal	10.70	2.70	Beta	2.34	1.30
Upper Post	Normal	10.70	2.69	Beta	2.34	1.30

[a]Standard Deviation.

[b]For every beta distribution, the parameters values are a = 3 and b = 128, which results in the tabulated mean and S.D. values.

TABLE D.3 Summary of Hemoglobin Risk Distributions for Children Aged 0-3 Years

Source 1, Hb* = 9.5 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-3.1800	0.1567	0.0404	0.0061	1.0000
	D	-3.6945	0.1671	0.0246	0.0040	1.0000
	E	-2.0239	0.4752	0.1252	0.0565	1.0000
UPPER BASE	C	-2.9741	0.1433	0.0490	0.0067	1.0000
	D	-2.9266	0.1879	0.0516	0.0093	1.0000
	E	-1.7214	0.4283	0.1594	0.0615	1.0000
LOWER PRE	C	-3.9254	0.2237	0.0198	0.0044	1.0000
	D	-8.7457	0.3518	0.0002	0.0001	1.0000
	E	-3.3975	0.7363	0.0408	0.0342	1.0000
UPPER PRE	C	-3.6385	0.1770	0.0260	0.0045	1.0000
	D	-4.4620	0.2845	0.0119	0.0034	1.0000
	E	-2.7415	0.5559	0.0683	0.0392	1.0000
LOWER POST	C	-3.9995	0.2764	0.0187	0.0052	1.0000
	E	-3.7820	0.9600	0.0331	0.0414	1.0000
UPPER POST	C	-4.0004	0.2764	0.0186	0.0052	1.0000
	E	-3.7829	0.9600	0.0330	0.0413	1.0000

TABLE D.3 (Cont'd)

Source 2, Hb* = 9.5 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-3.1718	0.1570	0.0407	0.0062	1.0000
	D	-3.6630	0.1675	0.0253	0.0042	1.0000
	E	-2.0128	0.4741	0.1264	0.0568	1.0000
UPPER BASE	C	-2.9671	0.1444	0.0494	0.0068	1.0000
	D	-2.8938	0.1890	0.0533	0.0096	1.0000
	E	-1.7078	0.4290	0.1612	0.0622	1.0000
LOWER PRE	C	-3.9172	0.2236	0.0200	0.0044	1.0000
	D	-8.1926	0.3519	0.0003	0.0001	1.0000
	E	-3.3824	0.7357	0.0414	0.0346	1.0000
UPPER PRE	C	-3.6118	0.1802	0.0267	0.0047	1.0000
	D	-4.4551	0.2844	0.0119	0.0034	1.0000
	E	-2.6880	0.5600	0.0718	0.0414	1.0000
LOWER POST	C	-3.9995	0.2764	0.0187	0.0052	1.0000
	E	-3.7820	0.9600	0.0331	0.0414	1.0000
UPPER POST	C	-4.0010	0.2764	0.0186	0.0052	1.0000
	E	-3.7836	0.9600	0.0330	0.0413	1.0000

TABLE D.3 (Cont'd)

Source 3, Hb* = 9.5 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-3.1912	0.1568	0.0399	0.0061	1.0000
	D	-3.7376	0.1669	0.0236	0.0039	1.0000
	E	-2.0373	0.4761	0.1238	0.0561	1.0000
UPPER BASE	C	-2.9823	0.1430	0.0487	0.0067	1.0000
	D	-2.9596	0.1868	0.0500	0.0090	1.0000
	E	-1.7337	0.4276	0.1577	0.0609	1.0000
LOWER PRE	C	-3.9270	0.2237	0.0198	0.0044	1.0000
	D	-8.8980	0.3518	0.0001	0.0001	1.0000
	E	-3.4004	0.7365	0.0407	0.0341	1.0000
UPPER PRE	C	-3.6642	0.1743	0.0253	0.0043	1.0000
	D	-4.4677	0.2845	0.0118	0.0034	1.0000
	E	-2.7959	0.5528	0.0649	0.0371	1.0000
LOWER POST	C	-3.9995	0.2764	0.0187	0.0052	1.0000
	E	-3.7820	0.9600	0.0331	0.0414	1.0000
UPPER POST	C	-4.0004	0.2764	0.0186	0.0052	1.0000
	E	-3.7829	0.9600	0.0330	0.0413	1.0000

TABLE D.3 (Cont'd)

Source 4, Hb* = 9.5 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-3.1912	0.1568	0.0399	0.0061	1.0000
	D	-3.7376	0.1669	0.0236	0.0039	1.0000
	E	-2.0373	0.4761	0.1238	0.0561	1.0000
UPPER BASE	C	-2.9823	0.1430	0.0487	0.0067	1.0000
	D	-2.9596	0.1868	0.0500	0.0090	1.0000
	E	-1.7337	0.4276	0.1577	0.0609	1.0000
LOWER PRE	C	-3.9283	0.2237	0.0198	0.0044	1.0000
	D	-9.0534	0.3518	0.0001	0.0000	1.0000
	E	-3.4030	0.7366	0.0406	0.0340	1.0000
UPPER PRE	C	-3.7472	0.2083	0.0235	0.0048	1.0000
	D	-5.7803	0.3530	0.0033	0.0012	1.0000
	E	-3.1071	0.7050	0.0525	0.0412	1.0000
LOWER POST	C	-3.9995	0.2764	0.0187	0.0052	1.0000
	E	-3.7820	0.9600	0.0331	0.0414	1.0000
UPPER POST	C	-3.9999	0.2764	0.0186	0.0052	1.0000
	E	-3.7824	0.9600	0.0331	0.0414	1.0000

TABLE D.3 (Cont'd)

Source 5, Hb* = 9.5 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-3.7472	0.2083	0.0235	0.0048	1.0000
	D	-5.7803	0.3530	0.0033	0.0012	1.0000
	E	-3.1071	0.7050	0.0525	0.0412	1.0000
UPPER BASE	C	-3.3970	0.1981	0.0330	0.0064	1.0000
	D	-4.2326	0.2230	0.0147	0.0033	1.0000
	E	-2.3392	0.5762	0.0985	0.0576	1.0000
LOWER PRE	C	-4.0002	0.2764	0.0186	0.0052	1.0000
	E	-3.7827	0.9600	0.0331	0.0414	1.0000
UPPER PRE	C	-3.8628	0.2186	0.0210	0.0046	1.0000
	D	-6.7643	0.3522	0.0012	0.0004	1.0000
	E	-3.3026	0.7138	0.0440	0.0353	1.0000
LOWER POST	C	-3.9995	0.2764	0.0187	0.0052	1.0000
	E	-3.7820	0.9600	0.0331	0.0414	1.0000
UPPER POST	C	-4.0015	0.2764	0.0186	0.0052	1.0000
	E	-3.7840	0.9599	0.0330	0.0413	1.0000

TABLE D.3 (Cont'd)

Source 1, Hb* = 11.0 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	A	-7.7633	0.6530	0.0005	0.0004	0.9997
	C	-2.0191	0.1550	0.1182	0.0163	1.0000
	D	-1.9726	0.2625	0.1249	0.0294	1.0000
	E	-1.7216	0.3184	0.1561	0.0435	1.0000
UPPER BASE	A	-3.0621	0.5087	0.0498	0.0261	1.0000
	C	-1.7932	0.1409	0.1435	0.0174	1.0000
	D	-1.5492	0.2493	0.1780	0.0373	1.0000
	E	-1.5361	0.2885	0.1809	0.0441	1.0000
LOWER PRE	C	-2.4527	0.2211	0.0807	0.0167	1.0000
	D	-3.0274	0.3888	0.0493	0.0191	1.0000
	E	-2.1708	0.4716	0.1103	0.0501	1.0000
UPPER PRE	C	-2.3090	0.1717	0.0914	0.0144	1.0000
	D	-2.7356	0.2976	0.0631	0.0181	1.0000
	E	-1.9982	0.3587	0.1244	0.0409	1.0000
LOWER POST	C	-2.5162	0.2720	0.0769	0.0198	1.0000
	D	-3.1747	0.4919	0.0445	0.0225	1.0000
	E	-2.3463	0.5991	0.0987	0.0606	1.0000
UPPER POST	C	-2.5171	0.2720	0.0768	0.0198	1.0000
	D	-3.1756	0.4919	0.0445	0.0225	1.0000
	E	-2.3473	0.5991	0.0986	0.0605	1.0000

TABLE D.3 (Cont'd)

Source 2, Hb* = 11.0 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	A	-7.6745	0.6530	0.0006	0.0004	0.9997
	C	-2.0123	0.1551	0.1189	0.0164	1.0000
	D	-1.9528	0.2635	0.1270	0.0299	1.0000
	E	-1.7153	0.3173	0.1568	0.0435	1.0000
UPPER BASE	A	-3.0602	0.5085	0.0499	0.0261	1.0000
	C	-1.7829	0.1415	0.1448	0.0177	1.0000
	D	-1.5344	0.2495	0.1802	0.0377	1.0000
	E	-1.5276	0.2894	0.1821	0.0445	1.0000
LOWER PRE	C	-2.4493	0.2211	0.0810	0.0167	1.0000
	D	-3.0210	0.3887	0.0496	0.0192	1.0000
	E	-2.1681	0.4716	0.1105	0.0502	1.0000
UPPER PRE	C	-2.2977	0.1742	0.0923	0.0147	1.0000
	D	-2.7048	0.2991	0.0650	0.0187	1.0000
	E	-1.9864	0.3598	0.1257	0.0414	1.0000
LOWER POST	C	-2.5162	0.2720	0.0769	0.0198	1.0000
	D	-3.1747	0.4919	0.0445	0.0225	1.0000
	E	-2.3463	0.5991	0.0987	0.0606	1.0000
UPPER POST	C	-2.5178	0.2720	0.0768	0.0197	1.0000
	D	-3.1763	0.4919	0.0445	0.0225	1.0000
	E	-2.3480	0.5990	0.0985	0.0605	1.0000

TABLE D.3 (Cont'd)

Source 3, Hb* = 11.0 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	A	-7.9091	0.6529	0.0005	0.0003	0.9997
	C	-2.0286	0.1550	0.1172	0.0162	1.0000
	D	-1.9969	0.2610	0.1222	0.0287	1.0000
	E	-1.7295	0.3194	0.1551	0.0434	1.0000
UPPER BASE	A	-3.0660	0.5091	0.0497	0.0260	1.0000
	C	-1.8038	0.1405	0.1422	0.0173	1.0000
	D	-1.5649	0.2489	0.1758	0.0369	1.0000
	E	-1.5439	0.2874	0.1797	0.0437	1.0000
LOWER PRE	C	-2.4533	0.2211	0.0807	0.0167	1.0000
	D	-3.0286	0.3888	0.0492	0.0190	1.0000
	E	-2.1713	0.4716	0.1102	0.0500	1.0000
UPPER PRE	C	-2.3206	0.1701	0.0904	0.0141	1.0000
	D	-2.7637	0.2964	0.0615	0.0176	1.0000
	E	-2.0081	0.3583	0.1233	0.0405	1.0000
LOWER POST	C	-2.5162	0.2720	0.0769	0.0198	1.0000
	D	-3.1747	0.4919	0.0445	0.0225	1.0000
	E	-2.3463	0.5991	0.0987	0.0606	1.0000
UPPER POST	C	-2.5171	0.2720	0.0768	0.0198	1.0000
	D	-3.1756	0.4919	0.0445	0.0225	1.0000
	E	-2.3473	0.5991	0.0986	0.0605	1.0000

TABLE D.3 (Cont'd)

Source 4, Hb* = 11.0 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	A	-7.9091	0.6529	0.0005	0.0003	0.9997
	C	-2.0286	0.1550	0.1172	0.0162	1.0000
	D	-1.9969	0.2610	0.1222	0.0287	1.0000
	E	-1.7295	0.3194	0.1551	0.0434	1.0000
UPPER BASE	C	-1.8038	0.1405	0.1422	0.0173	1.0000
	D	-1.5649	0.2489	0.1758	0.0369	1.0000
	E	-1.5439	0.2874	0.1797	0.0437	1.0000
LOWER PRE	C	-2.4539	0.2212	0.0807	0.0167	1.0000
	D	-3.0297	0.3888	0.0492	0.0190	1.0000
	E	-2.1718	0.4716	0.1102	0.0500	1.0000
UPPER PRE	C	-2.3810	0.2065	0.0860	0.0165	1.0000
	D	-2.9024	0.3674	0.0550	0.0199	1.0000
	E	-2.1345	0.4477	0.1130	0.0482	1.0000
LOWER POST	C	-2.5162	0.2720	0.0769	0.0198	1.0000
	D	-3.1747	0.4919	0.0445	0.0225	1.0000
	E	-2.3463	0.5991	0.0987	0.0606	1.0000
UPPER POST	C	-2.5166	0.2720	0.0769	0.0198	1.0000
	D	-3.1751	0.4919	0.0445	0.0225	1.0000
	E	-2.3467	0.5991	0.0986	0.0606	1.0000

TABLE D.3 (Cont'd)

Source 5, Hb* = 11.0 g/dL

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	C	-2.3810	0.2065	0.0860	0.0165	1.0000
	D	-2.9024	0.3674	0.0550	0.0199	1.0000
	E	-2.1345	0.4477	0.1130	0.0482	1.0000
UPPER BASE	C	-2.1832	0.1934	0.1026	0.0180	1.0000
	D	-2.4202	0.3221	0.0849	0.0259	1.0000
	E	-1.9013	0.4014	0.1364	0.0501	1.0000
LOWER PRE	C	-2.5169	0.2720	0.0769	0.0198	1.0000
	D	-3.1754	0.4919	0.0445	0.0225	1.0000
	E	-2.3470	0.5991	0.0986	0.0606	1.0000
UPPER PRE	C	-2.4254	0.2196	0.0828	0.0169	1.0000
	D	-2.9829	0.3819	0.0513	0.0194	1.0000
	E	-2.1509	0.4686	0.1121	0.0504	1.0000
LOWER POST	C	-2.5162	0.2720	0.0769	0.0198	1.0000
	D	-3.1747	0.4919	0.0445	0.0225	1.0000
	E	-2.3463	0.5991	0.0987	0.0606	1.0000
UPPER POST	C	-2.5183	0.2720	0.0768	0.0197	1.0000
	D	-3.1767	0.4919	0.0444	0.0225	1.0000
	E	-2.3485	0.5990	0.0985	0.0605	1.0000

techniques to the results of the convolution step. The regression r-square values were about equal to 1 in most cases, so they are not listed in this table. Such high r-square values resulted for two reasons. First, in the case of a low PbB scenario, the dose-response function for PbB = 5 μg/dL essentially became the overall result. In higher PbB scenarios, the distributions tend to become normal.

D.4 ESTIMATES OF THE RISKS OF LEAD-INDUCED IQ EFFECTS AMONG U.S. CHILDREN AGED 7 YEARS

Two IQ health effects were considered, IQ decrement and the increased probability of having IQ values less than specified critical levels. A critical level is denoted by IQ*. Calculations for the risk of IQ decrements are quite simple and similar to both the EP and hemoglobin risk calculations. They followed directly from the encoded judgments of the experts regarding IQ decrement. Calculation of risks of the second type is much more complicated and includes uncertainty in: the mean IQ of populations of children sheltered from lead exposure, within-group standard deviation, and IQ decrements among children exposed to lead. The probabilistic judgments of

numerical methods analogous to a convolution integral, to produce an overall risk distribution for a particular SES group. If desired, these distributions can also be combined across SES to produce further risk distributions based on the judgments of each expert. Table D.4 lists relevant information about the IQ decrement risk distributions, and Tab. D.5 lists relevant informaton about the increased probability of children having IQ levels $\leq$ IQ*.

TABLE D.4 Summary of Distributions for Expected IQ Decrement at Age 7 for Lower SES Populations of Children

Source 1, Expected IQ Decrement

SCENARIO	EXPERT	MU R	MU Y	SIG Y	MU Δ_{IQ}	SD Δ_{IQ}
LOWER BASE	G	Normal			2.2875	0.7794
	H	Normal			4.5786	1.1163
	I	Normal			1.6460	0.4579
	J	Normal			4.1610	0.8254
	K	Normal			1.9940	0.4443
UPPER BASE	G	Normal			3.3145	0.8823
	H	Normal			6.0238	1.1424
	I	Normal			2.8566	0.6046
	J	Normal			5.3974	0.7917
	K	Normal			3.1143	0.5629
LOWER PRE	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-6.1658	0.5315	0.0021	0.0012
	J	Normal			2.4045	0.8865
	K	Normal			0.0032	0.0011
UPPER PRE	G	Normal			1.0974	0.4876
	H	Normal			2.9885	0.8825
	I	Normal			0.3307	0.1764
	J	Normal			2.8175	0.7049
	K	Normal			0.4953	0.1623
LOWER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4018	0.8887
UPPER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-7.2644	0.5315	0.0007	0.0004
	J	Normal			2.4027	0.8879
	K	Normal			0.0011	0.0004

TABLE D.4 (Cont'd)

Source 2, Expected IQ Decrement

SCENARIO	EXPERT	MU R	MU Y	SIG Y	MU Δ_{IQ}	SD Δ_{IQ}
LOWER BASE	G	Normal			2.3400	0.7872
	H	Normal			4.6489	1.1141
	I	Normal			1.7080	0.4673
	J	Normal			4.2207	0.8219
	K	Normal			2.0522	0.4541
UPPER BASE	G	Normal			3.3637	0.8796
	H	Normal			6.0991	1.1429
	I	Normal			2.9144	0.6081
	J	Normal			5.4617	0.7878
	K	Normal			3.1669	0.5621
LOWER PRE	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal			0.0068	0.0040
	J	Normal			2.4106	0.8815
	K	Normal			0.0105	0.0036
UPPER PRE	G	Normal			1.1665	0.5091
	H	Normal			3.0888	0.9009
	I	Normal			0.4000	0.2076
	J	Normal			2.9010	0.7095
	K	Normal			0.5948	0.1910
LOWER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4018	0.8887
UPPER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-6.4378	0.5315	0.0016	0.0009
	J	Normal			2.4039	0.8870
	K	Normal			0.0025	0.0009

TABLE D.4 (Cont'd)

Source 3, Expected IQ Decrement

SCENARIO	EXPERT	MU R	MU Y	SIG Y	MU Δ_{IQ}	SD Δ_{IQ}
LOWER BASE	G	Normal			2.2356	0.7713
	H	Normal			4.5095	1.1177
	I	Normal			1.5849	0.4486
	J	Normal			4.1023	0.8282
	K	Normal			1.9358	0.4344
UPPER BASE	G	Normal			3.2589	0.8820
	H	Normal			5.9439	1.1408
	I	Normal			2.7919	0.5996
	J	Normal			5.3278	0.7941
	K	Normal			3.0537	0.5616
LOWER PRE	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-6.7254	0.5315	0.0012	0.0007
	J	Normal			2.4034	0.8874
	K	Normal			0.0019	0.0007
UPPER PRE	G	Normal			1.0384	0.4760
	H	Normal			2.9020	0.8796
	I	Normal			0.2721	0.1478
	J	Normal			2.7454	0.7113
	K	Normal			0.4095	0.1360
LOWER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4018	0.8887
UPPER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-7.6009	0.5315	0.0005	0.0003
	J	Normal			2.4025	0.8881
	K	Normal			0.0008	0.0003

TABLE D.4 (Cont'd)

Source 4, Expected IQ Decrement

SCENARIO	EXPERT	MU R	MU Y	SIG Y	MU Δ_{IQ}	SD Δ_{IQ}
LOWER BASE	G	Normal			2.2470	0.7716
	H	Normal			4.5264	1.1172
	I	Normal			1.5983	0.4495
	J	Normal			4.1164	0.8275
	K	Normal			1.9486	0.4352
UPPER BASE	G	Normal			3.2701	0.8821
	H	Normal			5.9601	1.1412
	I	Normal			2.8050	0.6007
	J	Normal			5.3419	0.7936
	K	Normal			3.0660	0.5619
LOWER PRE	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-6.1658	0.5315	0.0021	0.0012
	J	Normal			2.4045	0.8865
	K	Normal			0.0032	0.0011
UPPER PRE	G	Normal			0.9969	0.4729
	H	Normal			2.8407	0.8863
	I	Normal			0.2309	0.1274
	J	Normal			2.6945	0.7223
	K	Normal			0.3490	0.1171
LOWER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4018	0.8887
UPPER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-8.1117	0.5315	0.0003	0.0002
	J	Normal			2.4022	0.8884
	K	Normal			0.0004	0.0002

TABLE D.4 (Cont'd)

Source 5, Expected IQ Decrement

SCENARIO	EXPERT	MU R	MU Y	SIG Y	MU Δ_{IQ}	SD Δ_{IQ}
LOWER BASE	G	Normal			0.9464	0.4753
	H	Normal			2.7657	0.9047
	I	Normal			0.1811	0.1012
	J	Normal			2.6323	0.7434
	K	Normal			0.2747	0.0931
UPPER BASE	G	Normal			1.6103	0.6746
	H	Normal			3.6991	1.1076
	I	Normal			0.8664	0.3447
	J	Normal			3.4139	0.8357
	K	Normal			1.1977	0.3206
LOWER PRE	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4020	0.8885
	K	Normal			0.0002	0.0001
UPPER PRE	G	Normal			0.8477	0.5017
	H	Normal			2.6182	0.9740
	I	Normal			0.0845	0.0483
	J	Normal			2.5099	0.8084
	K	Normal			0.1288	0.0444
LOWER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	J	Normal			2.4018	0.8887
UPPER POST	G	Lognormal	-0.4816	0.6422	0.7593	0.5425
	H	Lognormal	0.8268	0.4087	2.4851	1.0596
	I	Lognormal	-6.4378	0.5315	0.0016	0.0009
	J	Normal			2.4039	0.8870
	K	Normal			0.0025	0.0009

TABLE D.5 Summary of Distributions for Increased Probability of IQ≤IQ* at Age 7 for Lower SES Populations of Children

Source 1, IQ* = 70

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-3.7137	0.3469	0.0252	0.0088	1.0000
	H	-3.4238	0.3823	0.0337	0.0130	1.0000
	J	-3.3198	0.2950	0.0363	0.0106	1.0000
	K	-3.1979	0.3685	0.0416	0.0153	1.0000
UPPER BASE	G	-3.2636	0.3000	0.0383	0.0114	1.0000
	H	-2.9822	0.3366	0.0506	0.0167	1.0000
	J	-2.9282	0.2442	0.0521	0.0123	1.0000
	K	-2.6357	0.2747	0.0689	0.0181	1.0000
LOWER PRE	G	-5.4148	0.7608	0.0059	0.0052	1.0000
	H	-4.7670	0.7706	0.0112	0.0101	1.0000
	J	-4.4231	0.6651	0.0146	0.0108	1.0000
UPPER PRE	G	-4.6162	0.4346	0.0107	0.0049	1.0000
	H	-4.1058	0.4560	0.0179	0.0085	1.0000
	J	-3.8932	0.3769	0.0213	0.0082	1.0000
	K	-3.7984	0.4650	0.0242	0.0117	1.0000
LOWER POST	G	-5.4123	0.7608	0.0059	0.0052	1.0000
	H	-4.7645	0.7706	0.0112	0.0101	1.0000
	J	-4.4206	0.6652	0.0147	0.0109	1.0000
UPPER POST	G	-5.4132	0.7608	0.0059	0.0052	1.0000
	H	-4.7654	0.7706	0.0112	0.0101	1.0000
	J	-4.4215	0.6651	0.0147	0.0109	1.0000

TABLE D.5 (Cont'd)

Source 2, IQ* = 70

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-3.6877	0.3451	0.0258	0.0090	1.0000
	H	-3.4020	0.3812	0.0344	0.0132	1.0000
	J	-3.2994	0.2928	0.0370	0.0107	1.0000
	K	-3.1709	0.3674	0.0427	0.0156	1.0000
UPPER BASE	G	-3.2471	0.2995	0.0389	0.0115	1.0000
	H	-2.9628	0.3365	0.0515	0.0170	1.0000
	J	-2.9105	0.2434	0.0529	0.0124	1.0000
	K	-2.6157	0.2729	0.0702	0.0182	1.0000
LOWER PRE	G	-5.4206	0.7607	0.0058	0.0052	1.0000
	H	-4.7728	0.7706	0.0111	0.0100	1.0000
	J	-4.4290	0.6651	0.0145	0.0108	1.0000
UPPER PRE	G	-4.5434	0.4370	0.0115	0.0052	1.0000
	H	-4.0560	0.4600	0.0188	0.0090	1.0000
	J	-3.8558	0.3770	0.0221	0.0085	1.0000
	K	-3.7851	0.4664	0.0245	0.0119	1.0000
LOWER POST	G	-5.4123	0.7608	0.0059	0.0052	1.0000
	H	-4.7645	0.7706	0.0112	0.0101	1.0000
	J	-4.4206	0.6652	0.0147	0.0109	1.0000
UPPER POST	G	-5.4143	0.7608	0.0059	0.0052	1.0000
	H	-4.7664	0.7706	0.0112	0.0101	1.0000
	J	-4.4225	0.6651	0.0146	0.0108	1.0000

TABLE D.5 (Cont'd)

Source 3, IQ* = 70

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-3.8755	0.4381	0.0222	0.0100	1.0000
	H	-3.5978	0.4820	0.0296	0.0148	1.0000
	J	-3.4468	0.3707	0.0328	0.0122	1.0000
	K	-3.2633	0.3861	0.0393	0.0152	1.0000
UPPER BASE	G	-3.2819	0.2994	0.0377	0.0111	1.0000
	H	-3.0035	0.3365	0.0496	0.0164	1.0000
	J	-2.9467	0.2444	0.0512	0.0121	1.0000
	K	-2.6587	0.2759	0.0675	0.0178	1.0000
LOWER PRE	G	-5.4138	0.7608	0.0059	0.0052	1.0000
	H	-4.7660	0.7706	0.0112	0.0101	1.0000
	J	-4.4221	0.6651	0.0146	0.0108	1.0000
UPPER PRE	G	-4.8645	0.5504	0.0089	0.0052	1.0000
	H	-4.3354	0.5772	0.0151	0.0094	1.0000
	J	-4.0670	0.4781	0.0188	0.0094	1.0000
	K	-5.2291	0.5802	0.0063	0.0040	1.0000
LOWER POST	G	-5.4123	0.7608	0.0059	0.0052	1.0000
	H	-4.7645	0.7706	0.0112	0.0101	1.0000
	J	-4.4206	0.6652	0.0147	0.0109	1.0000
UPPER POST	G	-5.4130	0.7608	0.0059	0.0052	1.0000
	H	-4.7651	0.7706	0.0112	0.0101	1.0000
	J	-4.4212	0.6651	0.0147	0.0109	1.0000

TABLE D.5 (Cont'd)

Source 4, IQ* = 70

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-3.7347	0.3481	0.0247	0.0087	1.0000
	H	-3.4401	0.3831	0.0332	0.0128	1.0000
	J	-3.3334	0.2969	0.0358	0.0105	1.0000
	K	-3.2196	0.3683	0.0408	0.0150	1.0000
UPPER BASE	G	-3.2784	0.2996	0.0378	0.0112	1.0000
	H	-2.9995	0.3364	0.0498	0.0165	1.0000
	J	-2.9431	0.2442	0.0513	0.0121	1.0000
	K	-2.6539	0.2756	0.0678	0.0178	1.0000
LOWER PRE	G	-5.4148	0.7608	0.0059	0.0052	1.0000
	H	-4.7670	0.7706	0.0112	0.0101	1.0000
	J	-4.4231	0.6651	0.0146	0.0108	1.0000
UPPER PRE	G	-4.9027	0.5509	0.0085	0.0051	1.0000
	H	-4.3573	0.5759	0.0148	0.0092	1.0000
	J	-4.0867	0.4810	0.0184	0.0093	1.0000
	K	-5.3788	0.5797	0.0054	0.0034	1.0000
LOWER POST	G	-5.4123	0.7608	0.0059	0.0052	1.0000
	H	-4.7645	0.7706	0.0112	0.0101	1.0000
	J	-4.4206	0.6652	0.0147	0.0109	1.0000
UPPER POST	G	-5.4127	0.7608	0.0059	0.0052	1.0000
	H	-4.7648	0.7706	0.0112	0.0101	1.0000
	J	-4.4209	0.6651	0.0147	0.0109	1.0000

TABLE D.5 (Cont'd)

Source 5, IQ* = 70

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-4.9481	0.5534	0.0082	0.0049	1.0000
	H	-4.3852	0.5760	0.0144	0.0090	1.0000
	J	-4.1138	0.4855	0.0180	0.0091	1.0000
	K	-5.6092	0.5791	0.0043	0.0027	1.0000
UPPER BASE	G	-4.2739	0.4805	0.0153	0.0077	1.0000
	H	-3.8419	0.4893	0.0235	0.0120	1.0000
	J	-3.6962	0.4115	0.0262	0.0110	1.0000
	K	-3.6965	0.4609	0.0267	0.0127	1.0000
LOWER PRE	G	-5.4125	0.7608	0.0059	0.0052	1.0000
	H	-4.7646	0.7706	0.0112	0.0101	1.0000
	J	-4.4207	0.6652	0.0147	0.0109	1.0000
UPPER PRE	G	-5.0453	0.5672	0.0075	0.0046	1.0000
	H	-4.4330	0.5853	0.0138	0.0088	1.0000
	J	-4.1446	0.5057	0.0176	0.0094	1.0000
	K	-6.3526	0.5777	0.0021	0.0013	1.0000
LOWER POST	G	-5.4123	0.7608	0.0059	0.0052	1.0000
	H	-4.7645	0.7706	0.0112	0.0101	1.0000
	J	-4.4206	0.6652	0.0147	0.0109	1.0000
UPPER POST	G	-5.4143	0.7608	0.0059	0.0052	1.0000
	H	-4.7664	0.7706	0.0112	0.0101	1.0000
	J	-4.4225	0.6651	0.0146	0.0108	1.0000

TABLE D.5 (Cont'd)

Source 1, IQ* = 85

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-2.6805	0.2977	0.0665	0.0190	1.0000
	H	-1.8746	0.2786	0.1362	0.0337	1.0000
	J			0.1114	0.0232	
	K			0.0605	0.0137	
UPPER BASE	G	-2.2625	0.2581	0.0966	0.0230	1.0000
	H	-1.5063	0.2471	0.1843	0.0380	1.0000
	J			0.1478	0.0231	
	K			0.0942	0.0171	
LOWER PRE	G	-4.2354	0.7058	0.0180	0.0144	1.0000
	H	-2.8860	0.5182	0.0589	0.0313	1.0000
	J			0.0616	0.0237	
UPPER PRE	G	-3.5192	0.3840	0.0308	0.0119	1.0000
	H	-2.4489	0.3155	0.0826	0.0247	1.0000
	J			0.0735	0.0191	
	K			0.0152	0.0052	
LOWER POST	G	-4.2328	0.7058	0.0181	0.0144	1.0000
	H	-2.8834	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER POST	G	-4.2337	0.7058	0.0180	0.0144	1.0000
	H	-2.8843	0.5182	0.0590	0.0313	1.0000
	J			0.0617	0.0237	

TABLE D.5 (Cont'd)

Source 2, IQ* = 85

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-2.6554	0.2955	0.0680	0.0192	1.0000
	H	-1.8563	0.2772	0.1383	0.0340	1.0000
	J			0.1130	0.0231	
	K			0.0622	0.0139	
UPPER BASE	G	-2.2450	0.2568	0.0981	0.0232	1.0000
	H	-1.4893	0.2467	0.1868	0.0383	1.0000
	J			0.1498	0.0231	
	K			0.0957	0.0171	
LOWER PRE	G	-4.2413	0.7057	0.0179	0.0143	1.0000
	H	-2.8922	0.5180	0.0586	0.0311	1.0000
	J			0.0613	0.0235	
UPPER PRE	G	-3.4542	0.3833	0.0327	0.0127	1.0000
	H	-2.4067	0.3180	0.0858	0.0258	1.0000
	J			0.0758	0.0193	
	K			0.0182	0.0061	
LOWER POST	G	-4.2328	0.7058	0.0181	0.0144	1.0000
	H	-2.8834	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER POST	G	-4.2348	0.7058	0.0180	0.0144	1.0000
	H	-2.8854	0.5182	0.0590	0.0313	1.0000
	J			0.0616	0.0237	

TABLE D.5 (Cont'd)

Source 3, IQ* = 85

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-2.8112	0.3738	0.0601	0.0220	1.0000
	H	-1.9831	0.3455	0.1257	0.0396	1.0000
	J			0.1073	0.0232	
	K			0.0570	0.0134	
UPPER BASE	G	-2.2813	0.2587	0.0950	0.0227	1.0000
	H	-1.5247	0.2469	0.1816	0.0375	1.0000
	J			0.1458	0.0231	
	K			0.0924	0.0171	
LOWER PRE	G	-4.2343	0.7058	0.0180	0.0144	1.0000
	H	-2.8850	0.5182	0.0590	0.0313	1.0000
	J			0.0617	0.0237	
UPPER PRE	G	-3.7270	0.4877	0.0262	0.0133	1.0000
	H	-2.5868	0.3907	0.0743	0.0282	1.0000
	J			0.0706	0.0191	
	K			0.0120	0.0043	
LOWER POST	G	-4.2328	0.7058	0.0181	0.0144	1.0000
	H	-2.8834	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER POST	G	-4.2335	0.7058	0.0181	0.0144	1.0000
	H	-2.8840	0.5182	0.0591	0.0314	1.0000
	J			0.0617	0.0237	

TABLE D.5 (Cont'd)

Source 4, IQ* = 85

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-2.6990	0.2991	0.0653	0.0188	1.0000
	H	-1.8901	0.2799	0.1345	0.0335	1.0000
	J			0.1101	0.0232	
	K			0.0592	0.0134	
UPPER BASE	G	-2.2775	0.2586	0.0953	0.0228	1.0000
	H	-1.5210	0.2471	0.1821	0.0376	1.0000
	J			0.1462	0.0231	
	K			0.0928	0.0171	
LOWER PRE	G	-4.2354	0.7058	0.0180	0.0144	1.0000
	H	-2.8860	0.5182	0.0589	0.0313	1.0000
	J			0.0616	0.0237	
UPPER PRE	G	-3.7631	0.4917	0.0254	0.0130	1.0000
	H	-2.6094	0.3910	0.0728	0.0277	1.0000
	J			0.0694	0.0194	
	K			0.0103	0.0037	
LOWER POST	G	-4.2328	0.7058	0.0181	0.0144	1.0000
	H	-2.8834	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER POST	G	-4.2332	0.7058	0.0181	0.0144	1.0000
	H	-2.8837	0.5182	0.0591	0.0314	1.0000
	J			0.0617	0.0237	

TABLE D.5 (Cont'd)

Source 5, IQ* = 85

SCENARIO	EXPERT	MU Y	SIG Y	MU R	SIG R	R SQR
LOWER BASE	G	-3.8108	0.4985	0.0243	0.0127	1.0000
	H	-2.6388	0.3925	0.0708	0.0271	1.0000
	J			0.0679	0.0199	
	K			0.0082	0.0030	
UPPER BASE	G	-3.1956	0.4214	0.0425	0.0181	1.0000
	H	-2.2355	0.3727	0.1014	0.0356	1.0000
	J			0.0886	0.0230	
	K			0.0356	0.0101	
LOWER PRE	G	-4.2330	0.7058	0.0181	0.0144	1.0000
	H	-2.8835	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER PRE	G	-3.8960	0.5234	0.0226	0.0125	1.0000
	H	-2.6816	0.4057	0.0684	0.0272	1.0000
	J			0.0648	0.0216	
	K			0.0039	0.0014	
LOWER POST	G	-4.2328	0.7058	0.0181	0.0144	1.0000
	H	-2.8834	0.5182	0.0591	0.0314	1.0000
	J			0.0618	0.0237	
UPPER POST	G	-4.2348	0.7058	0.0180	0.0144	1.0000
	H	-2.8854	0.5182	0.0590	0.0313	1.0000
	J			0.0616	0.0237	

encoded judgments and the fitted functions. Each table is organized so that each expert/hemoglobin-level/age-category combination is listed in the table, beginning with Expert A, continuing in alphabetical order, and ending with Expert E.

B.3.1 EXPERT A

Expert A chose to separate the population of U.S. children aged 0-6 years into two subpopulations, ages 0-3 and 4-6 years. This individual provided judgments only at the 11 g/dL hemoglobin level. Since his judgments indicated that there is a very small lead-induced hemoglobin effect at blood-lead levels below 45 μg/dL, it was not necessary to obtain probabilistic judgments at the 9.5 g/dL level.

The probabilistic judgments of Expert A regarding the effects of lead exposure on hemoglobin levels among U.S. children aged 0-3 years are listed in Table B.2 for the 11.0 g/dL hemoglobin level. The probability judgments are cumulative distribution functions (CDFs) of judgmental probability. That is, for each blood-lead level L, the entries in column F in Table B.2 represent the judged probability that the true response rate R_T is less than or equal to the rate R shown in column R of the table. There are four CDFs, one each for blood-lead levels of 45, 55, 65, and 75 μg/dL. Expert A did not feel that there was a measurable, lead-induced hemoglobin effect at blood-lead levels below 45 μg/dL. The curves display a wider range of plausible population response rates at successively higher blood-lead levels. This may be interpreted to mean the Expert A is less certain about what actual response rates may be at the higher blood-lead levels. There is no indication of a threshold for a hemoglobin effect in the range of 45-75 μg/dL.

For the reasons presented in Sec. 1, mathematical functions were fit to the judgments of Expert A (and the judgments of the other experts). For Expert A, the mathematical representations for the various blood-lead levels are guaranteed never to cross. They were obtained by fitting regression lines to transformations of the judgments. The transformation used was the normal-on-log-odds transformation. It

happens that this transformation leads to distributions that are normal distributions, with a high degree of accuracy. These distributions are also convenient for use in subsequent analysis. We call these fitted representations "normal-on-log-odds" (NOLO) functions. They are uniquely defined by the mean and variance of the underlying normal distributions, and the transformation functions; these are summarized in Sec. B.2. Table B.3 summarizes relevant parameters for each of the conditional CDFs fit to the judgments of Expert A. Included in the table for each lead level at which a CDF was assessed are the mean and standard deviation of the underlying normal distribution, the mean and standard deviation of the NOLO distribution, and the R-squares of regressions of the fitted representations on encoded judgments. In general, the regression R-square values are reassuringly high. Finally, Table B.4 summarizes and compares the judgments and the fitted representations. Included in the table are medians and several credible intervals. The median is that value that is exceeded with probability 0.5. The 90% credible interval is a set of response rate values such that there is a 0.9 probability of the true value falling within it. For example for children aged 0-3 having blood-lead levels of 55 μg/dL, the median encoded response rate is 9%, and the 90% credible interval is 1% to 15%; that is, with probability 0.05 the true response rate is less than or equal to 1%, and with probability 0.05 it is greater than 15%. Similar statements apply to the fitted representation. Note the the encoded median values are fairly close to the medians of the fitted representations, differing by two to four percentage points. However, the fitted representations exhibit more uncertainty than do the judgments. These features and differences were carefully pointed out to Expert A. He finally concluded that the fitted representations better captured his best judgments about the effects of lead on hemoglobin levels among U.S. children.

B.3.2 Expert B

Expert B did not feel comfortable with the process of giving probabilistic judgments about the effects of lead on hemoglobin levels among children. However, he was able to discuss qualitatively the effects of lead exposure on hemoglobin and EP levels. A summary of that discussion is presented in Sec. B.4.

B.3.3 Expert C

Expert C chose to provide judgments for the population of children aged 0-6 years. He acknowledged that there certainly were differences in metabolism between children aged 0-3 and 4-6, but he felt that these differences were difficult to quantify considering other uncertainties. Thus, he provided two sets of judgments, one set at each of the hemoglobin levels of interest.

Following the format introduced for discussing the judgments of Expert A, it can be seen that the fitted functions closely match the encoded judgments. The median response rate values are no more than 1% apart. The fitted functions do display larger uncertainty at the higher blood-lead levels. Expert C concluded, after study and reflection, that the mathematical representations did capture his best judgments.

B.3.4 Expert D

Expert D chose to divide the population into two subgroups, children aged 0-3 and 4-6 years, and provided judgments at both the 9.5 and 11.0 g/dL hemoglobin levels. This resulted in four sets of judgments for Expert D.

At the 9.5 g/dL hemoglobin level, there is evidence of a slight threshold between blood-lead levels 25 and 35 μg/dL, indicated by the wider spacing between these levels compared to the other levels. At the 11 g/dL hemoglobin level, a threshold is indicated between the 15 and 25 μg/dL blood-lead levels. This was recognized and agreed to by Expert D.

In general, there is again good agreement between the mathematical representations and the judgments for Expert D. R-square values for regressions of $\hat{F}$ on F are generally around 0.95.

B.3.5 EXPERT E

Like Expert D, Expert E provided four sets of judgments, one for each of the age and hemoglobin level combinations that result from dividing the age range into subgroups aged 0-3 and 4-6 years. The judgments and fitted mathematical representations are listed and compared in the tables already mentioned.

Only one case for Expert E showed evidence of a threshold effect. That was at the 9.5 g/dL hemoglobin level for children aged 4-6 years. The threshold appears between the blood-lead levels of 35 and 45 µg/dL. In this case, as well as the others for Expert E, the fitted mathematical representations matched the judgments quite well. the R-square values of the regressions of $\hat{F}$ on F were generally around 0.95.

B.4 SUMMARY DISCUSSIONS

B.4.1 Summary of Discussion with Expert A

1. There is absolutely no evidence that ZPP is toxic per se.
2. The biosynthenis of heme is regulated by a negative feedback process (Stanbury, "Metabolic Basis of Inherited Disease"). The rate limiting factors are heme, heme oxygenase, and ALA synthase. The enzymes in the pathway from ALAD dehydratase to ferrochelatase are present in substantial excess. There is normally a substantial excess capacity of these intermediary enzymes. For example, there is sixteen times as much ALAD-dehydratase as is necessary to metabolize the amount of ALA

produced by ALA synthase. Overall, for every 10,000 molecules of heme produced, the equivalent of one molecule is lost as ALA, copper porphyrin, and protoporphyrin.

3. Lead causes partial inhibition of ALA dehydratase and ferrochelatase. EP is a marker of the partial inhibition that occurs at the ferrochelatase step. In the presence of lead, it has been observed that increased FEP levels correlate with decreased ALAD levels. Some genetic studies suggest that high FEP levels are associated with decreased ALAD activities, even in the absence of lead.

4. ZPP is bound to the red cell and therefore remains throughout the life of the red cell, which is about 120 days. Insofar as is known, all cells synthesize the heme necessary for the cell's function, and lead inhibits heme synthesis in most other cells of the body, including liver, kidney, and brain. The heme enzymes in these other cells may have very short half lives. These are probably quite important in considering lead effects. One of these is the P-450 family of enzymes. Also of concern is the activation rate of an enzyme in the kidney, which may be inhibited to a harmful level. In the liver, the P-450 enzymes are essential in the metabolism of drugs.

5. There is a dose-dependent decrease in the level of 1,25 dihydroxy-vitamin D in serum. This relationship is nonlinear, and the level decreases at a much higher rate as lead rises above 35-40 μg/dL. Although the mechanism for the reduction of 1,25 dihydroxy-vitamin D in serum is not fully understood, it is possible that this may be realted to an effect of lead on production of P-450 in the

kidney. P-450 is one of the factors necessary for the hydroxylization of 1,25 dihydroxyvitamin D at the 1 position, which occurs in the kidney. Further research will be required to evaluate the biological significance of reduced serum levels of 1,25 dihydroxyvitamin D.

6. A new disease related to the virtual absence of ALAD has been described. This is a genetic disorder transmitted as a recessive trait. People so affected have symptoms similar to those having acute intermittent prophyria, another genetic disorder of porphyrin metabolism. It is extremely doubtful that this has anything to do with low-level lead exposure, although at high PbB levels (70-80 µg/dL or higher), it is possible that lead inhibits ALAD to the same extent as that found in this new disease. In this new disease, ALAD activities are about 1% of normal. Furthermore, people have been identified with ALAD levels at about 20% of normal, and these people are healthy and without any symptoms whatsoever.
7. Clinically speaking, children with blood leads around 40 µg/dL, elevated EP, and other evidence of iron deficiency can be treated with iron and this anemia and/or iron deficiency can be corrected, even though no specific treatment for lead is given. The symptoms associated with anemia would be the same irrespective of whether the cause was lead exposure or iron deficiency. However, the presence of iron deficiency increases the absorption of lead from the gut.
8. Another possible effect involves tryptophane pyrrolase. This is a heme-dependent enzyme in the liver. It has been postulated that

the reduced level of heme in porphyria may be responsible for some of the symptoms seen in acute intermittent porphyria. A *Science* article by Corriea about a year ago addresses this issue.

9. Regarding the possible effects of ALA itself, effects are not likely because ALA does not readily cross the blood-brain barrier.
10. Dick Bull and McCauley (EPA, Cincinnati) have found delay of myelinization of the nerve sheaths in the brain in conjunction with lead in the 30-40 µg/dL range in neonatal rats. It is known that lead interferes with oxidative phosphorylation and hence energy production. This lead-related disruption of energy production may be responsible for this delay in mylenization.
11. Patients with sickle-cell anemia appear to be more susceptible to the toxic effects of lead, although the mechanisms responsible for these are not understood.
12. In general, there seems to be good evidence that elevated EP levels indicate increased risk of adverse health effects. It seems to be a good indicator of both FeD and Pb exposure. This places a child in double jeopardy since both lead to the same result. EP level is useful for monitoring long-term exposure.
13. Epidemiological studies should be based on data acquired through extraction techniques rather than through the hematofluorometer. Extraction techniques are used to calibrate hemotofluoremeters. There are data showing that hemotofluoremeters give readings taht are systematically low. It would appear that hematofluoremeters should be recalibrated at the factory every three months or so. At this time, the technical difficulties with the hematofluoremeter have not been resolved.

B.4.2 Summary of Discussion with Expert B

The discussion with Expert B covered two main topics: hemoglobin and EP.

1. A child can have a high PbB (say, 500 mcg/dL), no reduction in hemoglobin level, and die. Thus, hemoglobin level is not a very useful index for adverse health effects of lead, at least from a clinical point of view.
2. Neurological and behavioral (aggression, hyperactivity) effects occur much more quickly in children than hematological effects, and at much lower levels (about 40 mcg/dL).
3. The body can also compensate for reduced hemoglobin; other processes will begin to transport oxygen, offsetting, to some degree, the possible adverse effects of reduced hemoglobin.
4. Anemia only becomes a problem when the body's ability to compensate is lost.
5. Damage to the heme synthetic pathway is crucial; but this is not attributable to reduced hemoglobin.
6. A rough dose/effect relationship between PbB and Hb was expressed, one for Fe deficient kids, and one for non-Fe deficient kids. Expert B estimated that, on the average, non-Fe deficient kids with PbB=10 mcg/dL would have hemoglobin levels of about 12 g/dL; at PbB=50, Hb=10.5. For Fe deficient kids, Expert B estimated Hb=10 at PbB=10, and Hb=6.5-7.0 at PbB=50. However, Expert B felt that it was quite possible that it could be even lower for the later kids at PbB=50, but just couldn't be sure, and could not express these feelings quantitatively, despite efforts to assist Expert B using judgmental probability encoding techniques.

B.4.3 Summary of Discussion with Expert C

1. Very high EP levels can cause porphyria due to errors in metabolism with clinically recognizable skin and CNS effects. However, there is no evidence of a lead worker ever developing porphyria as a result of lead exposure. Thus, it is reasonable to conclude that lead exposure does not cause porphyria.
2. There is no evidence in the literature that lead-induced elevated EP itself causes health problems; it simply is a mrker, or indicator, of adverse health effects due, possible, to lead exposure. EP levels > 1 sigma above the mean should be followed by a look at blood lead.
3. Since EP levels tend to lag changes in PbB levels, it is important to be looking at both over time to get a good idea of what is going on. The rate of rise of EP vs. an increase in PbB is greater than linear, but lagged in time. A plot of log (EP) versus PbB would be roughly linear under chronic relatively constant lead exposure conditions.
4. Lead-induced elevated EP is an indicator of interference with heme synthesis which may or may not be compensated by derepression of ALA synthetase, the rate-limiting enzyme involved in heme systhesis, depending on the degree of lead exposure.
5. It is possible that elevated EP is an indicator of an interference by lead of the body's ability to produce 1,25 dihydroxyvitamin D. The cytochrome P-450 enzymes, of which there are over 50 differenct isozymes, are related to the synthesis of 1,25 dihydroxyvitamin D. There are strong indications that lead

reduces synthesis of the P-450 enzymes. This in turn may reduce the metabolism of 25-(OH)D, a necessary ingredient in the biosynthesis of 1,25 dihydroxyvitamin D. Reference was made to an important review paper concerning 1,25 dihydroxyvitamin D metabolism [D.R. Fraser, Physiological Reviews, 60:551 (1983)].

6. There is also suggestive evidence, but not proof, that elevated EP indicated reduced ability of the liver to detoxify. this effect has been shown in a number of studies involving high lead exposure in workers and children.

7. Furthermore, elevated EP levels may indicate sufficient exposure to lead (and/or other metals) to accelerate the destruction (i.e., reduce their half-lives) of the hemoproteins [M.D. Maines, Annals of Clinical Research, 8:39 (1976)].

8. Regarding the CNS, it is not clear whether elvated EP per se causes any effects. There are some animal studies concerning CNS effects when very large doses of the EP precursor ALA are administered. Some report CNS effects and others not [M. Moore & P.A. Meredith, Trace Substances in Environmental Hlth. 10:363 (a977); S. Edwards, Neurosci. Lett. 50:169 (1984)].

9. In a lead worker study [Hammond, et al, J. Occ. Med., 2:475-484 (july, 1980)], ALA in plasma and in urine was found to be as good a marker of the degree of toxicity of Pb as was PbB; ALA was as good a marker as EP for hematopoietic system damage; for other effects (neurological, renal, etc.), ALA and PbB were better indicators of damage than EP. Another study by Ruth Lillis et al. at Mt. Sinai showed that EP was a better indicator of hematopoietic effects than ALA and PbB. The differences

between the two studies may be due to differential measurement reliabilities in the two labs.

10. There is no evidence of an EP vs. hemoglobin relationship in children at PbB's <50. There are factors other than lead that affect EP level.
11. The possibility of Pb exposure causing hypertension is a new issue.
12. To summarize, there is no evidence that elevated EP per se due to lead causes adverse health effects.

B.4.4 Summary of Discussions with Expert D

The discussion with Expert D included FEP, liver, calcium, and 1,25 dihydroxy-vitamin D ($1,25(OH)_2D$) effects.

FEP:

1. Elevated EP due to lead exposure indicates direct toxic effects of lead on heme, which is a basic physiological system that is common to many organs and cell types.
2. Piomelli has postulated, and Silbergeld have published some data, and both suggest that some of the CNS effects of lead may be attributable to altered metabolism of porphyrin compounds. Others have suggested that there may be perturbations in CNS intracellular calcium metabolism.

Liver:

3. There is evidence at 30-40 µgPb/dL that liver metabolism of model compounds (e.g., cortisol is altered. This implies a reduced ability to detoxify the blood stream. Since many drugs have

metabolic pathway that are similar to that of cortisol, their metabolism may also be affected by lead.

Calcium:

4. There is evidence that calcium-mediated basic enzymatic systems in all mamallian cells are affected at lead levels above about 5 micromolar. Any disruption of these systems may have pervasive effects. Calcium from calmodulin can be displaced by lead. Calmodulin is a central component of the normal chemistry of cells. Lead can purturb this process.

1,25 Dihydroxy-vitamin D:

5. A complex enzyme system is responsible for the production of this vitamin D hormone. Lead is known to impair several ingredients of this enzyme system. Some of these ingredients include electron transport, mitochondrial function, and the cytochrome P-450 family of enzymes. Observations of impairment in the production of the vitamin D hormone in lead-toxic children have been confirmed in experimental studies in vivo and in vitro.
6. At 33-55 µgPb/dL, there is evidence of about 66% decrease in the kidney's ability to produce the vitamin D hormone.
7. At levels above 62 µgPb/dL, there is evidence of a decrease in $1,25(OH)_2D$ levels to a degree that have been reported in children with inborn errors in metabolism.
8. Between 12 to 120 µgPb/dL, there is a statistically significant negative correlation between vitamin D level and lead level.

9. At blood-lead levels above 55 µgPb/dL, chelation therapy is immediately used. At 25-55 µgPb/dL, the results of the EDTA provocative test is used to determine which children would benefit from chelation therapy.

10. $1,25(OH)_2D$ affects the maturation of cells and enhances the differentiation of cells.

    - Rats and mice having leukemia were treated with picomolar concentrations of $1,25(OH)_2D$. Results indicated that the lifespans of the animals were extended.
    - In vitro experiments on a variety of human cells (e.g. lymphoma, myeloid leukemia cells, monocytes) have replicated such findings.

11. There is recent research showing that $1,25(OH)_2D$ evidences some immuno-regulatory functions like other steroid hormones.

12. There is early, evolving evidence that $1,25(OH)_2D$ has a role in controlling insulin secretion from the pancreas.

13. The actions of the vitamin D hormone involve not only mineral absorption, bone remodeling, and calcium homeostasis in virtually alll mammalian cells, but recently reported information indicates that the vitamin D hormone has even more pervasive effects in humans. Some of these effects, which are now recognized, include enhancement of cell differentiation (maturation), and immuno-regulatory capacity. This, in consideration also of points 608, reasonably suggests that a decrease of vitamin D hormone is likely to have pervasive effects on the function of other organs.

B.4.5 Summary of Discussions with Expert E

Discussions with Expert E focused primarily on FEP effects of lead.

1. There does not seem to be any evidence that elevated EP is itself an adverse health effect. However, it is possible because anytime the balance of an essential metabolite is upset, there is a risk of damage to other organs or functions. For example:
   - Phenylalanine is an essential amino acid, but is toxic at high levels. This is seen in the disease phenylketonuria, in which there is an inability to metabolize excess phenylalanine. Thus, (in the absence of a special diet) the phenylalanine builds up and affects the developing central nervous system, causing mental retardation.
   - Excess levels of fluoride cause teeth to become brittle, although proper levels increase hardness of teeth.

   Thus, there is the chance that future research might show that elevated FEP is an adverse health effect.
2. Elevated EP is a valuable index of adverse lead effects in other organs, e.g., the central nervous system, the liver, the kidney, etc.
3. The body has the ability to compensate for lead insult up to a point. For example, at low lead levels the inhibition of EP is compensated for, but not at slightly higher levels. Piomellis's analysis indicates a threshold around 16-17 µg/dL. ALAD shows no threshold.
4. Regarding measurement of EP, care must be taken if the hematoflorometer is used because the amount of EP that is zinc protoporphyrin (ZPP) decreases proportionately as total EP rises;

since the hematoflorometer measures ZPP, it may be underestimating total EP. Below 35 µg/dL there is good agreement between FEP (extraction) and ZPP (hemotofluorometer). Above 35 µg/dL there appears to be a 30% difference between ZPP (lower) and FEP (higher).

5. In assessing the need for treatment, both EP and PbB levels should be considered (as the Centers for Disease Control does).
6. Critical blood-lead levels for regulatory purposes may need to be less than those considered for clinical purposes in order to provide an adequate margin of safety for the population at large.

the health risks associated with various potential standards will aid the EPA administrator in selecting that standard which, in his judgment, protects the public health with an adequate margin of safety.

Because the risk estimates that EPA seeks are often based in part on dose-response relationships and lead exposure estimates which are uncertain, it is necessary to make probability judgments about relevant dose-response functions based on the available evidence, and to probabilistically estimate lead exposure under alternative NAAQS. Obtaining the health risk estimates then involves combining dose-response and exposure probability estimates.

The problem of estimating dose-response relationships is similar to that which exists in clinical medicine when there do not exist data that bear precisely on the patient's problem. In that case it is necessary to use scientific judgment to extrapolate from the data to make a best decision for the patient. Here, too, it is necessary to use scientific judgment to extrapolate from the available data. The extrapolation is not certain and, therefore, we will aid you to represent your opinion probabilistically. Furthermore, since the extrapolation depends on one's interpretation of the literature, different people will have different judgments. For each health effect we intend to obtain the probabilistic judgments of about five experts to sample the range of respected opinions. The model for estimating risks will not merge these judgments into a single average judgment, but rather will estimate the range of risks based on the range of judgments. If we as risk analysts do our job properly, then not only will we be able to show the EPA administrator the range of estimates based on the range of judgments, but we will also be able to show some of the sources of the disagreements. Indeed, a side benefit of this exercise in which we probe your knowledge in a structured manner may be to help identify sources of greatest disagreement.

Based on the evidence in the lead criteria document, two populations have been identified as being most susceptible to the effects of lead intoxication. One is children

APPENDIX C

INTELLIGENCE QUOTIENT

This appendix is organized in a fashion similar to App. B for hemoglobin. Section C.1 is the IQ protocol, and Sec. C.2 summarizes the mathematical formulations. Section C.3 tabulates the encoded judgments, specifications for functions fit to those judgments, various credible intervals, and comparisons of encoded judgments vs. fitted functions. Section C.4 summarizes of the discussions with each of the IQ experts.

C.1 IQ PROTOCOL

C.1.1 Introduction

The Environmental Protection Agency (EPA) is charged by the Clean Air Act with setting and revising National Ambient Air Quality Standards (NAAQS) for selected pollutants, at levels sufficient to protect the public health with an adequate margin of safety. As you know, the scientific bases for NAAQS are presented and reviewed in criteria documents. In support of the forthcoming review of the lead NAAQS, EPA has just prepared a new *Air Quality Criteria for Lead.* It presents scientific evidence from which the most susceptible populations can be determined and from which various adverse health effects can be identified. The criteria document summarizes and evaluates the available clinical, epidemiological, and animal or toxicological laboratory evidence with regard to the physiological and adverse health effects of lead, and therefore represents our most up to date knowledge on lead effects.

As one aspect of the review process, EPA assesses health risks by identifying the most sensitive populations for each pollutant, and estimating probabilistically the numbers of people in the populations who may suffer each of various well-defined adverse health effects attributable to the pollutant. It is believed that information about

from birth through the seventh birthday, and the other is pregnant women, or more precisely, the fetuses being carried by the pregnant women. A number of adverse health effects have been identified for which we would like to estimate dose-response functions including the one discussed below.

C.1.2 Lead-Induced IQ Decrements

C.1.2.1 Central Nervous System and Behavioral Effects of Lead

A large number of studies reviewed in the Criteria Document suggest that there are numerous central nervous system (CNS) and behavioral effects of lead exposure. Considerable uncertainty surrounds all of these effects, however, because of the enormous difficulty in defining and measuring them, and in isolating them from the effects of covariates. Since our goal is to obtain probabilistic estimates about the shape of a dose-response curve for a particular well-defined effect if a sufficient amount of pertinent data could be collected, our first task is to select one CNS or behavioral effect that is of acknowledged importance and that can be well specified.

The search for lead-induced effects has included studies of: EEG effects; sensory motor, perceptual, and attentional deficits; cognitive decrements of various sorts; hyperactivity; negative classroom behaviors; and other effects that so far have been studied only in animal models. From this large assortment of effects, we have selected IQ decrement as the adverse effect upon which to focus. We are not considering IQ to be the only nor necessarily the best measure of cognitive abilities. Nor is it being considered as a surrogate for the other systems in which effects have been explored. (Because of its multifaceted nature it probably involves many of them.) Rather, we have selected IQ decrement because there are more data on this effect than on any other, and because its functional or "clinical" significance is clear. It is quite conceivable that as research continues in future years, other more "pure" CNS or behavioral effects will

emerge that are obviously adverse and for which a substantial body of data has developed. Until such an event occurs, it appears that lead-induced IQ decrement is the most appropriate effect to consider.

C.1.2.2 Definition of IQ Decrement

Because it is generally thought that CNS effects of lead may be cumulative, it is necessary to specify the children's ages at which the IQ decrements are estimated. In all cases, assume that IQ tests are given on the children's seventh birthdays. Assume further that the WISC-R is the test employed.

A lead-induced IQ change cannot be measured directly for a given child, and therefore the change would have to be estimated statistically from suitable data. Thus, rather than ask you directly about dose-response functions, we will encode your judgments about the outcomes of hypothetical ideal experiments. If you agree with the assumptions on which the hypothetical experiments are based, then your judgments about the potential outcomes will lead naturally to probabilistic estimates about dose-response functions for IQ decrements.

There are differences of opinion, of course, as to what IQ decrements should be considered adverse. The Clean Air Act makes it clear that EPA should set standards to protect against adverse health effects, but the level defined as adverse may be different for regulatory than for clinical or remedial purposes. We will not focus on particular magnitudes of IQ decrements, but rather will encode your probabilistic judgments about IQ distributions given various blood-lead levels. Then we will use your judgments to derive probabilistic statements about the percent of children at each lead level whose IQ scores are below any value of interest, such as, for example, below IQs of 70, 80, 90, or 100.

C.1.2.3 Hypothetical Ideal Experiments

Assume that at conception, subjects are randomly assigned to groups differentiated by blood lead level targeted for the third birthday. Members of each treatment group are exposed to lead from conception until their seventh birthdays, while members of a control group are sheltered from lead from conception until their seventh birthdays. The level of environmental lead assigned to a child is roughly constant for the seven years, however, each child's lead uptake is *not* constant, due to the changes with age of his or her physiology and behavior. Nevertheless, the experimental conditions are such that at their *third* birthdays, all the children in each group have essentially the same measured blood lead. Thus, groups differ in terms of mean blood lead on the third birthday. Environmental lead levels necessary to yield a given blood-lead level at age three in a particular child remain constant through the seventh birthday.

The experimental manipulation affects only lead exposure, and no other aspect of the children's lives. Then the WISC-R IQ test is administered to all children on their *seventh* birthdays. The children in each group have a distribution of IQ scores with some mean and standard deviation. We are interested in your probabilistic judgments about the IQ distributions for groups of children on their seventh birthdays, all of whom had specific blood-lead levels on their third birthdays.

Note three features about this hypothetical experiment. First, it is similar to a longitudinal study; children are in a group from conception until their seventh birthdays. Second, it involves random assignment of children to groups, making it unnecessary to worry about covariates. Third, blood lead is measured at age *three*, and IQ at age *seven*.

If you believe IQ effects of lead to be different for lower SES than for middle and upper SES subpopulations of children under these experimental conditions, then we will consider separate hypothetical experiments for the two subpopulations. Otherwise, we will consider only a single such experiment.

SES is a complex variable. For simplicity, if we divide the population into low SES and middle or upper SES categories, let us define low SES as those children coming from households with incomes in the lowest 15 percent.

Unless you disagree, we will also assume that the distribution of IQ scores is normal within each group. (The WISC-R was developed to yield normally distributed IQ scores for the general population, with mean IQ equal to 100 and a standard deviation of 15.)

Furthermore, if we consider only a single hypothetical experiment, sampling from the full population of children, then, unless you disagree, we will assume that the IQ standard deviation is 15 at each blood-lead level and we will only have to encode your probabilistic judgment about the mean IQ at age seven for each of several blood-lead levels at age three.

If we consider two hypothetical experiments, each sampling from a different SES subpopulation, then the IQ standard deviation will possibly be less than 15. Assuming that IQ standard deviation is the same at all blood-lead levels within SES subpopulations, we will have to encode your probabilistic judgment about the standard deviation for each SES group.

Then, separately for each SES group or for the population as a whole, as appropriate, we will encode your probabilistic judgment about the mean IQ at age seven of the control group. Finally, we will encode your probabilistic judgments about the differences in mean IQ at age seven between children with negligible blood lead (the control group) and children at each of several elevated blood leads at age three (the treatment groups).

The following sections specify further the conditions to be assumed in this hypothetical experiment.

C.1.3 Population at Risk

Based on the Criteria Document, we can specify the most susceptible population as all United States children from conception through their seventh birthdays. We have already defined the effect as being measured on the seventh birthday, but on the assumption that younger children will reach their seventh birthdays, the effect is well-defined for the whole population.

However, as already discussed, you may consider IQ effects of lead to be different for lower SES than for middle and upper SES subpopulations of children, resulting in different experimental outcomes for the two groups. If so, then we will elicit your judgments separately for lower SES children and for middle and upper SES children.

C.1.4 Exposure Conditions

We will be asking your judgment about mean IQ values on the seventh birthdays at various blood-lead levels. Assume that the blood-lead levels under consideration for a given judgment have been measured on the children's third birthdays. Assume further that external environmental conditions supporting those levels have been more or less constant since birth, and that in interacting with the environment, the children exhibited the usual range of behaviors at each age. Thus, blood-lead levels were not necessarily constant from birth to age seven, but exposure and behavioral factors were such that at age three, blood lead was at a specific level. Finally, assume that the changes in blood lead levels from birth until the seventh birthday are distributed as you believe they actually are.

C.1.5 Physiological and Environmental Conditions

Because the effects of lead depend on many parameters, it is necessary to specify assumptions about those parameters in the population(s).

C.1.5.1 Physiological Conditions

The effect of lead in the system depends on the person's nutritional and metabolic status. Assume that the levels of iron, zinc, copper, vitamins A, C, D, and E, calcium, phosphorus, and magnesium are distributed within the population or within each of the two SES groups as you believe they in fact are, taking into account the wide range of diets and nutritional levels of children within the population or within each of the two SES groups.

C.1.5.2 Environmental Conditions

The effect of lead on IQ also may depend on numerous environmental and caregiving factors that vary within SES level, such as those assessed in the HOME scale (Home Observation for Measurement of the Environment), parental IQ, and so forth. Assume that these factors are distributed within the population or within each of the two SES groups as you believe they in fact are.

C.1.6 Factors to Consider

In order to help you bring to mind the relevant evidence so that you may consider it systematically, and also in order to help us to interpret your judgments, we would like to ask you to discuss briefly your interpretations of various aspects of the literature. How do you evaluate the research concerning the effects of low level lead exposure on cognitive and behavioral development? Is it your feeling that lead has a deleterious effect on this development over and above that which can be explained by other factors frequently associated with lead exposure? If so, do you believe that the effects of lead simply add to the effects of other factors, or that the effects of lead depend on the levels of other factors? In the latter case, what are the variables that lead exposure interacts with? A related question concerns your opinion about the relationship between exposure and susceptibility. That is, do you think that those children who are at greatest

risk of exposure due to their living in deteriorating pre-1950 housing or in urban areas with large amounts of vehicular traffic or due to other reasons, are also most susceptible to the effects of lead? For example, these may be the same children who also have poorer diets, less access to medical care, poorer care-giving environments, and less intellectual resources to fall back on.

What is your opinion about the time course of lead exposure and of lead effects on CNS, cognitive and behavioral development? What are the relative effects of cumulative versus current exposure. Are the effects reversible? Is there a threshold for the effects? What are the implications of the animal model research, both the behavioral and the morphological, for human CNS, cognitive, and behavioral effects of lead? Are there other factors to consider in thinking about the dose-response functions for lead-induced IQ effects that we should discuss now?

C.1.7 Factors to Keep in Mind When Making Probability Judgments

There is usually uncertainty associated with conclusions that we draw from research, and more generally in our everyday thinking. However, not everyone is aware of all the sources that should contribute to their uncertainty, nor are most people familiar with the process of actually expressing their uncertainty in probabilistic terms. When an expert is asked to make probability judgments on socially important matters, it is particularly important that he or she consider the relevant evidence in a systematic and effective manner and provide judgments that represent his or her opinions well.

Experimental psychologists and decision analysts have amassed a considerable amount of data concerning the way people form and express probabilistic judgments. The evidence suggests that when considering large amounts of complex information, most people employ simplifying heuristics and demonstrate certain systematic distortions of thought, i.e., cognitive biases, which adversely affect their judgments. The purpose of this section is to make you aware of these biases and heuristics so that, as much as

possible, you can avoid them in making probability judgments. We will first review the most widespread biases and heuristics, and then offer some suggestions to help you mitigate their effects.

C.1.7.1 Sequential Consideration of Information

Generally, the order in which evidence is considered influences the final judgment, although logically that should not be the case. Of necessity, pieces of information are considered one by one in a sequential fashion. However, those considered first and last tend to dominate judgment. In part, initial information has its undue influence because it provides the framework that subsequent information is then tailored to fit. For example, people usually search for evidence to *confirm* their initial hypotheses; they rarely look for evidence that *weighs against* them. The later evidence has its undue effect simply because it is fresher in memory.

Related to these sequential effects is the phenomenon of *anchoring and adjustment*. Based on early partial information, one forms an initial probability estimate regarding the event in question. This anchor judgment is then adjusted as subsequent information is considered. Unfortunately, such adjustments tend to be too conservative. In other words, too little weight is attached to information considered subsequent to the formation of the initial judgment.

C.1.7.2 Effects of Memory on Judgment

It is difficult for most people to conceptualize and make judgments about large, abstract universes or populations. A natural tendency is to recall specific members and then to consider them representative of the population as a whole. However, the specific instances often are recalled precisely because they stand out in some way, such as being familiar, unusual, especially concrete, or of personal significance. Unfortunately, the specific characteristics of these singular examples are then attributed, often incorrectly,

to all the members of the population of interest. Moreover, these memory effects are often combined with the sequential phenomena discussed earlier. For example, in considering the evidence regarding the dose-response curve of a particular pollutant, you might naturally first think of a study you or a personal friend recently completed. Or you might think of a study you recently read, or one that was unusual and therefore stands out. The tendency might then be to treat the recalled studies as typical of the population of relevant research, ignoring important differences among studies. Subsequent attempts to recall information could result in thinking primarily of evidence consistent with the initial items you thought of.

C.1.7.3 Estimating Reliability of Information

People tend to overestimate the reliability of information, ignoring factors such as sampling error and imprecision of measurement. Rather they summarize evidence in terms of simple and definite conclusions, causing them to be overconfident in their judgments. This tendency is stronger when one has a considerable amount of intellectual and/or personal involvement in a particular field. In such cases, information is often interpreted in a way which is consistent with one's beliefs and expectations, results are over-generalized, and contradictory evidence is ignored or underestimated.

C.1.7.4 Relation Between Event Importance and Probability

Sometimes the importance of events, or their possible costs or benefits, influence judgments about the certainty of the events when, rationally, importance should not affect probability. In other words, one's attitudes towards risk tend to affect one's ability to make accurate probability judgments. For example, many physicians tend to overestimate the probability of very severe diseases, because they feel it is important to detect and treat them; and similarly, many smokers underestimate the probability of adverse consequences of smoking, because they feel that the odds do not apply to themselves personally.

C.1.7.5 Estimation of Probabilities

Another limitation is related to one's ability to discriminate between levels of uncertainty, and to use appropriate criteria of discrimination for different ranges of probability. One result of this fact is that people tend to estimate both extreme and mid-range probabilities in the same fashion, usually doing a poor job in the extremes. It helps here to think in terms of odds as well as probabilities. Thus, for example, changing a probability estimate from 0.510 to 0.501 is equivalent to a change in odds from 1.041:1 to 1.004:1, but a change from an estimate of 0.999 to 0.990 changes the odds by a factor of about 10 from 999:1 to 99:1 The closer to the extremes (either 0 or 1) that one is estimating probabilities, the greater the impact of small changes.

C.1.7.6 Recommendations

Although extensive and careful training would be necessary to eliminate all the problems mentioned above, some relatively simple suggestions can help minimize them. Most important is to be aware of one's natural cognitive biases, and to try consciously to avoid them.

To avoid *sequential effects* keep in mind that the order in which you think of information should not influence your final judgment. It may be helpful to actually note on paper the important facts you are considering and then to reconsider them in two or more sequences, checking the consistency of your judgments. Try to keep an open mind until you have gone through all the evidence, and don't let the early information you consider sway you more than is appropriate.

To avoid adverse *memory effects*, define various classes of information that you deem relevant, and then search your memory for examples of each. Don't restrict your thinking only to items that stand out for specific reasons. Make a special attempt to consider conflicting evidence, and to think of data that may be inconsistent with a particular theory. Also, be careful to concentrate on the given probability judgment and

do not let your own values (how you would make the decision yourself) affect those judgments.

To accurately *estimate the reliability of information*, pay attention to such matters as sample size and power of the statistical tests. Keep in mind that data are probabilistic in nature, subject to elements of random error, imprecise measurements, and subjective evaluation and interpretation. In addition, the further one must extrapolate, or generalize, from a particular study to a situation of interest, the less reliable is the conclusion and the less certainty should be attributed to it. Rely more heavily on information which you consider more reliable, but do not treat it as "absolute truth."

Keep in mind that *the importance of an event or an outcome should not influence its judged probability*. It is rational to let the costliness or severity of an outcome influence the point at which action is taken with respect to it, but not the judgment that is made about the outcome's likelihood.

Finally, in making probability judgments, think primarily in terms of the measure (probability or odds) with which you feel more comfortable, but sometimes translate to the alternative scale, or even to measures of other events (e.g., the probability of the event not happening). When estimating very small or very large likelihoods, it is usually best to think in terms of odds, which are unbounded, instead of probabilities, which are bounded. For example, one can more easily conceptualize odds of 1:200 than a probability of .005.

C.1.8 Final Preparation for Elicitation of Probability Judgments

The outcomes of the ideal experiments described above are uncertain. Existing data are relevant, but do not allow exact predictions. Our goal is to have you represent probabilistically your own uncertainty about the experimental outcomes based on your expertise and the available knowledge. In responding to the questions we will ask you,

please think carefully about the relevant information reviewed in the criteria document, and consult it as well as the literature or your files as you deem appropriate.

The previous section suggests ways to think about the relevant data. The purpose of that section is to help minimize the biasing effects that frequently accompany the information overload naturally resulting from rapid consideration of large amounts of complex evidence. You may find it helpful to review the section or raise questions about the points made in it before we begin.

Uncertainty about the effects of lead exposure on mean decrement in IQ (relative to mean IQ at zero blood lead) can be represented probabilistically in two different ways.

1. Uncertainty about the mean IQ decrement that would result when the exposed group has a *given blood-lead concentration.*
2. Uncertainty about the blood-lead concentration that would be required to *cause a given mean IQ* decrement.

We will concentrate on one way at a time, focusing primarily on the first one.

For these purposes recall that everyone in the exposed population has a specified blood-lead level while everyone in the zero blood-lead population has negligible blood lead, as described above under Exposure Conditions. Then, in order for us to determine your uncertainty about the mean IQ decrement of the exposed group with a *given blood-lead level* relative to the negligible blood-lead group, we must introduce a definition. Let L be the given blood-lead concentration in question.

Definition: D(L) is the mean IQ decrement of the exposed group with blood-lead level L relative to the negligible group.

The value of D(L) for a given L is uncertain, and we would like to obtain probability judgments from you about its possible values. We will elicit your judgments about the possible values of D(L) by specifying a particular mean IQ decrement d and having you consider how likely it is that D(L) is less than that value. To help you make

your probability judgments, we will make use of a device called a probability wheel, which has adjustable sectors of blue and orange. We can read on the back of the wheel the percentage of the wheel that is each color. In making your judgments you are to imagine that the wheel is a perfectly fair random device, and that therefore the probability of its stopping with the pointer on blue is exactly represented by the relative area that is blue.

We will then proceed as follows. For each blood lead concentration L, we will specify a particular mean IQ decrement d, and also set the probability wheel to have a specific relative area of blue. Then you are to consider carefully the *question*:

> Do you consider it more probable that D(L) is less than d *or* that the wheel would stop with the pointer on blue (on a random spin)?

You can give one of three responses:

1. You judge it to be more probable that D(L) is less than d;
2. You judge it to be more probable that the wheel would stop with the pointer on blue;
3. You cannot judge either event as more probable than the other.

For a particular concentration L and mean IQ decrement d, some wheel settings will have a small enough relative area of blue that you will feel confident making response (1). Other wheel settings will have a large enough relative area of blue that you will feel confident making response (2). The intermediate settings will be more difficult to judge. However, for a fixed blood lead L, and mean IQ decrement d, we will manipulate the wheel settings to find the one for which you feel most comfortable making response (3).

Once we have determined the point at which you are most comfortable with response (3), and still focusing on the given blood lead L, we will specify a new mean IQ decrement d', reset the wheel to a new relative area of blue, and proceed to obtain your judgments regarding the probabilistic relation between d' and D(L). This will continue for

the given L until we have obtained comparisons of various mean values to D(L). Then we will have elicited the probabilistic representation of your uncertainty about the mean IQ decrement for one blood-lead level. We will obtain the probabilistic representation of your uncertainty for the outcome of another experiment by specifying a new L and continuing as before. We will do this for a number of values of L.

It frequently happens that an expert's judgments alter somewhat over the course of a session such as this, as he or she considers the evidence from various perspectives and thinks about the various responses called for. Hence, we will graph your responses and, at appropriate times, show them to you for your consideration and comparison. At these times you may wish to change some of the judgments you give earlier.

Also, although the probability judgments are entirely your own, we must introduce one logical constraint: namely, that your *final* judgments are coherent in a sense that we can explain as we go along. It is quite common for initial judgments to exhibit some incoherencies. That is another reason we will consider together the graphs of your judgments. Our objective is to obtain a coherent set of judgments that represent your opinions well by the end of this elicitation process. You are not expected to give us such a set immediately. All of this will become clearer as we go along.

We should emphasize that the judgments we are asking you to make are not simple ones, nor of course are there known correct answers. Rather, we want your best and most considered judgment in light of the available relevant scientific data. Therefore, please reflect on the available data carefully, feeling free to consult the lead criteria document or other sources as you wish as you formulate your judgments.

(Encode judgments for two blood-lead concentrations, then continue with instructions.)

Recall that uncertainty about the experimental outcomes can also be expressed in terms of the blood-lead concentration necessary to produce a given mean IQ decrement D, if the entire exposed population had the same blood-lead concentration,

established in the manner discussed earlier. More specifically, proceeding as before, let us introduce a definition:

Definition: L(D) is the blood-lead level that would cause a mean IQ decrement D in an exposed group relative to the negligible group.

The value of L(D) is uncertain for a given D, and we want to obtain your judgments about its possible values. Analogously to what we have already done in obtaining the D(L) judgments, for a given D we will specify concentrations ℓ and ask you to consider how likely it is that the L(D) is less than the specified ℓ. More specifically, utilizing the wheel as before, we will ask you the *question*:

Do you consider it more likely that L(D) is less than ℓ *or* that the wheel would stop with the pointer on blue (on a random spin)?

You can give one of three responses:

1. You judge it more probable that L(D) is less than ℓ;
2. You judge it more probable that the wheel would stop with the pointer on blue;
3. You cannot judge either event as more probable than the other.

After determining the wheel setting at which you feel most comfortable with response (3), we will specify a new ℓ and repeat the process. As before, for each D we will elicit your comparative judgments for various values of ℓ.

We must encode your probabilistic judgment about the mean IQ of children in the population, or in the two SES groups, who have negligible blood-level under the exposure conditions described previously. We start with a definition:

Definition: Let μ_o^* be the true population mean IQ for children with negligible blood lead. (Or, let μ_{oi}^* be the true population mean IQ for children with negligible blood lead in SES group i.)

The value of μ_o^* or μ_{oi}^* is uncertain, and we will obtain your judgment about its possible values. Analogously to what we have already done, we will specify a mean u and

ask you to consider how likely it is that μ_o^* or μ_{oi}^* is less than μ. We will elicit your judgments about various values of μ with the aid of the probability wheel, in the same manner as was done previously.

If you felt it necessary to consider separate experiments for two SES groups, then you may believe that the within-group IQ standard deviation is different from 15. If so, we must encode your probabilistic judgment about that value. If you agree, we will assume that the standard deviation is unaffected by lead-level. However, you may feel that the standard deviation is different for the two SES groups. If so, we must encode your opinion separately about each case.

As before, we start with a definition:

Definition: Let σ_i^* be the true IQ standard deviation for SES group i.

The value of σ_i^* is uncertain, and we will obtain your judgment about its possible values. Analogously to what we have already done, we will specify a standard deviation σ and ask you to consider how likely it is that σ_i^* is less than σ. We will elicit your judgments about various values of σ with the aid of the probability wheel, in the same manner as was done previously.

C.2 SUMMARY OF FUNCTIONS FIT TO IQ JUDGMENTS

Normal-on-log-odds (NOLO) probability functions were fit to all of the judgments of the hemoglobin experts. Because of the diverse nature of the judgments obtained from the IQ experts, it was necessary to choose among normal, lognormal, and NOLO probability functions to adequately represent those judgments. Fitting normal and lognormal distributions to judgments are intermediate steps in the process of fitting a NOLO distribution, so it is unnecessary to repeat those details here. The reader can refer to Sec. B.2 for all of the necessary information.

In fitting lognormal distributions to judgments, the variable of interest is that designated X in Sec. B.2.2; for fitting normal distributions, the variable of interest is

that designated Y. Lognormal distributions were found to be particularly useful in fitting functions to judgments about mean IQ levels of children sheltered from lead exposure, and normal distributions were particularly useful in fitting functions to judgments about population standard deviation.

C.3 RESULTS

Results are presented in seven tables which cover:

- Encoded judgments for the mean IQ, $E[\overline{IQ}_o]$ of the distribution over IQ of children sheltered from lead (Table C.1),
- Encoded judgments for within-group standard deviation, σ_{IQ} (Table C.2),
- Encoded judgments for mean IQ decrements, $\Delta_{\overline{IQ}}$, for exposed children relative to unexposed children (Table C.3),
- Specifications (form, parameters, moments) for functions fit to the judgments for $\Delta_{\overline{IQ}_o}$ (Table C.4),
- Specifications for functions fit to judgments for σ_{IQ} (Table C.5),
- Specifications for functions fit to judgments for $\Delta_{\overline{IQ}}$ (Table C.6), and
- Comparisons of medians and credible intervals for encoded judgments vs. fitted functions (Table C.7).

C.3.1 Expert F

Expert F did not feel that SES level was a significant factor influencing the effects of lead on IQ. Thus, only one set of judgments was obtained from Expert F. In addition, Expert F did not feel that it was necessary to use a probability distribution to characterize population standard deviation and simply gave a point estimate for that quantity.

TABLE C.1 Encoded Judgments about the Mean IQ of Children Sheltered from Lead Exposure

Expert F		Expert G		Expert H		Expert J		Expert K	
Both SES Levels		Lower SES Levels							
$\overline{IQ_o}$	F	$\overline{IQ_o}$	F	$\overline{IQ_o}$	F	$\overline{IQ_o}$	F	$\overline{IQ_o}$	F
100	0.01	90	0.01	92	0.01	90.0	0.01	78	0.01
100	0.01	93	0.25	94	0.20	93.5	0.25	83	0.25
101	0.42	94	0.50	98	0.50	95.0	0.50	85	0.50
101	0.80	96	0.76	99	0.65	96.0	0.75	88	0.75
102	0.99	98	0.92	102	0.99	98.0	0.90	91	0.99
		100	0.99						
		Higher SES Levels							
		$\overline{IQ_o}$	F	$\overline{IQ_o}$	F	$\overline{IQ_o}$	F	$\overline{IQ_o}$	F
		100	0.01	100	0.02	100	0.001	100	0.01
		104	0.15	102	0.11	102	0.1	102	0.17
		106	0.50	104	0.30	104	0.77	104	0.42
		108	0.85	106	0.49	106	0.90	106	0.60
		110	0.99	108	0.75	108	0.99	108	0.85
				110	0.96	110	0.999	110	0.99

TABLE C.2 Encoded Judgments about Population Standard Deviation

Expert F		Expert G		Expert H		Expert J		Expert K	
Both SES Levels		Lower SES Levels							
$\overline{\sigma_{IQ}}$	F	$\overline{\sigma_{IQ}}$	F	$\overline{\sigma_{IQ}}$	F	$\overline{\sigma_{IQ}}$	F	$\overline{\sigma_{IQ}}$	F
15	1	14	1	10	0.04	11	0.001	10	0.01
				11	0.1	12	0.125	11	0.05
				12	0.35	13	0.5	12	0.14
				13	0.5	14	0.875	13	0.53
				14	0.85	15	0.999	14	0.93
				15	0.999			15	0.999
		Higher SES Levels							
		14	1	13	0.04	11	0.001	12	0.2
				13.5	0.07	12	0.125	13	0.25
				14	0.275	13	0.5	13.5	0.5
				14.3	0.5	14	0.875	14	0.75
				15	0.9	15	0.999	15	0.99
				15.5	0.99				

TABLE C.3 Encoded Judgments about Mean IQ Decrements of Children Exposed to Lead

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert F, for Both SES Levels											
25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL		65 µg/dL			
0.25	0.5	0.5	0.5	0.9	0.5	1	0.2	1	0.150		
1.0	0.999	2.0	0.999	1.0	0.55	2.0	0.8	2.0	0.500		
				2.0	0.93	3.0	0.95	3.0	0.900		
				3.0	0.99	4.0	0.99	4.0	0.960		
								5.0	0.990		
Expert G, Lower SES Level											
5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
0.2	0.01	0.5	0.01	1.0	0.01	2.0	0.01	3.5	0.01	14.0	0.010
0.3	0.25	0.7	0.25	1.3	0.25	2.5	0.25	4.0	0.25	5.0	0.250
0.5	0.5	1.5	0.5	3.0	0.5	3.5	0.5	5.0	0.5	7.0	0.500
1.0	0.75	2.0	0.75	4.0	0.75	5.0	0.75	7.0	0.75	10.0	0.750
3.0	0.99	5.0	0.99	8.5	0.99	10.0	0.99	12.0	0.99	15.0	0.990

TABLE C.3 (Cont'd)

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert H, Lower SES Level											
5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
1.0	0.03	2.0	0.06	2.0	0.01	4.0	0.02	6.0	0.02	8.0	0.020
2.0	0.4	3.0	0.5	4.0	0.35	6.0	0.25	8.0	0.35	1.0	0.400
3.0	0.55	4.0	0.6	6.0	0.6	8.0	0.46	1.0	0.55	12.0	0.770
4.0	0.9	5.0	0.75	8.0	0.84	1.0	0.79	12.0	0.85	14.0	0.900
5.0	0.99	6.0	0.93	1.0	0.99	12.0	0.95	14.0	0.95	16.0	0.950
Expert I, Lower SES Level											
15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL			
0.001	0.025	1.0	0.025	2.0	0.025	3.0	0.025	4.0	0.025		
0.5	0.25	1.9	0.25	3.1	0.25	4.5	0.25	5.5	0.25		
0.7	0.5	2.3	0.5	3.6	0.5	5.3	0.5	6.3	0.5		
1.1	0.75	3.2	0.75	5.1	0.75	6.9	0.75	8.2	0.75		
2.0	0.975	5.0	0.975	7.0	0.975	10.0	0.975	12.0	0.975		

TABLE C.3 (Cont'd)

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert J, Lower SES Level											
5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
0.001	0.0001	0.002	0.001	0.003	0.001	2.0	0.001	4.0	0.001	6.0	0.001
1.0	0.25	2.25	0.25	4.0	0.25	6.0	0.25	7.5	0.25	9.0	0.250
2.0	0.5	3.5	0.5	5.0	0.5	7.0	0.5	9.0	0.5	11.0	0.500
3.0	0.75	4.5	0.75	5.75	0.75	8.0	0.75	10.5	0.75	12.5	0.750
5.0	0.995	6.5	0.995	8.0	0.99	1.0	0.99	12.0	0.99	14.0	0.990
Expert K, Lower SES Level											
15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL		65 µg/dL	
0.25	0.01	0.5	0.01	1.0	0.01	4.0	0.02	5.0	0.02	7.0	0.010
1.0	0.36	1.0	0.06	2.0	0.05	5.0	0.11	6.0	0.1	8.0	0.090
1.5	0.5	2.0	0.27	3.0	0.18	6.0	0.5	7.0	0.28	1.0	0.500
2.0	0.97	3.0	0.47	4.0	0.5	7.0	0.64	8.0	0.47	12.0	0.790
		4.0	0.8	5.0	0.86	8.0	0.82	9.0	0.82	14.0	0.990
		5.0	0.98	6.0	0.99	9.0	0.99	10.0	0.970		

TABLE C.3 (Cont'd)

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert G, Higher SES Level											
5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
		0.3	0.01	0.5	0.01	1.0	0.01	2.0	0.01	3.5	0.010
		0.4	0.25	0.7	0.25	1.5	0.25	2.5	0.25	4.0	0.250
		0.5	0.5	1.0	0.5	2.5	0.5	3.5	0.5	5.0	0.500
		1.0	0.75	2.0	0.75	4.0	0.75	5.0	0.75	7.5	0.750
		3.0	0.99	4.0	0.99	8.0	0.99	1.0	0.99	12.0	0.990
Expert H, Higher SES Level											
5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
0.5	0.45	1.0	0.03	2.4	0.02	3.8	0.04	4.8	0.01	6.0	0.010
1.0	0.8	2.0	0.3	3.4	0.13	4.8	0.3	5.8	0.09	7.0	0.100
1.5	0.98	2.5	0.5	4.4	0.47	5.5	0.5	6.5	0.5	8.0	0.500
		3.0	0.6	4.5	0.5	5.8	0.61	6.8	0.7	9.0	0.850
		3.5	0.73	5.4	0.7	6.8	0.86	7.8	0.85	10.0	0.990
		4.0	0.93	6.4	0.93	8.3	0.99	8.8	0.990		
				7.4	0.99						

TABLE C.3 (Cont'd)

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert I, Higher SES Level											
15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL			
0.2	0.25	0.7	0.25	1.0	0.025	2.0	0.025	2.0	0.025		
0.3	0.5	1.0	0.5	1.9	0.25	3.1	0.25	3.5	0.25		
0.5	0.75	2.0	0.75	2.3	0.5	3.6	0.5	4.3	0.5		
1.0	0.975	3.0	0.975	3.2	0.75	4.7	0.75	5.8	0.75		
				5.0	0.975	7.0	0.975	9.0	0.975		
Expert J, Higher SES Level											
15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL			
0.5	0.014	1.0	0.01	2.0	0.05	3.00	0.005	5.0	0.010		
1.0	0.14	2.75	0.25	3.5	0.25	4.5	0.25	6.5	0.250		
1.25	0.25	3.0	0.5	4.0	0.5	5.0	0.5	7.0	0.500		
2.0	0.5	3.5	0.75	5.0	0.75	5.75	0.75	7.5	0.750		
2.3	0.75	7.0	0.999	7.0	0.999	8.0	0.99	10.0	0.990		
3.0	0.95										
4.0	0.99										
6.0	0.999										

TABLE C.3 (Cont'd)

Blood-Lead Level											
$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F	$\bar{\Delta}_{IQ}$	F
Expert K, Higher SES Levels											
15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL		65 µg/dL	
		0.5	0.01	1.0	0.01	2.0	0.01	3.0	0.02	4.0	0.010
		1.0	0.35	1.5	0.11	3.0	0.2	4.0	0.25	5.0	0.170
		1.5	0.5	2.0	0.5	3.5	0.5	4.5	0.5	6.0	0.500
		2.0	0.99	2.5	0.9	4.0	0.75	5.0	0.78	7.0	0.820
				3.0	0.99	5.0	0.99	6.0	0.98	8.0	0.990

[a] $\bar{\Delta}_{IQ}$ denotes the mean IQ decrement among children with the specified blood-lead levels, referenced to unexposed children.

[b] F denotes cumulative probability.

TABLE C.4 Parameters of Functions Fit to the Probabilistic Judgments of Experts Regarding Mean IQ Levels among Children Sheltered from Lead Exposure

Expert	SES Case	Functional Form	Parameters of Normal-on-Ln (IQ_0) Distribution		Moments of Distribution Over IQ_0		r^2 of $\hat{F}$ on F
			Mean	SD[a]	$E[\overline{IQ}_0]$[b]	$SD[\overline{IQ}_0]$[c]	
F	Combined	Lognormal	4.6140	0.0046	100.9	0.47	0.931
G	Lower	Lognormal	4.5495	0.0232	94.6	2.20	0.994
H	Lower	Lognormal	4.5763	0.0249	97.2	2.42	0.966
J	Lower	Normal	-	-	94.8	1.93	0.996
K	Lower	Normal	-	-	85.0	2.92	0.993
G	Higher	Lognormal	4.6591	0.0206	105.6	2.17	0.987
H	Higher	Lognormal	4.6596	0.0265	105.6	2.80	0.992
J	Higher	Lognormal	4.6454	0.0158	104.1	1.65	0.960
K	Higher	Normal	-	-	104.9	2.36	0.988

[a] SD denotes standard deviation.

[b] $E[\overline{IQ}_0]$ denotes the mean of the distribution for $\overline{IQ}_0$.

[c] $SD[\overline{IQ}_0]$ denotes the standard deviation of the distribution for $\overline{IQ}_0$.

TABLE C.5 Summary of Functions Fit to the Probabilistic Judgments of Experts Regarding Population Standard Deviation for IQ Levels

Expert	SES Case	Functional Form	Parameters of Normal-on-Ln (σ_{IQ}) Distribution		Moments of Distribution Over σ_{IQ}		r^2 of $\hat{F}$ on F
			Mean	SD[a]	$E[\sigma_{IQ}]$[b]	$SD[\sigma_{IQ}]$[c]	
F	Combined	Point	-	-	15.0	-	-
G	Lower	Point	-	-	14.0	-	-
H	Lower	Lognormal	2.5056	0.0915	12.3	1.13	0.999
J	Lower	Normal	-	-	13.0	0.68	0.999
K	Lower	Normal	-	-	12.6	0.93	0.999
G	Higher	Point	-	-	14.0	-	-
H	Higher	Lognormal	2.6541	0.0416	14.2	0.59	0.999
J	Higher	Lognormal	2.5590	0.0531	12.9	0.69	0.999
K	Higher	Normal	-	-	13.5	0.69	0.999

[a] SD denotes standard deviation.

[b] $E[\sigma_{IQ}]$ denotes the mean of the distribution of σ_{IQ}.

[c] $SD[\sigma_{IQ}]$ denotes the standard deviation of the distribution for σ_{IQ}.

TABLE C.6 Summary of Function Fit to the Probabilistic Judgments of Experts Regarding Mean IQ Decrements among Children Exposed to Lead

Blood-Lead Level (μg/dL)	Functional Form	Parameters of Normal on LN ($\Delta_{\overline{IQ}}$) Distribution		Moments of Distribution Over $\Delta_{\overline{IQ}}$		r^2 of $\hat{F}$ on F
		Mean	SD[a]	$E[\Delta_{\overline{IQ}}]$[b]	$SD[\Delta_{\overline{IQ}}]$[c]	
Expert F, Both SES Levels						
25	Lognormal	-1.4059	0.4628	0.2729	0.1334	0.9990
35	Lognormal	-0.7127	0.4628	0.5457	0.2667	0.9990
45	Lognormal	-0.0329	0.4628	1.0770	0.5264	0.9986
55	Lognormal	0.3350	0.4628	1.5559	0.7604	0.9984
65	Lognormal	0.5574	0.4628	1.9434	0.9498	0.9885
Expert G, Lower SES Levels						
5	Lognormal	-0.4816	0.6422	0.7593	0.5425	0.9789
15	Lognormal	0.3316	0.5429	1.6145	0.9452	0.9789
25	Lognormal	0.9775	0.5229	3.0471	1.7089	0.9689
35	Lognormal	1.3548	0.3807	4.1675	1.6459	0.9825
45	Lognormal	1.7359	0.3011	5.9369	1.8287	0.9701
55	Lognormal	1.9905	0.3174	7.6970	2.5060	0.9782
Expert H, Lower SES Levels						
5	Lognormal	0.8268	0.4087	2.4851	1.0596	0.9715
15	Lognormal	1.2481	0.3989	3.7724	1.5669	0.9765
25	Lognormal	1.5860	0.3749	5.2397	2.0356	0.877
35	Lognormal	2.0323	0.3087	8.0040	2.5307	0.9879
45	Lognormal	2.2422	0.2366	9.6807	2.3224	0.9898
55	Lognormal	2.4045	0.1887	11.2714	2.1458	0.9888
Expert I, Lower SES Levels						
15	Lognormal	-0.3276	0.5351	0.8316	0.4789	0.9984
25	Lognormal	0.8495	0.4110	2.5445	1.0915	0.9945
35	Lognormal	1.3361	0.3282	4.0148	1.3540	0.9926
45	Lognormal	1.7009	0.3088	5.7465	1.8178	0.9980
55	Lognormal	1.9041	0.2845	6.9908	2.0298	0.9949

TABLE C.6 (Cont'd)

Blood-Lead Level (μg/dL)	Functional Form	Parameters of Normal on LN ($\Delta_{\overline{IQ}}$) Distribution		Moments of Distribution Over $\Delta_{\overline{IQ}}$		r^2 of $\hat{F}$ on F
		Mean	SD[a]	$E[\Delta_{\overline{IQ}}]$[b]	$SD[\Delta_{\overline{IQ}}]$[c]	
Expert J, Lower SES Levels						
5	Normal	-	-	2.3874	0.8956	0.9608
15	Normal	-	-	3.4704	1.1878	0.9907
25	Normal	-	-	4.7747	1.4872	0.9927
35	Normal	-	-	6.8264	1.5025	0.9949
45	Normal	-	-	8.8357	1.5642	0.9889
55	Normal	-	-	10.7412	1.6007	0.9791
Expert K, Lower SES Levels						
15	Normal	-	-	1.2764	0.4426	0.9540
25	Normal	-	-	2.8910	1.1027	0.9913
35	Normal	-	-	3.7668	1.0816	0.9900
45	Normal	-	-	6.4349	1.2201	0.9800
55	Normal	-	-	7.7608	1.3071	0.9889
65	Normal	-	-	10.3694	1.5855	0.9920
Expert G, Higher SES Levels						
15	Lognormal	-0.3430	0.5671	0.8335	0.5133	0.9546
25	Lognormal	0.2059	0.5045	1.3954	0.7512	0.9668
35	Lognormal	0.9575	0.4847	2.9299	1.5079	0.9854
45	Lognormal	1.3548	0.3807	4.1675	1.6459	0.9825
55	Lognormal	1.7497	0.3068	6.0298	1.8942	0.9613
Expert H, Higher SES Levels						
5	Lognormal	-0.5767	0.5207	0.6433	0.3590	0.9961
15	Lognormal	0.8960	0.4498	2.7106	1.2837	0.9833
25	Lognormal	1.4785	0.2607	4.5382	1.2035	0.9899
35	Lognormal	1.6873	0.1955	5.5090	1.0873	0.9968
45	Lognormal	1.8859	0.1297	6.648a	0.8661	0.9842
55	Lognormal	2.0688	0.1103	7.9635	0.8809	0.9967

REFERENCES

APPENDIX A

FITTING FUNCTIONS TO DATA ON LEAD-INDUCED ELEVATED ERYTHROCYTE PROTOPORPHYRIN LEVELS

Section 2 discusses published data (Piomelli et al., 1982) about dose-response relationships between PbB level and EP level. Because the data are complete and reliable, probability encoding of the judgments of experts was not necessary. The available data, when combined with sample-size data, are sufficient to develop probabilistic dose-response functions. Those functions are specified in this appendix.

Figure 2.1 presents published dose-response data about lead-induced elevated EP levels, and Table 2.1 presents sample-size data. For the two EP levels (33 μg/dL and 53 μg/dL) addressed, expressions can be determined for the mean dose-response curves shown.

$$R_{EP} = \begin{cases} 0.107, & \text{for } L \leq 16.6 \text{ μg/dL} \\ \Phi^{-1}(0.10348 \cdot (L-16.6) - 1.22, & \text{for } L > 16.6 \text{ μg/dL} \end{cases}$$

for EP levels greater than or equal to 33 μg/dL, and

$$R_{EP} = \begin{cases} 0.024, & \text{for } L \leq 16.6 \text{ μg/dL} \\ \Phi^{-1}(0.10348 \cdot (L-16.6) - 2), & \text{for } L > 16.6 \text{ μg/dL} \end{cases}$$

for EP levels greater than or equal to 53 μg/dL,

where:

R = population response rate (fraction of children having EP levels greater than or equal to a specific EP level),

L = PbB level (μg/dL), and

$\Phi^{-1}(\cdot)$ = inverse of the cumulative distribution function of a standardized normal random variable.

Piomelli et al.'s (1982) data indicate a threshold for EP effects at a PbB level about 16.6

µg/dL. Further, response rates of 0.024 and 0.107 were found to be "natural frequencies" for the occurrence of EP levels above 53 µg/dL and 33 µg/dL, respectively.

Probability distributions over the population response rate determined for 10 (PbB, EP) level combinations are summarized in Table A.1. These distributions are either beta

$$\beta(R) = \frac{(N-1)!}{(X-1)!(N-X-1)!} R^{X-1} (1-R)^{N-X-1}$$

or normal at each PbB level, where N is the number of children at a specific PbB level, and X is the number of children having EP levels greater than or equal to either 33 µg/dL or 53 µg/dL. The cumulative probability function for the beta distribution is

$$B(R) = 1 - \sum_{i=o}^{X-1} \binom{N-1}{i} R^i (1-R)^{N-1-i}$$

The mean E(R) and variance V(R) of the beta distribution and, where appropriate, its normal approximation are

$$E(R) = \frac{X}{N} \quad \text{and} \quad V(R) = \frac{X(N-X)}{N^2(N+1)}$$

For most PbB and EP combinations, the quantity NR(1-R) was large enough to allow use of the normal approximation.

TABLE A.1 Probability Distributions and Parameters for Piomelli's Data on Erythrocyte Protoporphyrin (EP) Levels among New York City Children

PbB (μg/dL)	EP ≥ 33 μg/dL			EP ≥ 53 μg/dL		
	Functional Form	Parameter Values		Functional Form	Parameter Values	
5	Normal	$E(R)^a=0.107$	$V(R)^b=0.029^2$	Beta	X=3	N=128
15	Normal	E(R)=0.107	$V(R)=0.009^2$	Normal	E(R)=0.024	$V(R)=0.004^2$
25	Normal	E(R)=0.337	$V(R)=0.020^2$	Normal	E(R)=0.124	$V(R)=0.014^2$
35	Normal	E(R)=0.728	$V(R)=0.043^2$	Normal	E(R)=0.449	$V(R)=0.047^2$
50	Beta	X=42	N=43	Beta	X=40	N=43

[a]E(R) denotes the expected value of R.

[b]V(R) denotes the variance of R.

Source: Data are from Piomelli et al. (1982).

APPENDIX B

HEMOGLOBIN

Appendix B is organized as follows. Section B.1 reproduces the protocol used during the hemoglobin encoding sessions. Section B.2 details the mathematical formulations for the dose-response models. Section B.3 tabulates encoded judgments, specifications for functions fit to those judgments, and credible intervals for encoded judgments and fitted functions. Section B.4 summarizes the discussions held with each of the interviewed experts.

B.1 HEMOGLOBIN PROTOCOL

B.1.1 Introduction

The Environmental Protection Agency (EPA) is charged by the Clean Air Act with setting and revising National Ambient Air Quality Standards (NAAQS) for selected pollutants, at levels sufficient to protect the public health with an adequate margin of safety. As you know, the scientific bases for NAAQS are presented and reviewed in criteria documents. In support of the forthcoming review of the lead NAAQS, EPA has just prepared a new *Air Quality Criteria for Lead*. It presents scientific evidence from which the most susceptible populations can be determined and from which various adverse health effects can be identified. The criteria document summarizes and evaluates the available clinical, epidemiological, and animal or toxicological laboratory evidence with regard to the physiological and adverse health effects of lead, and therefore represents our most up to date knowledge on lead effects.

As one aspect of the review process, EPA assesses health risks by identifying the most sensitive populations for each pollutant, and estimating probabilistically the numbers of people in the populations who may suffer each of various well-defined adverse health effects attributable to the pollutant. It is believed that information about

the health risks associated with various potential standards will aid the EPA administrator in selecting that standard which, in his judgment, protects the public health with an adequate margin of safety.

Because the risk estimates that EPA seeks are often based in part on dose-response relationships and lead exposure estimates which are uncertain, it is necessary to make probability judgments about relevant dose-response functions based on the available evidence, and to probabilistically estimate lead exposure under alternative NAAQS. Obtaining the health risk estimates then involves combining dose-response and exposure probability estimates.

The problem of estimating dose-response relationships is similar to that which exists in clinical medicine when there do not exist data that bear precisely on the patient's problem. In that case it is necessary to use scientific judgment to extrapolate from the data to make a best decision for the patient. Here, too, it is necessary to use scientific judgment to extrapolate from the available data. The extrapolation is not certain and, therefore, we will aid you to represent your opinion probabilistically. Furthermore, since the extrapolation depends on one's interpretation of the literature, different people will have different judgments. For each health effect we intend to obtain the probabilistic judgments of about five experts to sample the range of respected opinions. The model for estimating risks will not merge these judgments into a single average judgment, but rather will estimate the range of risks based on the range of judgments. If we as risk analysts do our job properly, then not only will we be able to show the EPA administrator the range of estimates based on the range of judgments, but we will also be able to show some of the sources of the disagreements. Indeed, a side benefit of this exercise in which we probe your knowledge in a structured manner may be to help identify sources of greatest disagreement.

Based on the evidence in the lead criteria document, two populations have been identified as being most susceptible to the effects of lead intoxication. One is children

from birth through the seventh birthday, and the other is pregnant women, or more precisely, the fetuses being carried by the pregnant women. A number of adverse health effects have been identified for which we would like to estimate dose-response functions including the one discussed below.

B.1.2 Lead-Induced Hemoglobin Decrements

We would like to represent in probabilistic form your opinion about the location and shape of dose-response functions for certain well defined levels of lead-induced hemoglobin reduction in the population of U.S. children from conception through the seventh birthday. We realize that available data do not fully define these functions; if they did, then it would be unnecessary to obtain your and other experts' opinions about them. Nevertheless, if the population, the exposure conditions, and the health effects are all precisely defined, then such functions in fact exist and we would like your best judgment about what they would look like if the data could be collected.

B.1.2.1 Definition of Hemoglobin Decrement as an Adverse Health Effect

There are differences of opinion, of course, as to what degree of reduction in hemoglobin level constitutes a health risk. The Clean Air Act makes it clear that EPA should set its standards to protect against adverse health effects. However, the reduction in hemoglobin level considered to be adverse may be different for regulatory purposes than for clinical action. Since it is generally agreed for children that a hemoglobin level of about 12 grams per deciliter is normal and about 9 grams per deciliter is anemic, we will specify two hemoglobin levels in that interval, namely 9.5 and 11 gm/dl, and treat each as the physiological effect for purposes of specifying a dose-response function. We will elicit from you probability judgments about the shape and location of the dose-response curve for each of the two levels of effects in the sensitive population.

B.1.3 Population at Risk

We are defining the most susceptible group as all United States children from conception through their seventh birthday. However, you may believe on various grounds that younger and older children in this group have different dose-response functions for lead-induced hemoglobin reduction. If so, then we will elicit your judgments separately for the age groups 0-3 and 4-6.

B.1.4 Exposure Conditions

We will be asking your judgment about population response rates at various blood-lead levels. Assume that the blood-lead level under consideration for a given judgment is in equilibrium as a result of a sufficiently long-term constant lead exposure without gross excursions above or below the stated level.

B.1.5 Physiological Conditions

Because the effect of lead on hemoglobin level depends on many parameters of a person's system, it is necessary to specify assumptions about those parameters in the population. The effect of lead in the system depends on the person's nutritional and metabolic status. The incidence of iron deficiency is greater in children than in adults, and greater yet in children age 0-3 than age 4-6, with a wide range of iron levels in children up to their seventh birthday. Assume that the distribution of iron levels is that which you believe it actually to be in the age groups 0-3 and 4-6. The effect of lead also depends on the levels of zinc, copper, vitamins A, C, D, and E, calcium, phosphorous and magnesium. Assume in all cases that these nutrients are distributed in the two age groups as you believe they in fact are, taking into account the wide range of diets and nutritional levels of U.S. children.

It is well established, of course, that erythrocyte protoporphyrin (EP) level is positively correlated with blood-lead level. However, the correlation is not perfect, and

there is a range of EP values at any given blood lead. You may believe that the effect of lead varies with EP level. If so, assume that EP is distributed at each blood-lead level as you believe it in fact is.

B.1.6 Factors to Consider

In order to help you bring to mind the relevant evidence so that you may consider it systematically, and also in order to help us to interpret your judgments, we would like to ask you to discuss briefly your interpretations of various aspects of the literature. First of all, could you tell us something about how, in your judgment, lead and iron interact to reduce hemoglobin levels and something about the relation between iron and lead levels in the body? Similarly, we would be interested in your opinions about the same questions with regard to zinc and lead and to calcium and lead. It appears that lead affects hemoglobin levels through three somewhat separate routes: interference with heme synthesis, interference with globin synthesis, and decreased life expectancy of erythrocytes. Briefly, what is your interpretation of the literature on these various effects of lead? EP level increases with iron deficiency, but it also increases with blood lead and is considered to be a proxy measure for the amount of lead recently cumulated in body tissue. Considering only that portion of EP elevation due to lead, in your judgment, is the relationship between blood lead and hemoglobin different at different EP levels?

Finally, we know that some groups of children are at greater risk of exposure to lead due to their living in deteriorating pre-1950 housing or in urban areas with large amounts of vehicular traffic. These children would tend to have high lead levels, but do you believe that there are also independent reasons to think that their dose-effect curves will be different from children who live in other circumstances? For example, these may be the same children who tend to have poorer diets and less access to medical care. What is your opinion regarding the possibility of a correlation between increased

exposure and increased susceptibility? Are there other factors to consider in thinking about the dose-response functions for lead-induced hemoglobin decrements that we should discuss now?

B.1.7 Factors to Keep in Mind When Making Probability Judgments

There is usually uncertainty associated with conclusions that we draw from research, and more generally in our everyday thinking. However, not everyone is aware of all the sources that should contribute to their uncertainty, nor are most people familiar with the process of actually expressing their uncertainty in probabilistic terms. When an expert is asked to make probability judgments on socially important matters, it is particularly important that he or she consider the relevant evidence in a systematic and effective manner and provide judgments that represent his or her opinions well.

Experimental psychologists and decision analysts have amassed a considerable amount of data concerning the way people form and express probabilistic judgments. The evidence suggests that when considering large amounts of complex information, most people employ simplifying heuristics and demonstrate certain systematic distortions of thought, i.e., cognitive biases, which adversely affect their judgments. The purpose of this section is to make you aware of these biases and heuristics so that, as much as possible, you can avoid them in making probability judgments. We will first review the most widespread biases and heuristics, and then offer some suggestions to help you mitigate their effects.

B.1.7.1 Sequential Consideration of Information

Generally, the order in which evidence is considered influences the final judgment, although logically that should not be the case. Of necessity, pieces of information are considered one by one in a sequential fashion. However, those considered first and last tend to dominate judgment. In part, initial information has its

undue influence because it provides the framework which subsequent information is then tailored to fit. For example, people usually search for evidence to *confirm* their initial hypotheses; they rarely look for evidence that *weighs against* them. The later evidence has its undue effect simply because it is fresher in memory.

Related to these sequential effects is the phenomenon of *anchoring and adjustment*. Based on early partial information, one forms an initial probability estimate regarding the event in question. This anchor judgment is then adjusted as subsequent information is considered. Unfortunately, such adjustments tend to be too conservative. In other words, too little weight is attached to information considered subsequent to the formation of the initial judgment.

B.1.7.2 Effects of Memory on Judgment

It is difficult for most people to conceptualize and make judgments about large, abstract universes or populations. A natural tendency is to recall specific members and then to consider them representative of the population as a whole. However, the specific instances often are recalled precisely because they stand out in some way, such as being familiar, unusual, especially concrete, or of personal significance. Unfortunately, the specific characteristics of these singular examples are then attributed, often incorrectly, to all the members of the population of interest. Moreover, these memory effects are often combined with the sequential phenomena discussed earlier. For example, in considering the evidence regarding the dose-response curve of a particular pollutant, you might naturally first think of a study you or a personal friend recently completed. Or you might think of a study you recently read, or one that was unusual and therefore stands out. The tendency might then be to treat the recalled studies as typical of the population of relevant research, ignoring important differences among studies. Subsequent attempts to recall information could result in thinking primarily of evidence consistent with the initial items you thought of.

B.1.7.3 Estimating Reliability of Information

People tend to overestimate the reliability of information, ignoring factors such as sampling error and imprecision of measurement. Rather they summarize evidence in terms of simple and definite conclusions, causing them to be overconfident in their judgments. This tendency is stronger when one has a considerable amount of intellectual and/or personal involvement in a particular field. In such cases, information is often interpreted in a way which is consistent with one's beliefs and expectations, results are overgeneralized, and contradictory evidence is ignored or underestimated.

B.1.7.4 Relation Between Event Importance and Probability

Sometimes the importance of events, or their possible costs or benefits, influence judgments about the certainty of the events when, rationally, importance should not affect probability. In other words, one's attitudes towards risk tend to affect one's ability to make accurate probability judgments. For example, many physicians tend to overestimate the probability of very severe diseases, because they feel it is important to detect and treat them; and similarly, many smokers underestimate the probability of adverse consequences of smoking, because they feel that the odds do not apply to themselves personally.

B.1.7.5 Estimation of Probabilities

Another limitation is related to one's ability to discriminate between levels of uncertainty, and to use appropriate criteria of discrimination for different ranges of probability. One result of this fact is that people tend to estimate both extreme and mid-range probabilities in the same fashion, usually doing a poor job in the extremes. It helps here to think in terms of odds as well as probabilities. Thus, for example, changing a probability estimate from 0.510 to 0.501 is equivalent to a change in odds from 1.041:1 to 1.004:1, but a change from an estimate of 0.999 to 0.990 changes the odds by a factor

of about 10? from 999:1 to 99:1 The closer to the extremes (either 0 or 1) that one is estimating probabilities, the greater the impact of small changes.

B.1.7.6 Recommendations

Although extensive and careful training would be necessary to eliminate all the problems mentioned above, some relatively simple suggestions can help minimize them. Most important is to be aware of one's natural cognitive biases, and to try consciously to avoid them.

To avoid *sequential effects* keep in mind that the order in which you think of information should not influence your final judgment. It may be helpful to actually note on paper the important facts you are considering and then to reconsider them in two or more sequences, checking the consistency of your judgments. Try to keep an open mind until you have gone through all the evidence, and don't let the early information you consider sway you more than is appropriate.

To avoid adverse *memory effects*, define various classes of information that you deem relevant, and then search your memory for examples of each. Don't restrict your thinking only to items that stand out for specific reasons. Make a special attempt to consider conflicting evidence, and to think of data that may be inconsistent with a particular theory. Also, be careful to concentrate on the given probability judgment and do not let your own values (how you would make the decision yourself) affect those judgments.

To accurately *estimate the reliability of information*, pay attention to such matters as sample size and power of the statistical tests. Keep in mind that data are probabilistic in nature, subject to elements of random error, imprecise measurements, and subjective evaluation and interpretation. In addition, the further one must extrapolate, or generalize, from a particular study to a situation of interest, the less reliable is the conclusion and the less certainty should be attributed to it. Rely more

heavily on information which you consider more reliable, but do not treat it as "absolute truth."

Keep in mind that *the importance of an event or an outcome should not influence its judged probability.* It is rational to let the costliness or severity of an outcome influence the point at which action is taken with respect to it, but not the judgment that is made about the outcome's likelihood.

Finally, in making probability judgments, think primarily in terms of the measure (probability or odds) with which you feel more comfortable, but sometimes translate to the alternative scale, or even to measures of other events (e.g,, the probability of the event not happening). When estimating very small or very large likelihoods, it is usually best to think in terms of odds, which are unbounded, instead of probabilities, which are bounded. For example, one can more easily conceptualize odds of 1:200 than a probability of .005.

B.1.8 Possible Adverse Health Effects of Elevated Erythrocyte Protoporphyrin (EP)

It seems well established that there is a positive dose-response relationship between blood lead and EP. Although this elevated EP can be traced to heme synthesis interference, there is disagreement as to what, if any adverse health effects are associated with lead induced elevated EP. We would like to discuss your views on this issue in our second visit after we have completed encoding your probabilistic judgments concerning the hemoglobin blood-lead dose-response function.

In your judgment, is lead-induced elevated EP associated with alterations in any of the following, and if so, at what level or range of levels of EP would you consider the effects to become adverse? These are alterations in:

1. Neurochemistry or other CNS functions,
2. Liver detoxification capabilities, or

3. Renal or endocrine function (in particular, reduced biosynthesis of 1,25-dihydroxyvitamin D).

Is there a relation between lead induced elevated EP and anemia of any sort beyond that which can be indexed by blood-lead level alone? Does elevated EP suggest that the individual is more susceptible to lead toxicity due to subsequent exposures? Are there other adverse health effects that you believe may be associated with elevated EP?

B.1.9 Final Preparation for Elicitation of Probability Judgments

The shape and location of the dose-response curve as we defined it above is uncertain, because the existing data do not determine the dose-response relationship exactly. Yet, if we have defined the dose-response curve precisely, in a mathematical sense such a relationship does exist. Our goal is to have you represent probabilistically your own uncertainty about the location and shape of this mathematically existent function, based on your expertise and the available knowledge. In responding to the questions we will ask you, please think carefully about the relevant information reviewed in the criteria document, and consult it, the literature, or your files as you deem appropriate.

The previous section suggests ways to think about the relevant data. The purpose of that section is to help minimize the biasing effects that frequently accompany the information overload naturally resulting from rapid consideration of large amounts of complex evidence. You may find it helpful to review the section or raise questions about the points made in it before we begin.

Uncertainty about a dose response relationship can be represented probabilistically in two different ways.

1. Uncertainty about the percentage of the defined population that would be affected by *given blood-lead concentrations* can be represented probabilistically.

2. Uncertainty about the blood-lead concentrations that would be required to affect a *given percentage of the defined population* can be represented probabilistically.

We will concentrate on one way at a time, focusing primarily on the first one.

For these purposes it is helpful to imagine that everyone in the population has a specified blood-lead level that has become stabilized in the manner described above under *Exposure Conditions*. Then, as a result, some percentage of the population will suffer the hemoglobin response.

Now, in order for us to determine your uncertainty about the percentages of the population that would be affected by *given blood lead levels* we must introduce a definition. Let C be the given blood-lead concentration in question.

Definition: R(C) is the percentage of the population for which a blood-lead level of C would cause the defined hemoglobin health effect.

Thus, R(C) is precisely the percent of the population that would show a response if the entire population had blood-lead levels of C under the conditions defined above. R(C) is usually called the population response rate.

The value of R(C) for a given C is uncertain, and we would like to obtain probability judgments from you about its possible values. We will elicit your judgments about the possible values of R(C) by specifying a particular percentage and having you consider how likely it is that R(C) is less than that value. To help you make your probability judgments, we will make use of a device called a probability wheel, which has adjustable sectors of blue and orange. We can read on the back of the wheel the percentage of the wheel that is each color. In making your judgments you are to imagine that the wheel is a perfectly fair random device, and that therefore the probability of its stopping with the pointer on blue is exactly represented by the relative area that is blue.

We will then proceed as follows. For each blood lead concentration C, we will specify a particular percent r, and also set the probability wheel to have a specific relative area of blue. Then you are to consider carefully the *question*:

Do you consider it more probable that the true population response rate R(C) is less than r *or* that the wheel would stop with the pointer on blue (on a random spin)?

You can give one of three responses:

1. You judge it to be more probable that R(C) is less than r;
2. You judge it to be more probable that the wheel would stop with the pointer in blue;
3. You cannot judge either event as more probable than the other.

For a particular concentration C and percent r, some wheel settings will have a small enough relative area of blue that you will feel confident making response (1). Other wheel settings will have a large enough relative area of blue that you will feel confident making response (2). The intermediate settings will be more difficult to judge. However, we will manipulate the wheel settings to find the one for which you feel most comfortable making response (3).

Once we have determined the point at which you are most comfortable with response (3), and still focusing on the given blood lead C, we will specify a new percent r', and repeat the procedure. This will continue for the given C until we have specified various percents. Then we will have elicited one of the probabilistic representations of your uncertainty about the dose-response relationship we need. We will obtain the next probabilistic representation by specifying a new C and continuing as before. We will do this for a number of values of C.

It frequently happens that an expert's judgments alter somewhat over the course of a session such as this, as he or she considers the evidence from various perspectives and thinks about the various responses called for. Hence, we will graph your responses and, at appropriate times, show them to you for your consideration and comparison. At these times you may wish to change some of the judgments you gave earlier.

We should emphasize that the judgments we are asking you to make are not simple ones, nor of course are there known correct answers. Rather, we want your best and most considered judgment in light of the available relevant scientific data, Therefore, please reflect on the available data carefully, feeling free to consult the lead criteria document or other sources as you wish as you formulate your judgments.

(Encode judgments for at least two concentrations of one dose-response function, then continue with instructions.)

Recall that uncertainty about the dose-response function can also be expressed in terms of the blood-lead concentration necessary to produce the effect in a given percentage of the population r, if the entire population had the same blood-lead concentration, stabilized as the result of exposure in the manner discussed earlier. More specifically, proceeding as before, let us introduce a definition:

Definition: C(R) is the blood-lead level that would cause the defined hemoglobin health effect in R percent of the population.

The value of C(R) is uncertain for a given R, and we want to obtain your judgments about its possible values. Analogously to what we have already done, for a given R we will specify concentrations c and ask you to consider how likely it is that the true C(R) is less than the specified c. More specifically, utilizing the wheel as before, we will ask you the *question*:

Do you consider it more likely that the true concentration C(R) is less than c, or that the wheel would stop with the pointer on blue (on a random spin)?

You can give one of three responses:

1. You judge it more probable that C(R) is less than c;
2. You judge it more probable that the wheel would stop with the pointer in blue;
3. You cannot judge either event as more probable than the other.

After determining the wheel setting at which you feel most comfortable with response (3), we will specify a new c and repeat the process. As before, for each R we will elicit your comparative judgments for various values of c.

B.2 DETAILED MATHEMATICAL FORMULATIONS

This section discusses in detail the mathematical techniques used to represent probability judgments about dose-response relationships. First, basic definitions and the notation are introduced. Then the normal-on-log-odds function is discussed as a distribution that can be fitted to probability judgments and that meets certain criteria. Methods are presented for obtaining least-squares estimates of the parameters of a normal-on-log-odds distribution and for assessing the goodness of fit. A family of equal-variance normal-on-log-odds distributions may sacrifice some goodness of fit to the judgments, but it meets all the criteria specified. This family is presented next, along with methods for obtaining least-squares estimates of its parameters and assessing its goodness of fit.

B.2.1 Definitions and Notation

This section presents definitions and notation. The reader is assumed to be familiar with the basic concepts of probability theory (e.g., definitions of random variables, probability density functions, and expectation and higher moments).

For uncertain quantity X, let

$f_X(x)$ = judgmental probability density function (PDF) for X and

$F_X(x_o)$ = cumulative judgmental density function (CDF) for X

$$= \int_{-\infty}^{x_o} f_X(x)dx.$$

Thus $F_X(x_o)$ is the judgmental probability* that X is less than or equal to x_o. Furthermore, let

$$\mu_x = E(X) = \text{expected value of } X$$

$$= \int_{-\infty}^{\infty} x_o f_X(x_o) dx_o$$

$$\sigma_X^2 = V(X) \text{ variance of } X$$

$$= \int_{-\infty}^{\infty} [x_o - E(X)]^2 f_X(x_o) dx_o$$

The uncertain quantities of interest in this report are population response rates, denoted R, at each of several blood-lead levels, denoted L. During encoding sessions with recognized experts, judgments representing a cumulative density function (CDF) for a response rate R given a population blood-lead level L [i.e., judgments representing $F_{R/L}(R)$] will be elicited for a number of blood-lead levels. Taken together, these judgments will provide a family of probabilistic relationships that represents the judgmental probability that the population response rate (a fraction between 0 and 1) will be less than or equal to a particular value R, given a blood-lead level of L, for a specific adverse health effect, population, and exposure conditions.

As already discussed, the calculation of risk generally requires an interpolation between specifically assessed points. Such interpolation is best accomplished by fitting a function to the judged probabilities. Section B.2.2 describes a method for doing this.

*Hereafter, all probabilities referred to are judgmental probabilities, even if not explicitly stated.

B.2.2 Normal-on-Log-Odds Function

Because probabilistic judgments about dose-response relationships are generally "s-shaped" over the closed [0,1] interval, they can be difficult to represent with closed-form mathematical functions. One particularly useful function that is relatively easy to work with is the normal-on-log-odds distribution. The normal-on-log-odds distribution is obtained by fitting a normal distribution to the natural log of the odds implied by the population response rates R. Thus, the following relationships are defined:

$$X = \frac{R}{1 - R}, \quad 0 \leq X \leq \infty$$

$$Y = \ln(X), \quad -\infty \leq Y \leq \infty$$

where X is the odds variable and Y is the log of the odds variable. The variable Y is assumed to be normally distributed with mean μ and variance σ^2. The degree to which the assumption is appropriate can be tested with each set of judgments. If Y is normally distributed, X is log-normally distributed. Although no closed-form expression is readily available for the distribution on R, all probabilistic and statistical results of interest on R can be obtained through the distribution on Y since

$$\text{pr}(R \leq r) = \text{pr}\ X \leq \frac{r}{1 - r} = \text{pr}\ Y \leq \ln \frac{r}{1 - r}$$

and

$$\text{pr}(Y \leq y_o) = \text{pr}\ Z \leq \frac{y_o - \mu}{\sigma} = \Phi \frac{y_o - \mu}{\sigma}$$

where pr(·) denotes probability and Z is the unit normal random deviate for which $\Phi(Z)$ is extensively tabulated and available in computer libraries. Thus

$$F_R(r) = F_X\left(\frac{r}{1-r}\right)$$

$$= F_Y \ln\left(\frac{r}{1-r}\right)$$

$$= F_Z \frac{\ln\left(\frac{r}{1-r}\right) - \mu}{\sigma}$$

$$= \Phi \frac{\ln\left(\frac{r}{1-r}\right) - \mu}{\sigma} \qquad \text{(B.1)}$$

Table B.1 summarizes quantities of interest for these variables.

The parameters μ and σ can be estimated in a least-squares sense from an assessed distribution. Let the n assessed points for a CDF be denoted (R_i, F_i) for i = 1, ..., n, where R_i is the population response rate and F_i is the associated cumulative judgmental probability.

Least squares estimates $\hat{\mu}$ and $\hat{\sigma}$ for μ and σ can be obtained by linearly regressing the z_i on the y_i. The reciprocal of the slope of the regression equation is $\hat{\sigma}$ and $\hat{\mu}$ is the y-intercept, i.e., the value of y corresponding to z = 0 in the regression equation. Since m judgmental distributions are assessed, one for each blood-lead level j for j = 1, ..., m, there are m means $(\hat{\mu}_1, \ldots, \hat{\mu}_m)$ and standard deviations $(\hat{\sigma}_1, \ldots, \hat{\sigma}_m)$, one pair for each of the judged distributions.

The goodness of fit of each distribution fit to the judgments for each blood-lead level is given by the standard regression r^2 statistic obtained by regressing the $\hat{F}_i$ on the F_i, where

$$r^2 = 1 - \frac{\sum(z_i - \hat{z}_i)^2}{\sum(z_i - \bar{z})^2}$$

TABLE B.1 *Summary of Variables Pertaining to the Normal-On-Log-Odds Function*

Quantity of Interest	Variable		
	Y	X	R
Defining Equations	$Y = \ln(X)$	$X = e^Y = R/(1 - R)$	$R = \frac{X}{1 + X}$
Distribution[a]	$N(\mu,\sigma^2)$	$\Lambda(\mu,\sigma^2)$	-
Mean[b]	μ	$\mu_x = e^{\mu+\sigma^2/2}$	$\frac{\mu_x}{1 + \mu_x} - \frac{\sigma_x^2}{(1 + \mu_x)^3}$
Variance[b]	σ^2	$\sigma_x^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$	$\sigma_x^2 \left(\frac{1}{1 + \mu_x}\right)^4$
Median	μ	e^{μ}	$\frac{e^{\mu}}{1 + e^{\mu}}$
Mode	μ	$e^{\mu-\sigma^2}$	$\frac{e^{\mu-\sigma^2}}{1 + e^{\mu-\sigma^2}}$

[a] $N(\mu,\sigma^2)$ denotes the normal PDF; $\Lambda(\mu,\sigma^2)$ denotes the lognormal PDF; the PDF for R cannot be expressed in closed form.

[b] Mean and variance expressions for R are approximate; they were derived with the method of moments; other estimates can be obtained with numerical methods on $F_Y = N(\mu,\sigma^2)$.

and $\hat{F}_i$ denotes the estimated F values {i.e., $\hat{F}_i = \Phi(y_i - \hat{\mu})/\hat{\sigma}$, from Eq. B.1}, and obtaining the r^2 value of this regression. If the r^2 value is sufficiently high, the normal-on-log-odds distribution with parameters $\hat{\mu}_j$ and $\hat{\sigma}_j$ describes the judgments well for lead level j. This distribution can then be used for interpolation.

To summarize, formulas were presented to calculate the parameters of a normal-on-log-odds distribution, which can be used to represent a judged probability distribution over population response rate for a given blood-lead level. Goodness of fit is given by a regression r^2 statistic. The normal-on-log-odds distribution variable and its PDF can be used to estimate judgmental probabilities for the response rate variable R, which is then used to estimate risk.

B.2.3 Representing a Probabilistic Dose-Response Surface with a Family of Equal-Variance Normal-on-Log-Odds Distributions

If the $\hat{\sigma}_j$ are approximately equal, it is desirable to set them equal to a single pooled value before proceeding further (i.e., set $\hat{\sigma}_j = \hat{\sigma}'_p = \sigma_y = \sigma$ for all j). This step assures that the fitted distributions for two different lead levels never cross one another. (Crossed distributions would imply that exceeding a specified response rate is more probable at blood-lead level L_1 than at L_2, where L_2 is greater than L_1.) Because of the manner in which judgments were obtained from the pilot subjects, crossings did not occur. Furthermore, judgments will be obtained from the health experts over a wider probability range so that crossings would only be possible at cumulative probabilities very near to 0 and 1. However, even crossings like these represent an inconsistency the experts ought to avoid and therefore should not be allowed.

A pooled estimate of σ, denoted $\hat{\sigma}'_p$, can be obtained by first subtracting the mean μ_j for each set of assessed points,

$$y'_{j,i} = y_{j,i} - \hat{\mu}_j \text{ for } i = 1, \ldots, n_j \text{ and } j = 1, \ldots, m$$

where n_j is the number of assessed points for F_{R/L_j} and $y_{j,i}$ is an assessed point on F_{R/L_j}. The $y'_{j,i}$ can then be used to calculate a mean (which should be very close to zero) and a variance (which will be the pooled, least-squares estimate $\hat{\sigma}'_p$) by regressing the $Z_{j,i}$ on the $y'_{j,i}$. The individual, least-squares estimates of the means can be recalculated at each lead level by finding the best-fitting line (in the least squares sense) with a slope of $1/\hat{\sigma}'_p$. Under these conditions, the result is simply

$$\hat{\mu}'_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{j,i}.$$

Denote the results of these last steps as $\hat{\mu}'_1, \ldots, \hat{\mu}'_m$. Goodness of fit again can be obtained by regressing the recalculated estimates (which may be denoted as the $\hat{F}'_{j,i}$) on the original $F_{j,i}$.

If a functional relationship can be established between blood-lead levels and the $\hat{\mu}'_j$, a judgmental probability density function over R can be calculated for any blood-lead level, should that be necessary for interpolating between blood-lead levels. This continuum of density functions in essence is a probability surface over the dose-response plane.

Any blood-lead distribution can be combined with the dose-response curve corresponding to judgmental cumulative probability F, $D_F(L)$, to yield an estimate of the overall fraction of the population R suffering the adverse health effect with cumulative probability F. Such calculations result in the risk estimates summarized in Sec. 5 and described in detail in App. D.

B.3 RESULTS

Results are presented in three tables: Table B.2 covers encoded judgments, Table B.3 summarizes the functions fit to the judgments, and Table B.4 compares the

TABLE B.2 Encoded Probability Judgments of Experts Regarding Lead-Induced Hemoglobin Decrements

Expert A, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years

Blood-Lead Levels

45 µg/dL		55 µg/dL		65 µg/dL		75 µg/dL	
R[a]	F[b]	R	F	R	F	R	F
0.01	0.1	0.01	0.05	0.04	0.05	0.05	0.05
0.02	0.25	0.03	0.18	0.07	0.15	0.1	0.2
0.04	0.4	0.06	0.4	0.1	0.35	0.15	0.4
0.05	0.5	0.09	0.5	0.13	0.45	0.2	0.5
0.07	0.8	0.12	0.6	0.16	0.6	0.25	0.6
0.09	0.95	0.15	0.95	0.19	0.9	0.3	0.7
0.1	0.99	0.16	0.99	0.22	0.98	0.35	0.95
	0.19	0.999			0.4	0.99	

Expert A, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years

25 µg/dL		35 µg/dL		45 µg/dL	
R	F	R	F	R	F
0.01	0.05	0.03	0.02	0.04	0.05
0.02	0.4	0.05	0.2	0.05	0.1
0.03	0.5	0.09	0.45	0.1	0.4
0.05	0.9	0.12	0.55	0.15	0.5
0.06	0.95	0.15	0.6	0.2	0.8
0.08	0.99	0.18	0.98	0.25	0.98

[a]R denotes population response rate (fraction having hemoglobin levels ≤ 9.5 or 11 g/dL).

[b]F denotes cumulative probability.

TABLE B.2 (Cont'd)

Expert C, Hemoglobin ≤ 9.5 g/dL, Ages 0–6 Years

Blood-Lead Levels

5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
R	F	R	F	R	F	R	F	R	F	R	F
0.005	0.01	0.015	0.001	0.025	0.01	0.035	0.01	0.045	0.01	0.05	0.01
0.01	0.05	0.02	0.02	0.03	0.02	0.04	0.04	0.05	0.05	0.06	0.05
0.0175	0.5	0.03	0.5	0.035	0.25	0.05	0.5	0.06	0.5	0.07	0.5
0.03	0.8	0.04	0.75	0.04	0.5	0.06	0.75	0.07	0.8	0.08	0.8
0.035	0.95	0.05	0.99	0.05	0.75	0.07	0.97	0.08	0.97	0.09	0.95
0.045	0.99	0.06	0.999	0.06	0.97	0.075	0.99	0.085	0.99	0.1	0.99
				0.0675	0.99						

Expert C, Hemoglobin ≤ 11 g/dL, Ages 0–6 Years

0.03	0.01	0.05	0.03	0.05	0.001	0.1	0.01	0.15	0.01	0.15	0.001
0.06	0.28	0.08	0.2	0.09	0.11	0.14	0.2	0.2	0.5	0.2	0.08
0.08	0.5	0.09	0.5	0.12	0.5	0.16	0.5	0.25	0.7	0.25	0.3
0.09	0.5	0.11	0.6	0.13	0.7	0.18	0.7	0.3	0.98	0.27	0.5
0.12	0.96	0.14	0.91	0.17	0.98	0.22	0.98	0.35	0.999	0.3	0.8
0.15	0.99	0.17	0.99	0.21	0.99	0.25	0.99			0.35	0.98
		0.2	0.999							0.4	0.999

Expert D, Hemoglobin ≤ 9.5 g/dL, Ages 0.3 Years

		0.01	0.5	0.005	0.01	0.03	0.001	0.07	0.001	0.1	0.001
		0.03	0.999	0.03	0.5	0.06	0.25	0.11	0.25	0.16	0.25
				0.07	0.999	0.08	0.5	0.16	0.5	0.2	0.5
						0.11	0.75	0.19	0.75	0.25	0.75
						0.15	0.999	0.25	0.999	0.33	0.999

TABLE B.2 (Cont'd)

Expert D, Hemoglobin ≤ 9.5 g/dL, Ages 4-6 Years

Blood-Lead Levels

5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
R	F	R	F	R	F	R	F	R	F	R	F
				0.00005	0.001	0.02	0.001	0.04	0.001	0.06	0.001
				0.005	0.5	0.05	0.5	0.06	0.25	0.08	0.25
				0.015	0.999	0.06	0.75	0.08	0.5	0.1	0.5
						0.08	0.999	0.1	0.75	0.11	0.75
								0.12	0.999	0.15	0.999

Expert D, Hemoglobin ≤ 11 g/dL, Ages 0.3 Years

0.01	0.01	0.02	0.01	0.11	0.01	0.15	0.01	0.17	0.05	0.23	0.125
0.02	0.25	0.04	0.25	0.115	0.25	0.16	0.25	0.19	0.25	0.26	0.25
0.04	0.5	0.06	0.5	0.15	0.5	0.18	0.5	0.22	0.5	0.3	0.5
0.1	0.75	0.12	0.75	0.178	0.75	0.22	0.75	0.26	0.75	0.4	0.75
0.118	0.99	0.15	0.99	0.21	0.99	0.24	0.99	0.29	0.95	0.47	0.875

Expert D, Hemoglobin ≤ g/dL, Ages 4-6 Years

0.01	0.01	0.01	0.01	0.05	0.01	0.07	0.01	0.08	0.01	0.1	0.01
0.01	0.25	0.02	0.25	0.07	0.5	0.09	0.5	0.1	0.5	0.12	0.25
0.02	0.5	0.03	0.5	0.085	0.75	0.105	0.75	0.16	0.99	0.14	0.5
0.06	0.99	0.09	0.99	0.12	0.99	0.14	0.99			0.155	0.75
										0.17	0.99

Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 0.3 Years

0.001	0.01	0.002	0.01	0.003	0.01	0.025	0.01	0.03	0.01	0.05	0.01
0.01	0.1	0.03	0.05	0.035	0.05	0.05	0.08	0.06	0.06	0.08	0.1
0.03	0.5	0.07	0.5	0.04	0.1	0.1	0.27	0.1	0.2	0.22	0.5
0.1	0.9	0.11	0.61	0.07	0.25	0.15	0.56	0.19	0.5	0.36	0.75
0.15	0.99	0.15	0.83	0.09	0.5	0.2	0.81	0.31	0.75	0.4	0.9
		0.19	0.9	0.15	0.75	0.26	0.98	0.41	0.99	0.47	0.99
		0.23	0.99	0.21	0.9						
				0.25	0.99						

TABLE B.3 Summary of Functions Fit to the Probabilistic Judgments of Experts Regarding Lead-Induced Hemoglobin Decrements

Blood-Lead Level (µg/dL	Parameters of Normal-on-Ln (R/(1-R)) Distribution		Moments of NOLO Distribution Over R		r^2 for Regression of F on F
	Mean	SD[a]	E[R][b]	SD[R][c]	
Expert A, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years					
45	-3.3546	0.6820	0.0412	0.0311	0.8914
55	-2.9338	0.6820	0.0607	0.0454	0.8388
65	-2.0724	0.6820	0.1275	0.0910	0.9305
75	-1.5754	0.6820	0.1869	0.1263	0.9407
Expert A, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years					
55	-3.6826	0.5945	0.0288	0.0184	0.9734
65	-2.2713	0.5945	0.1051	0.0635	0.8968
75	-2.0617	0.5945	0.1254	0.0745	0.9464

[a]SD denotes standard deviation.

[b]E[R] denotes the mean of the distribution for R.

[c]SD[R] denotes the standard deviation of the distribution for R.

TABLE B.2 (Cont'd)

Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 4-6 Years

Blood-Lead Levels

5 µg/dL		15 µg/dL		25 µg/dL		35 µg/dL		45 µg/dL		55 µg/dL	
R	F	R	F	R	F	R	F	R	F	R	F
0.004	0.15	0.01	0.15	0.02	0.17	0.015	0.1	0.02	0.03	0.06	0.05
0.01	0.5	0.02	0.5	0.04	0.5	0.03	0.22	0.065	0.1	0.11	0.4
0.06	0.78	0.07	0.78	0.1	0.9	0.06	0.54	0.105	0.42	0.15	0.5
0.09	0.98	0.1	0.98	0.13	0.99	0.1	0.85	0.165	0.6	0.22	0.75
						0.13	0.95	0.215	0.8	0.26	0.9
								0.265	0.99	0.3	0.98

Expert E, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years

0.02	0.01	0.03	0.01	0.035	0.01	0.04	0.01	0.06	0.01	0.1	0.01
0.04	0.15	0.06	0.15	0.06	0.1	0.08	0.08	0.1	0.07	0.2	0.35
0.08	0.5	0.1	0.5	0.095	0.25	0.15	0.27	0.16	0.25	0.3	0.5
0.12	0.61	0.14	0.61	0.15	0.5	0.2	0.56	0.22	0.5	0.4	0.62
0.16	0.83	0.18	0.83	0.195	0.75	0.25	0.81	0.35	0.75	0.5	0.94
0.2	0.9	0.22	0.9	0.25	0.9	0.3	0.98	0.45	0.99	0.55	0.99
0.24	0.98	0.26	0.98	0.28	0.99						

Expert E, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years

0.02	0.15	0.01	0.01	0.05	0.06	0.05	0.07	0.05	0.03	0.05	0.01
0.05	0.47	0.04	0.15	0.08	0.16	0.08	0.1	0.1	0.1	0.1	0.05
0.08	0.55	0.07	0.47	0.11	0.5	0.11	0.27	0.15	0.43	0.15	0.4
0.11	0.78	0.1	0.55	0.17	0.9	0.14	0.54	0.2	0.6	0.2	0.5
0.14	0.98	0.13	0.78	0.2	0.99	0.17	0.85	0.25	0.8	0.25	0.75
		0.16	0.98			0.2	0.95	0.3	0.99	0.3	0.9
										0.35	0.98

TABLE B.3 (Cont'd)

Blood-Lead Level (µg/dL	Parameters of Normal-on-Ln (R/(1-R)) Distribution Mean	SD[a]	Moments of NOLO Distribution Over R E[R][b]	SD[R][c]	r^2 for Regression of F on F
Expert D, Hemoglobin ≤ 9.5 g/dL, Ages 0-3 Years					
15	-4.5852	0.3557	0.0107	0.0039	1.0000
25	-3.8759	0.3557	0.0215	0.0078	0.8421
35	-2.4991	0.3557	0.0797	0.0272	0.9805
45	-1.7769	0.3557	0.1500	0.0475	0.9654
55	-1.4097	0.3557	0.2019	0.0602	0.9935
Expert D, Hemoglobin ≤ 9.5 g/dL, Ages 4-6 Years					
25	-6.4604	1.0409	0.0027	0.0037	0.8374
35	-3.0443	0.2177	0.0464	0.0098	0.9536
45	-2.5123	0.2177	0.0764	0.0156	0.9437
55	-2.2433	0.2177	0.0976	0.0195	0.9820
Expert D, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years					
5	-3.1747	0.4919	0.0445	0.0225	0.9285
15	-2.7097	0.4919	0.0686	0.0340	0.9565
25	-1.7442	0.4919	0.1586	0.0720	0.9693
35	-1.4655	0.4919	0.1975	0.0858	0.9674
45	-1.2485	0.4919	0.2322	0.0967	0.9966
55	-0.7254	0.4919	0.3312	0.1196	0.9729

TABLE B.3 (Cont'd)

Blood-Lead Level (µg/dL	Parameters of Normal-on-Ln (R/(1-R)) Distribution Mean	SD[a]	Moments of NOLO Distribution Over R E[R][b]	SD[R][c]	r^2 for Regression of $F_{\text{F on F}}$
Expert C, Hemoglobin ≤ 9.5 g/dL, Ages 0-6 Years					
5	-3.9995	0.2764	0.0187	0.0052	0.9806
15	-3.4481	0.2764	0.0319	0.0087	0.9899
25	-3.1298	0.2764	0.0433	0.0117	0.9906
35	-2.9187	0.2764	0.0529	0.0142	0.9917
45	-2.7424	0.2764	0.0624	0.0166	0.9900
55	-2.5781	0.2764	0.0727	0.0191	0.9713
Expert C, Hemoglobin ≤ 11 g/dL, Ages 0-6 Years					
5	-3.6826	0.5945	0.0288	0.0184	0.9734
15	-2.2713	0.5945	0.1051	0.0635	0.8968
25	-2.0617	0.5945	0.1254	0.0745	0.9464
35	-1.6965	0.2720	0.1582	0.0372	0.9857
45	-1.3190	0.2720	0.2144	0.0471	0.9683
55	-1.0498	0.2720	0.2624	0.0542	0.9816

TABLE B.3 (Cont'd)

Blood-Lead Level (µg/dL	Parameters of Normal-on-Ln (R/(1-R)) Distribution Mean	SD[a]	Moments of NOLO Distribution Over R E[R][b]	SD[R][c]	r^2 for Regression of F on F
Expert D, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years					
5	-3.8759	0.4894	0.0227	0.0116	0.9448
15	-3.4866	0.4894	0.0331	0.0168	0.9968
25	-2.5049	0.1777	0.0764	0.0127	0.9579
35	-2.2446	0.1777	0.0969	0.0157	0.9674
45	-2.0993	0.1777	0.1104	0.0176	0.9358
55	-1.8573	0.1777	0.1364	0.0212	0.9749
Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 0-3 Years					
5	-3.7820	0.9600	0.0331	0.0414	0.9817
15	-2.7992	0.9600	0.0773	0.0987	0.9553
25	-2.4725	0.9600	0.0994	0.1280	0.9586
35	-1.9602	0.9600	0.1413	0.1836	0.9646
45	-1.5643	0.9600	0.1786	0.2301	0.9854
55	-1.4002	0.9600	0.1951	0.2486	0.9866
Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 4-6 Years					
5	-4.1623	0.8224	0.0209	0.0206	0.9802
15	-3.6856	0.8224	0.0329	0.0323	0.9884
25	-3.3376	0.8224	0.0454	0.0444	0.9804
35	-3.0218	0.8224	0.0602	0.0588	0.9503
45	-2.1141	0.8224	0.1275	0.1217	0.9848
55	-1.9121	0.8224	0.1481	0.1398	0.9327

TABLE B.3 (Cont'd)

Blood-Lead Level (µg/dL	Parameters of Normal-on-Ln (R/(1-R)) Distribution Mean	SD[a]	Moments of NOLO Distribution Over R E[R][b]	SD[R][c]	r^2 for Regression of F on F
Expert E, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years					
5	-2.3463	0.5991	0.0987	0.0606	0.9802
15	-2.1130	0.5991	0.1203	0.0725	0.9884
25	-1.9311	0.5991	0.1398	0.0828	0.9804
35	-1.6552	0.5991	0.1739	0.0995	0.9503
45	-1.3014	0.5991	0.2260	0.1218	0.9848
55	-0.9199	0.5991	0.2924	0.1437	0.9327
Expert E, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years					
5	-2.8545	0.5912	0.0626	0.0388	0.9033
15	-2.6381	0.5912	0.0761	0.0468	0.9457
25	-2.2151	0.5912	0.1101	0.0658	0.9679
35	-1.9784	0.5912	0.1342	0.0785	0.9395
45	-1.7096	0.5912	0.1665	0.0943	0.9703
55	-1.5286	0.5912	0.1912	0.1055	0.9821

TABLE B.4 Summary and Comparison of Judgments and Fitted Functions Concerning Lead-Induced Hemoglobin Decrements

Index	45 µg/dL	55 µg/dL	65 µg/dL	75 µg/dL
Expert A, Hemoglobin ≤ 11 g/dL, Ages 0–3 Years				
	Blood-Lead Level			
	Encoded Judgments			
Median	0.05	0.09	0.14	0.20
50% C.I.[a]	(0.02, 0.07)	(0.04, 0.13)	(0.09, 0.18)	(0.11, 0.31)
90% C.I.	(0.01, 0.09)	(0.01, 0.15)	(0.04, 0.21)	(0.05, 0.35)
	Fitted Functions			
Median	0.03	0.05	0.11	0.17
50% C.I.	(0.02, 0.05)	(0.03, 0.08)	(0.07, 0.17)	(0.12, 0.25)
90% C.I.	(0.01, 0.10)	(0.02, 0.14)	(0.04, 0.28)	(0.06, 0.39)
98% C.I.	(0.01, 0.15)	(0.01, 0.21)	(0.03, 0.38)	(0.04, 0.50)
Expert A, Hemoglobin ≤ 11 g/dL, Ages 4–6 Years				
	Encoded Judgments			
Median	0.03	0.11	0.15	
50% C.I.	(0.02, 0.04)	(0.06, 0.16)	(0.08, 0.19)	
90% C.I.	(0.01, 0.06)	(.03, 0.18)	(0.04, 0.24)	
	Fitted Functions			
Median	0.02	0.09	0.11	
50% C.I.	(0.02, 0.04)	(0.06, 0.13)	(0.08, 0.16)	
90% C.I.	(0.01, 0.06)	(0.04, 0.22)	(0.05, 0.25)	
98% C.I.	(0.01, 0.09)	(0.03, 0.29)	(0.03, 0.34)	

[a]C.I. denotes Credible Interval.

TABLE B.4 (Cont'd)

Index	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL)
Expert C, Hemoglobin ≤ 9.5 g/dL, Ages 0-~~3~~ 10 Years						
			Blood-Lead Level			
			Encoded Judgments			
Median	0.02	0.03	0.04	0.05	0.06	0.07
50% C.I.	(0.01, 0.03)	(0.02, 0.04)	(0.04, 0.05)	(0.04, 0.06)	(0.05, 0.07)	(0.06, 0.08)
90% C.I.	(0.01, 0.04)	(0.02, 0.05)	(0.03, 0.06)	(0.04, 0.07)	(0.05, 0.08)	(0.06, 0.09)
			Fitted Functions			
Median	0.02	0.03	0.04	0.05	0.06	0.07
50% C.I.	(0.01, 0.02)	(0.03, 0.04)	(0.04, 0.05)	(0.04, 0.06)	(0.05, 0.07)	(0.06, 0.08)
90% C.I.	(0.01, 0.03)	(0.02, 0.05)	(0.03, 0.06)	(0.03, 0.08)	(0.04, 0.09)	(0.05, 0.11)
98% C.I.	(0.01, 0.04)	(0.02, 0.06)	(0.02, 0.08)	(0.03, 0.09)	(0.03, 0.11)	(0.04, 0.13)
Expert C, Hemoglobin ≤ 11 g/dL, Ages 0-6 Years						
			Encoded Judgments			
Median	0.08	0.09	0.12	0.16	0.20	0.27
50% C.I.	(0.06, 0.10)	(0.08, 0.12)	(0.10, 0.14)	(0.14, 0.19)	(0.17, 0.26)	(0.24, 0.29)
90% C.I.	(0.03, 0.10)	(0.05, 0.13)	(0.07, 0.15)	(0.11, 0.20)	(0.15, 0.27)	(0.18, 0.31)
			Fitted Functions			
Median	0.07	0.10	0.12	0.15	0.21	0.26
50% C.I.	(0.06, 0.09)	(0.08, 0.11)	(0.10, 0.14)	(0.13, 0.18)	(0.18, 0.24)	(0.23, 0.30)
90% C.I.	(0.05, 0.11)	(0.06, 0.14)	(0.08, 0.17)	(0.10, 0.22)	(0.15, 0.29)	(0.18, 0.35)
98% C.I.	(0.04, 0.13)	(0.05, 0.17)	(0.06, 0.20)	(0.09, 0.26)	(0.12, 0.33)	(0.16, 0.40)

TABLE B.4 (Cont'd)

Expert D, Hemoglobin ≤ 9.5 g/dL, Ages 0-3 Years

Index	Blood-Lead Level 15 ~~5~~ µg/dL	25 ~~15~~ µg/dL	35 ~~25~~ µg/dL	45 ~~35~~ µg/dL	55 ~~45~~ µg/dL	~~55 µg/dL)~~
Encoded Judgments						
Median	0.01	0.03	0.08	0.16	0.20	
50% C.I.	(--, 0.02)	(0.02, 0.05)	(0.06, 0.11)	(0.11, 0.19)	(0.16, 0.25)	
90% C.I.	(--, 0.03)	(0.01, 0.07)	(0.04, 0.14)	(0.08, 0.24)	(0.11, 0.31)	
Fitted Functions						
Median	0.01	0.02	0.08	0.14	0.20	
50% C.I.	(0.01, 0.01)	(0.02, 0.03)	(0.06, 0.09)	(0.12, 0.18)	(0.16, 0.24)	
90% C.I.	(0.01, 0.02)	(0.01, 0.04)	(0.04, 0.13)	(0.09, 0.23)	(0.12, 0.30)	
98% C.I.	(0.00, 0.02)	(0.01, 0.04)	(0.03, 0.16)	(0.07, 0.28)	(0.12, 0.36)	

Expert D, Hemoglobin ≤ 9.5, Ages 4-6 Years

Encoded Judgments				
Median	0.005	0.05	0.08	0.10
50% C.I.	(0.00, 0.01)	(0.03, 0.06)	(0.06, 0.10)	(0.08, 0.11)
90% C.I.	(0.00, 0.01)	(0.02, 0.08)	(0.04, 0.12)	(0.06, 0.14)
Fitted Functions				
Median	0.002	0.05	0.07	0.10
50% C.I.	(0.00, 0.00)	(0.04, 0.05)	(0.07, 0.09)	(0.08, 0.11)
90% C.I.	(0.00, 0.01)	(0.03, 0.06)	(0.05, 0.10)	(0.07, 0.13)
98% C.I.	(0.00, 0.02)	(0.03, 0.07)	(0.05 0.12)	(0.06, 0.15)

TABLE B.4 (Cont'd)

Index	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45	55
Expert D, Hemoglobin ≤ 11 g/dL, Ages 0-3 Years						
			Blood-Lead Level			
			Encoded Judgments			
Median	0.04	0.06	0.15	0.18	0.22	0.30
50% C.I.	(0.02, 0.10)	(0.04, 0.12)	(0.11, 0.18)	(0.16, 0.22)	(0.19, 0.26)	(0.26, 0.40)
90% C.I.	(0.01, 0.03)	(0.02, 0.15)	(0.11, 0.21)	(0.14, 0.26)	(0.17, 0.29)	(0.21, 0.52)
			Fitted Functions			
Median	0.04	0.06	0.15	0.19	0.22	0.33
50% C.I.	(0.03, 0.06)	(0.05, 0.08)	(0.11, 0.20)	(0.14, 0.24)	(0.17, 0.29)	(0.26, 0.40)
90% C.I.	(0.02, 0.09)	(0.03, 0.13)	(0.07, 0.28)	(0.09, 0.34)	(0.11, 0.39)	(0.18, 0.52)
98% C.I.	(0.01, 0.12)	(0.02, 0.17)	(0.05, 0.35)	(0.07, 0.42)	(0.08, 0.47)	(0.13, 0.60)
Expert D, Hemoglobin ≤ 11 g/dL, Ages 4-6 Years						
			Encoded Judgments			
Median	0.02	0.03	0.07	0.09	0.10	0.14
50% C.I.	(0.01, 0.04)	(0.02, 0.06)	(0.06, 0.09)	(0.08, 0.11)	(0.09, 0.13)	(0.12, 0.16)
90% C.I.	(0.01, 0.06)	(0.01, 0.09)	(0.05, 0.11)	(0.07, 0.13)	(0.08, 0.16)	(0.10, 0.17)
			Fitted Functions			
Median	0.02	0.03	0.08	0.10	0.11	0.14
50% C.I.	(0.01, 0.03)	(0.02, 0.04)	(0.07, 0.08)	(0.09, 0.11)	(0.10, 0.12)	(0.12, 0.15)
90% C.I.	(0.01, 0.04)	(0.01, 0.06)	(0.06, 0.10)	(0.07, 0.12)	(0.08, 0.14)	(0.10, 0.17)
98% C.I.	(0.01, 0.06)	(0.01, 0.09)	(0.05, 0.11)	(0.07, 0.14)	(0.07, 0.16)	(0.09, 0.19)

TABLE B.4 (Cont'd)

Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 0-3 Years

Blood-Lead Level

Index	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL)
Median	0.03	0.07	0.09	0.14	0.19	0.22
50% C.I.	(0.02, 0.07)	(0.05, 0.13)	(0.06, 0.15)	(0.09, 0.19)	(0.11, 0.31)	(0.13, 0.36)
90% C.I.	(0.01, 0.12)	(0.03, 0.21)	(0.03, 0.23)	(0.04, 0.24)	(0.05, 0.39)	(0.06, 0.44)

Fitted Functions

Median	0.02	0.06	0.08	0.12	0.17	0.20
50% C.I.	(0.01, 0.04)	(0.03, 0.10)	(0.04, 0.14)	(0.07, 0.21)	(0.10, 0.29)	(0.11, 0.32)
90% C.I.	(0.00, 0.10)	(0.01, 0.23)	(0.02, 0.29)	(0.03, 0.41)	(0.04, 0.50)	(0.05, 0.54)
98% C.I.	(0.00, 0.18)	(0.01, 0.36)	(0.01, 0.44)	(0.01, 0.57)	(0.02, 0.66)	(0.03, 0.70)

Expert E, Hemoglobin ≤ 9.5 g/dL, Ages 4-6 Years

Encoded Judgments

Median	0.01	0.02	0.04	0.06	0.13	0.15
50% C.I.	(0.01, 0.05)	(0.01, 0.06)	(0.01, 0.07)	(0.03, 0.09)	(0.08, 0.20)	(0.09, 0.22)
90% C.I.	(0.00, 0.09)	(0.00, 0.10)	(0.00, 0.11)	(0.01, 0.13)	(0.03, 0.25)	(0.06, 0.29)

Fitted Functions

Median	0.02	0.02	0.03	0.05	0.11	0.13
50% C.I.	(0.01, 0.06)	(0.01, 0.09)	(0.02, 0.12)	(0.03, 0.16)	(0.06, 0.32)	(0.08, 0.36)
90% C.I.	(0.00, 0.06)	(0.02, 0.09)	(0.01, 0.12)	(0.01, 0.16)	(0.03, 0.32)	(0.04, 0.36)
98% C.I.	(0.00, 0.10)	(0.00, 0.15)	(0.01, 0.19)	(0.01, 0.25)	(0.02, 0.45)	(0.02, 0.50)

TABLE B.4 (Cont'd)

Index	5 µg/dL	15 µg/dL	25 µg/dL	35 µg/dL	45 µg/dL	55 µg/dL)
Expert E, Hemoglobin ≤ ~~9.5~~ 11 g/dL, Ages 0–3 Years						
			Blood-Lead Level			
			Encoded Judgments			
Median	0.08	0.10	0.15	0.19	0.22	0.30
50% C.I.	(0.05, 0.14)	(0.07, 0.16)	(0.09, 0.19)	(0.14, 0.24)	(0.16, 0.35)	(0.17, 0.44)
90% C.I.	(0.03, 0.22)	(0.04, 0.24)	(0.05, 0.26)	(0.06, 0.28)	(0.09, 0.43)	(0.11, 0.51)
			Fitted Functions			
Median	0.09	0.11	0.13	0.16	0.21	0.28
50% C.I.	(0.06, 0.13)	(0.07, 0.15)	(0.09, 0.18)	(0.11, 0.22)	(0.15, 0.29)	(0.21, 0.37)
90% C.I.	(0.03, 0.20)	(0.04, 0.24)	(0.05, 0.28)	(0.07, 0.34)	(0.09, 0.42)	(0.13, 0.52)
98% C.I.	(0.02, 0.28)	(0.03, 0.33)	(0.03, 0.37)	(0.05, 0.44)	(0.06, 0.52)	(0.09, 0.62)
Expert E, Hemoglobin ≤ 11 g/dL, Ages 4–6 Years						
			Encoded Judgments			
Median	0.06	0.08	0.11	0.13	0.16	0.20
50% C.I.	(0.03, 0.10)	(0.05, 0.12)	(0.09, 0.15)	(0.10, 0.16)	(0.12, 0.24)	(0.13, 0.25)
90% C.I.	(0.01, 0.14)	(0.02, 0.16)	(0.04, 0.19)	(0.04, 0.20)	(0.06, 0.29)	(0.10, 0.33)
			Fitted Functions			
Median	0.05	0.07	0.10	0.12	0.15	0.18
50% C.I.	(0.04, 0.08)	(0.05, 0.10)	(0.07, 0.14)	(0.08, 0.17)	(0.11, 0.21)	(0.13, 0.24)
90% C.I.	(0.02, 0.13)	(0.03, 0.16)	(0.04, 0.22)	(0.05, 0.27)	(0.06, 0.32)	(0.08, 0.36)
98% C.I.	(0.01, 0.09)	(0.02, 0.22)	(0.03, 0.30)	(0.03, 0.35)	(0.04, 0.42)	(0.05, 0.46)

For all the experts, judgments about mean control group IQ and within-group IQ standard deviation were fit well by normal or lognormal distributions. The goodness of fit measure here, as before and in what follows, was the r-square statistic comparing actual judgments to those predicted by the fitted function.

With regard to mean IQ differences, some experts gave judgments for each lead level that were roughly symmetric around a particular value, and others gave judgments that were positively skewed, i.e., attributed small probabilities to very large differences. Generally, the former judgments were better fit with normal distributions and the latter with lognormal distributions, i.e., distributions over log IQ difference.

Specifically, four families of distributions were fit to each expert's judgments about mean IQ differences. One was the normal distribution allowing a separate mean and variance for each lead level, another was the lognormal allowing a separate mean and variance for each lead level, and the remaining two were equal variance normal and lognormal distributions. In the latter two cases, normal and lognormal distributions were fit respectively using a single pooled variance estimate for each, but allowing separate mean values at each lead level.

Generally, the normal function fit better than the equal-variance normal function, and the lognormal function fit better than the equal-variance lognormal function. However, each set of judgments were distinctly better fit by either the normal or the lognormal distributions, and only the better of the two types of distributions was considered further for that set.

As argued in Sec. 3.5, the equal-variance distributions have the virtue of never crossing and therefore never leading to inconsistencies. However, the equal variance fits were so poor for all experts but one, that there was no point in pursuing them in those cases. In order for the equal-variance fits to be acceptable, the slopes of the transformed judgments must be about the same. That generally was not the situation for the case of IQ decrements, making it undesirable to utilize equal-variance fits. Although

this causes the curves to cross at some point, the distributions are spaced sufficiently far apart, that the crossings occur in the very extreme tails. This is a small cost for representing the judgments accurately within the ranges of interest.

4.6 THE EXPERTS

In March of 1985 the individuals listed in Table 4.2 were selected by EPA/OAQPS staff for participation in the study. All agreed, again with the understanding that their judgments would be anonymous. The judgments presented here are attributed to Experts F through K, with these letter designations having been randomly assigned. We are grateful to these people for their hard work and the many hours they spent with us.

4.7 RESULTS

This section utilizes the functions fit to each individual's probabilistic judgments to present a summary of the quantitative results. The functions provide an accurate representation of the underlying encoded values, both because goodness of fit was excellent and because each expert endorsed the output of the respective functions as representing his or her judgments.

Summaries of the qualitative discussions about the IQ and behavioral effects of lead are in Sec. C.4 of App. C. Detailed quantitative results, including the encoded judgments, parameter values for the fitted functions, goodness of fit measures, and derived subjective probabilities about dose-response functions, are listed in Appendix C.

Recall that the experts were given the option of making separate judgments for lower and higher SES children. The lower SES group was defined as those children whose family incomes do not exceed the fifteenth percentile; the remaining children were to be considered in the higher SES group. All the experts except F believe that at the doses under consideration, lead interacts with variables that contribute to SES level. Therefore, all except F elected to provide separate judgments for the two SES levels.

It is convenient to present simultaneously the judgments about the lower and higher SES control group mean IQ values, and to do the same with the judgments about the within-group standard deviations. However, the judgments about the mean IQ decrements are most easily comprehended if they are shown separately for the two SES levels. F's judgments will be repeated in the cases for purposes of comparison.

4.7.1 Control Group Mean IQ

Expert I felt unable to provide judgments about the mean IQ of a group of children sheltered from lead exposure. The other experts' judgments are shown in Figure 4.1, categorized by SES level. The top four bars represent the low SES judgments of G, H, J, and K; the central bar represents the single overall judgment of F; and the lower four bars represent the high SES judgments of G, H, J, and K.

The symbol within each bar is the median judged mean IQ value. In other words, for a given bar, according to that expert, there is a 0.50 probability that the control group mean IQ would be above the value indicated by the symbol and a 0.50 probability that it would be below the indicated value. The brackets immediately to the left and right of the symbol denote each expert's central 50% credible interval, i.e., the interval such that there is a 0.25 probability of the mean IQ falling below it, a 0.50 probability of the mean IQ falling within it, and a 0.25 probability of the mean IQ falling above it. The other pair of brackets denotes the central 90% credible interval.

Experts G, H, and J show considerable agreement in their judgments regarding the low SES group. Their median judgments about the mean IQ range from 95 to 97 points, and (from their 90% credible intervals), with probability 0.05, the mean IQ exceeds a value that is between 98 and 103 points. Expert K's median judgment is 85 points, and according to him or her, the mean IQ exceeds 92 points with probability 0.05.

The judgments of all four experts overlap considerably regarding the high SES group. Their median judgments range only from 104 to about 106 points. The upper limits of the 90% credible intervals range from 108 to 112 points.

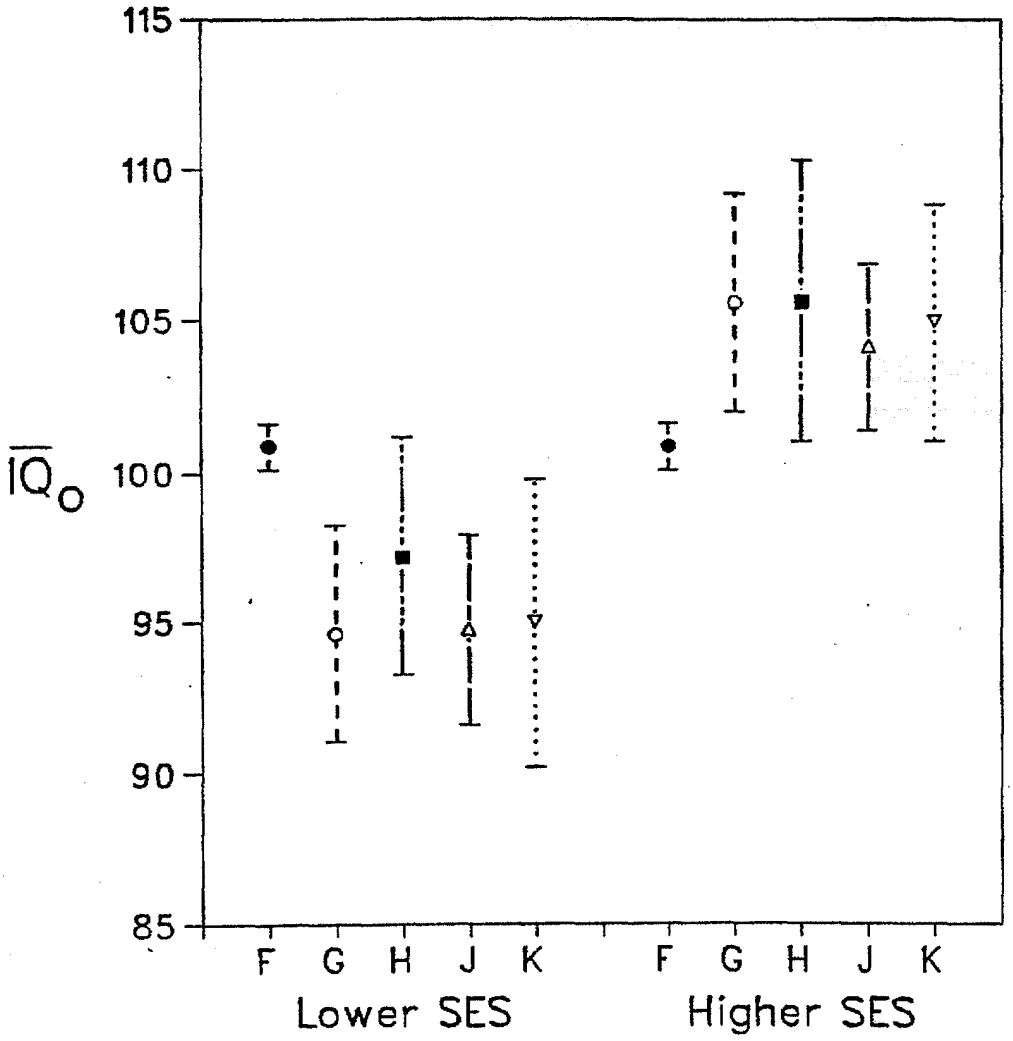

FIGURE 4.1 Judgments about Control Group Mean IQ

Expert F displays the least uncertainty. His or her judgments about the entire population fall between those of the other experts for the low and high SES levels, as they should. The median judgment is 101 points; with probability 0.05, the mean IQ exceeds 102 points.

4.7.2 Within-Group IQ Standard Deviation

Expert I felt unable to provide judgments about this variable, as well. The judgments of the remaining experts are shown in Fig. 4.2, arranged according to the same format as Fig. 4.1.

Expert F was certain that the within-group IQ standard deviation would be identical to that normed for the population as a whole, 15 points. Expert G, considering each of the two SES groups to be somewhat more homogeneous than the overall population, was certain that the within-group IQ standard deviation would be 14 points in each case.

The judgments of experts H, J, and K were relatively similar. They thought that the IQs of the low SES group would be more homogeneous than those of the high SES group, and therefore estimated lower standard deviations. The median judged standard deviation for the lower SES level range from 12.4 to 13 points; upper limits of the 90% credible intervals range from 14 to 15 points. Expert J was certain that the standard deviation would be 14 points for the upper SES level, while H and K have median estimates of 13.5 and 14.3 points, respectively. The upper limits of their 90% credible intervals are 15.6 and 15.1 points, respectively.

4.7.3 Mean IQ Decrements for the Low SES Group

Figure 4.3 summarizes the judgments of all six experts regarding mean IQ decrements for the low SES group. Each person's judgments are shown in a separate panel. Blood lead is on the abscissa and mean IQ decrement (mean control group IQ minus mean exposed group IQ) is on the ordinate.

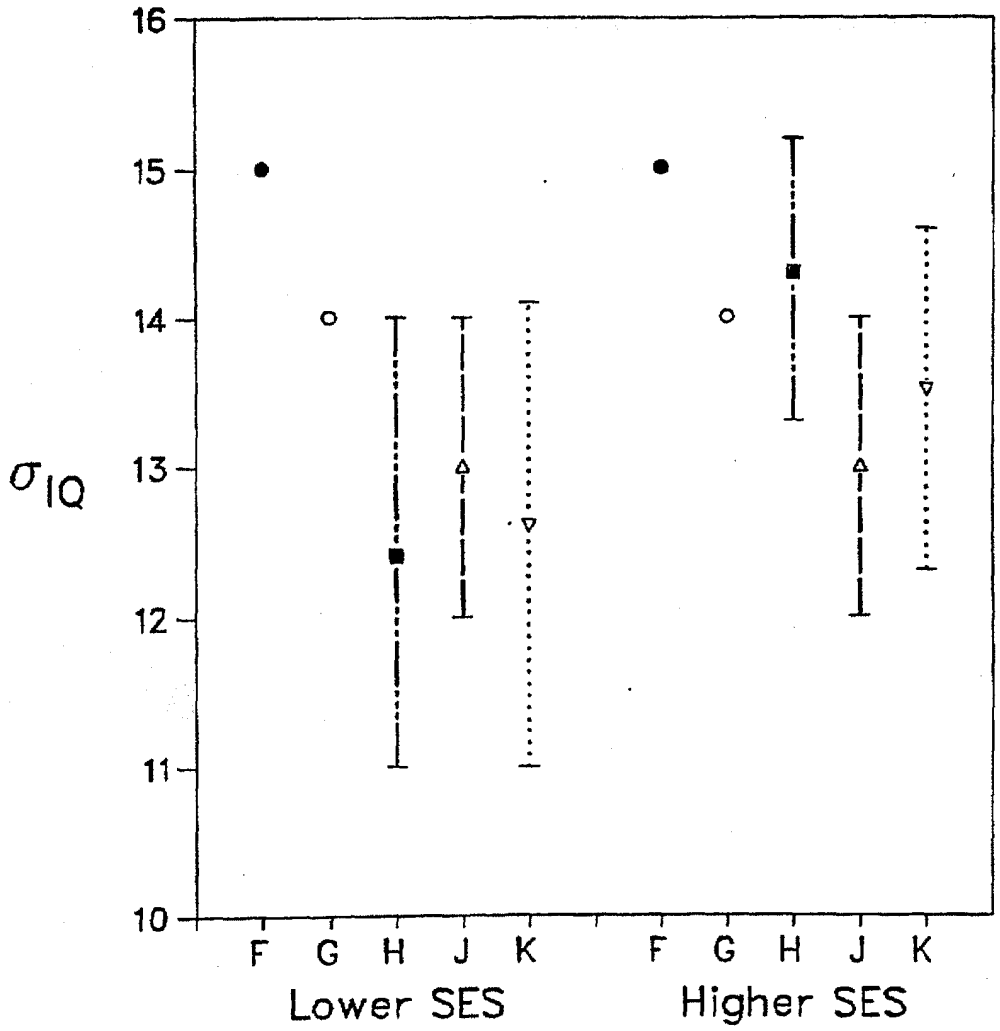

FIGURE 4.2 Judgments about σ_{IQ}

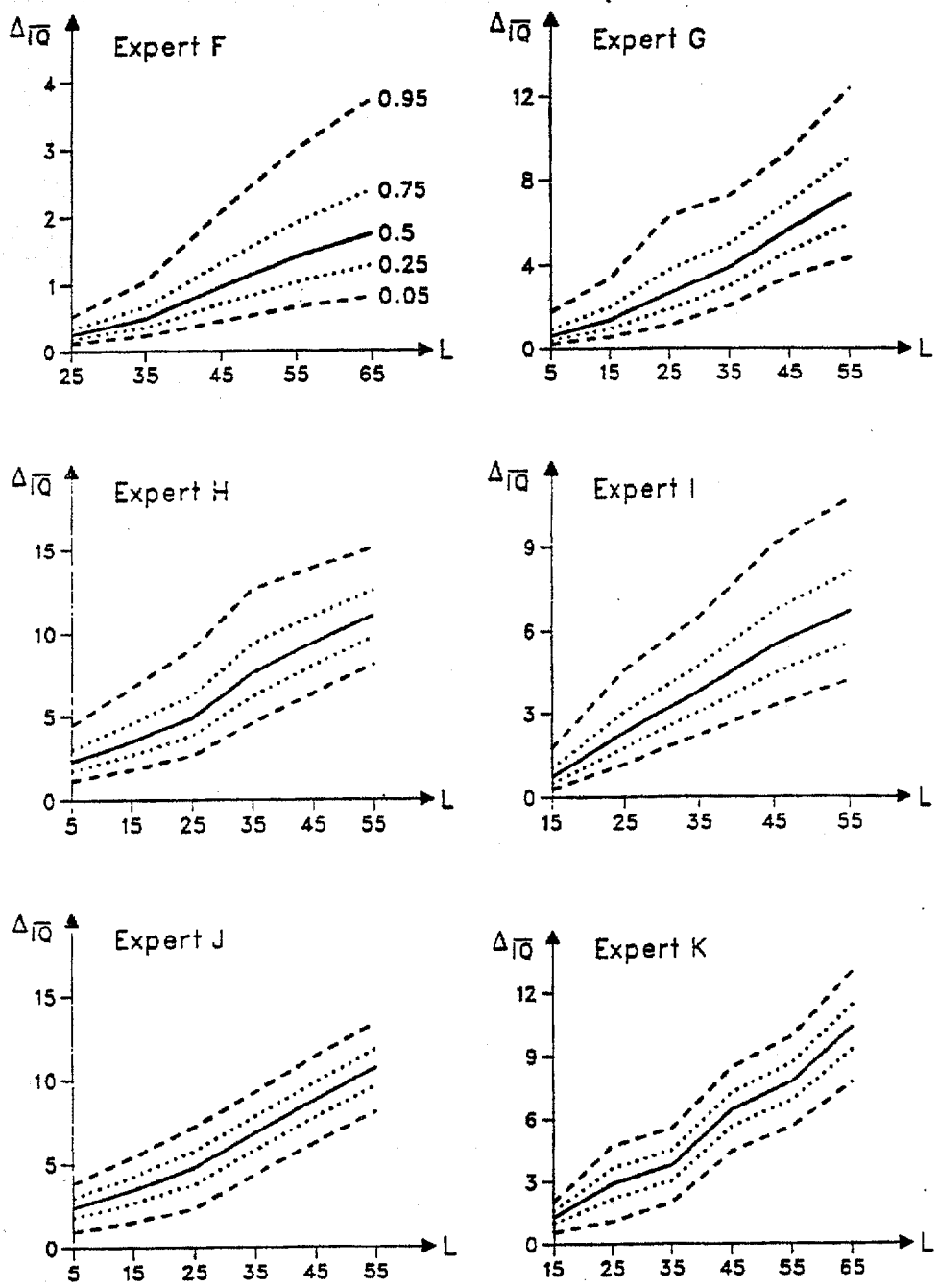

FIGURE 4.3 Judgments about $\Delta_{\overline{IQ}}$ for a Lower SES Population

The dark, central curve in each panel shows the median judged IQ decrement for each lead level. In other words, for a given panel, according to that expert there is a 0.50 probability that the actual mean IQ decrement would be greater than the indicated value, and a 0.50 probability that it would be less. The successively lighter pairs of curves that bracket the median curve represent central 50%, and 90% credible intervals.

Figure 4.4 allows a comparison of judgments across the experts. The axes are the same as in Figure 4.3. The vertical bars at each lead level represent each expert's central 90% credible interval, and the symbols are his or her median judgments.

Expert F consistently judges the IQ effects of lead to be less than do the other experts, and evidences considerably less uncertainty about the magnitude of these effects than do the others. Expert F is certain that there is no IQ effect of lead up to at least 15 μg/dL. At 25 μg/dL, the median judged IQ decrement is about 0.5 point, and this increases to a median judged IQ decrement of just under 2 points at 65 μg/dL. According to F, at 25 μg/dL, the IQ decrement exceeds approximately 1 point with probability 0.05, while at 65 μg/dL, it exceeds 5 points with the same probability.

There are overriding similarities in the judgments of the other experts, although there are also small, consistent differences among them. Thus, only G, H, and J give any credibility to there being IQ effects as low as 5 μg/dL, while I does so at 15 μg/dL, and K does at 25 μg/dL. The judgments of H and J are consistently very close, as are those of G, I, and K, which as a group are somewhat lower than those of H and J. Considering the five sets of judgments, the median judged IQ decrement at 5 μg/dL ranges from 0 to 2.5 points. The median judged IQ decrement at 55 μg/dL is from about 7 to 11 points. According to the judgments of G, H, and J, with probability 0.05, the IQ decrement exceeds 3 to 5.5 points at 5 μg/dL, while according to all 5 experts, it exceeds 11 to 17 points at 55 μg/dL with the same probability.

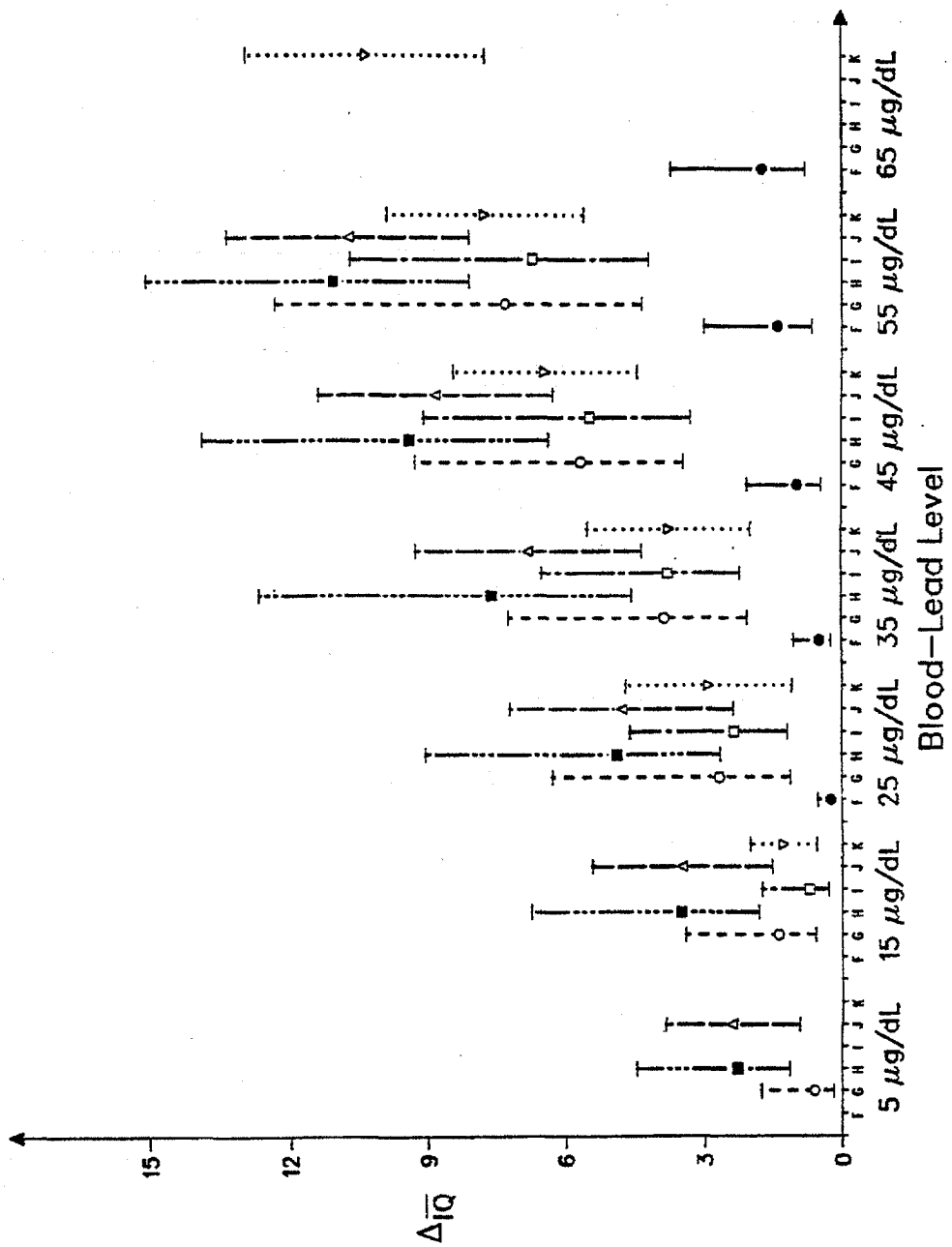

FIGURE 4.4 Comparison of Judgments about $\Delta_{\overline{IQ}}$ for a Lower SES Population

4.7.4 Mean IQ Decrements for the High SES Group

Figures 4.5 and 4.6 display the judgments about mean IQ decrement for the high SES group, and are read in the same fashion as are Figs. 4.3 and 4.4, respectively. As with the low SES population, F consistently judges the IQ effect to be less than do the other experts. From 25 μg/dL on, the judgments of the others overlap, with, as before, those of H and J being very similar and somewhat greater than those of G, K, and I, which themselves are similar. Only H gives any credibility to the existence of an IQ effect at 5 μg/dL, G, I, and J do so at 15 μg/dL, and F and K concur at 25 μg/dL.

Considering all the experts simultaneously, the median judged IQ decrement at 15 μg/dL ranges from 0 to 2.5 points; at 55 μg/dL, it ranges from 1.5 to 7.5 points. According to G, H, I, and J, with probability 0.05 it exceeds 1 to 7 points at 15 μg/dL, while according to all the experts it exceeds 4 to 11.5 points at 55 μg/dL with probability 0.05.

4.7.5 Change in Percentage of Low SES Group with IQs $\leq$ 85

The judgments about control group mean IQ, mean IQ decrement at each exposure level, and within-group IQ standard deviation can be used along with the assumption of within group normal IQ distributions to calculate subjective probabilities about the lead induced change in percentage of the population with IQ's below any critical level. These would correspond to judgments about IQ dose-response functions, and are the results necessary ultimately for the risk calculations. The mathematical details are in Appendix D. Here, we illustrate the results for a critical IQ of 85 points. We point out, however, that greater uncertainty is evidenced in these dose-response functions than in the separate judgments shown above, because uncertainty about three parameters is being combined. Note that calculations are not possible for expert I, because he or she declined to judge control group mean IQ and within-group IQ standard deviation.

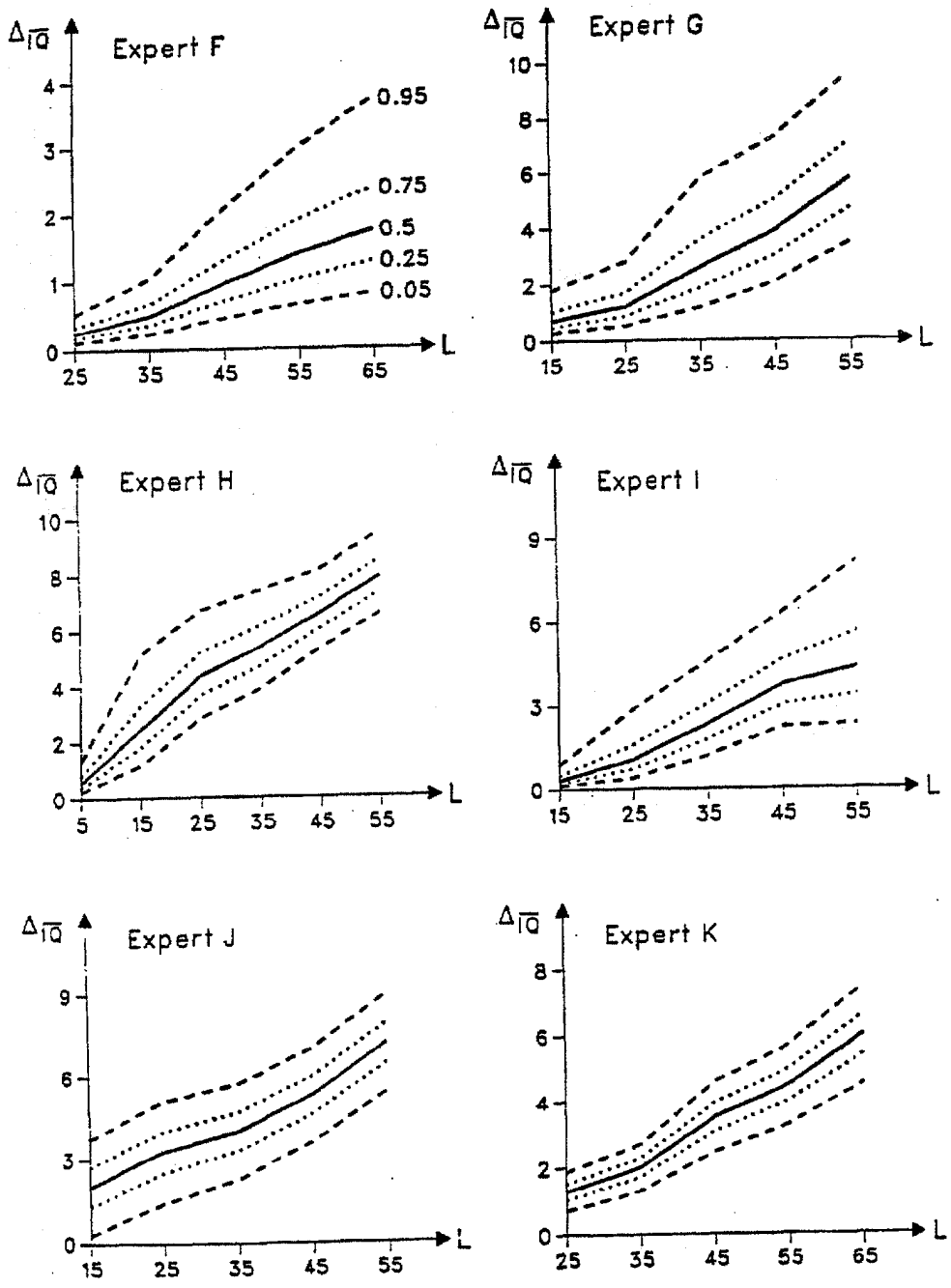

FIGURE 4.5 Judgments about $\Delta_{\overline{IQ}}$ for a Higher SES Population

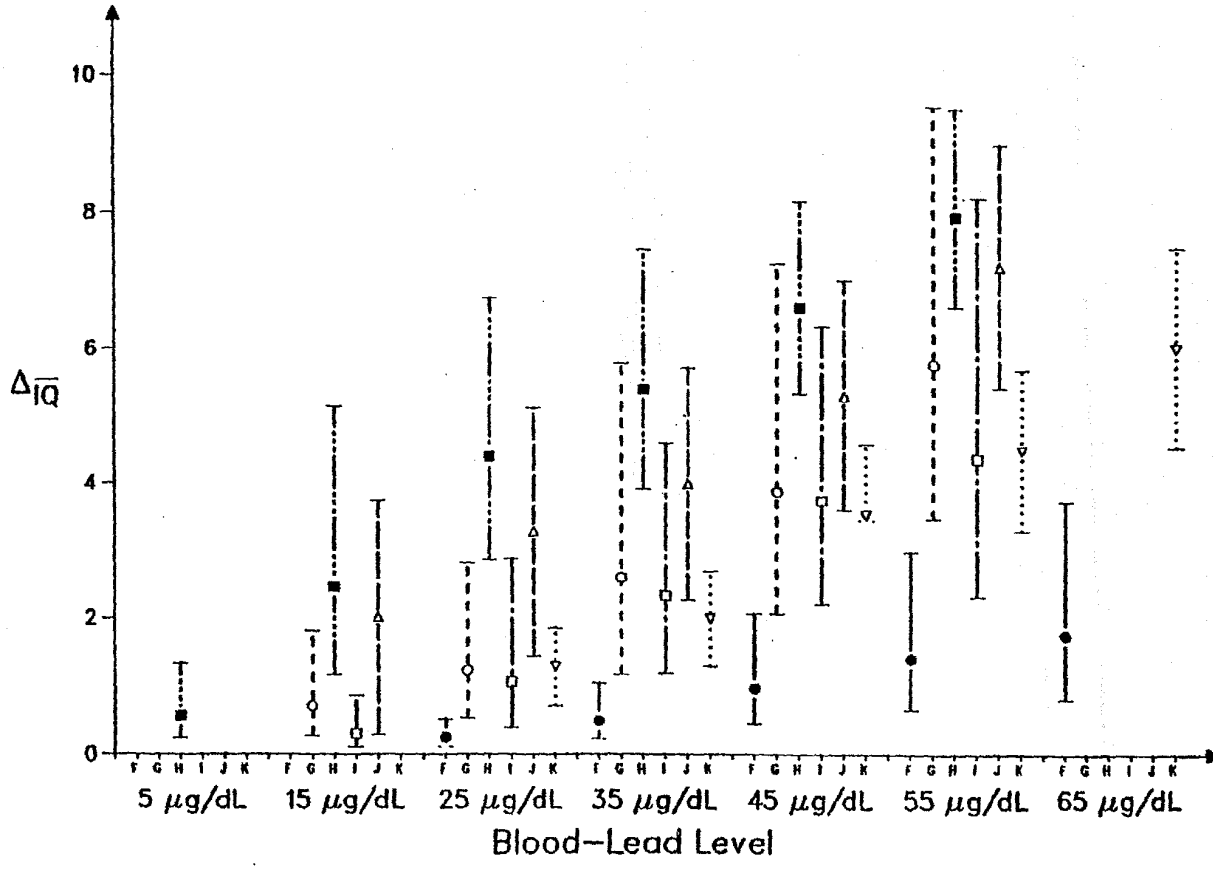

FIGURE 4.6 Comparisons of Judgments about $\Delta_{\overline{IQ}}$ for a Higher SES Population

Figures 4.7 and 4.8 show the calculated subjective probabilities about the dose-response function for increase in percentage of IQ's less than or equal to 85 points in the low SES population. The abscissa in both cases is blood lead, the ordinate is population response rate (i.e., lead induced increase in percentage below 85 points), and the curves are analogous to those in the preceding figures.

As would be expected, F's judgments suggest much lower dose-response functions than do those of the others. The results from G, H, J, and K overlap considerably, with those from H and J being very similar and slightly greater than those from G and K, which themselves are similar.

According to F, the median judged response rate is less than 1% at 25 μg/dL, rising to 4% at 65 μg/dL. With probability 0.05, it exceeds 1% at 25 μg/dL, and with the same probability, it exceeds 7% at 65 μg/dL. For G, H, and J, the median response rate is from 2.5% to 6% at 5 μg/dL, while including K, it is between 21% and 32% at 55 μg/dL. For the same experts, the response rate exceeds about 4% to 11% at 5 μg/dL, 14% to 26% at 25 μg/dL, and 29% to 45% at 55 μg/dL, each with probability 0.05.

4.7.6 Change in Percentage of High SES Group with IQs $\leq$ 85

The pattern of similarities and differences across experts is the same here as it was above, so corresponding figures are not shown. Generally, higher SES children are less at risk than lower SES children. The results for F are the same as those shown previously, because they are based on the same judgments. According to G, H, and J, the median judged response rate is from about 0.5% to 3% at 15 μg/dL, while according to G, H, J, and K, it is from about 5% to 12% at 55 μg/dL. According to the same experts, the median response rate exceeds 2% to 7.5% at 15 μg/dL, 4% to 11% at 25 μg/dL, and 9.5% to 18% at 55 μg/dL, all with probability equal to 0.05.

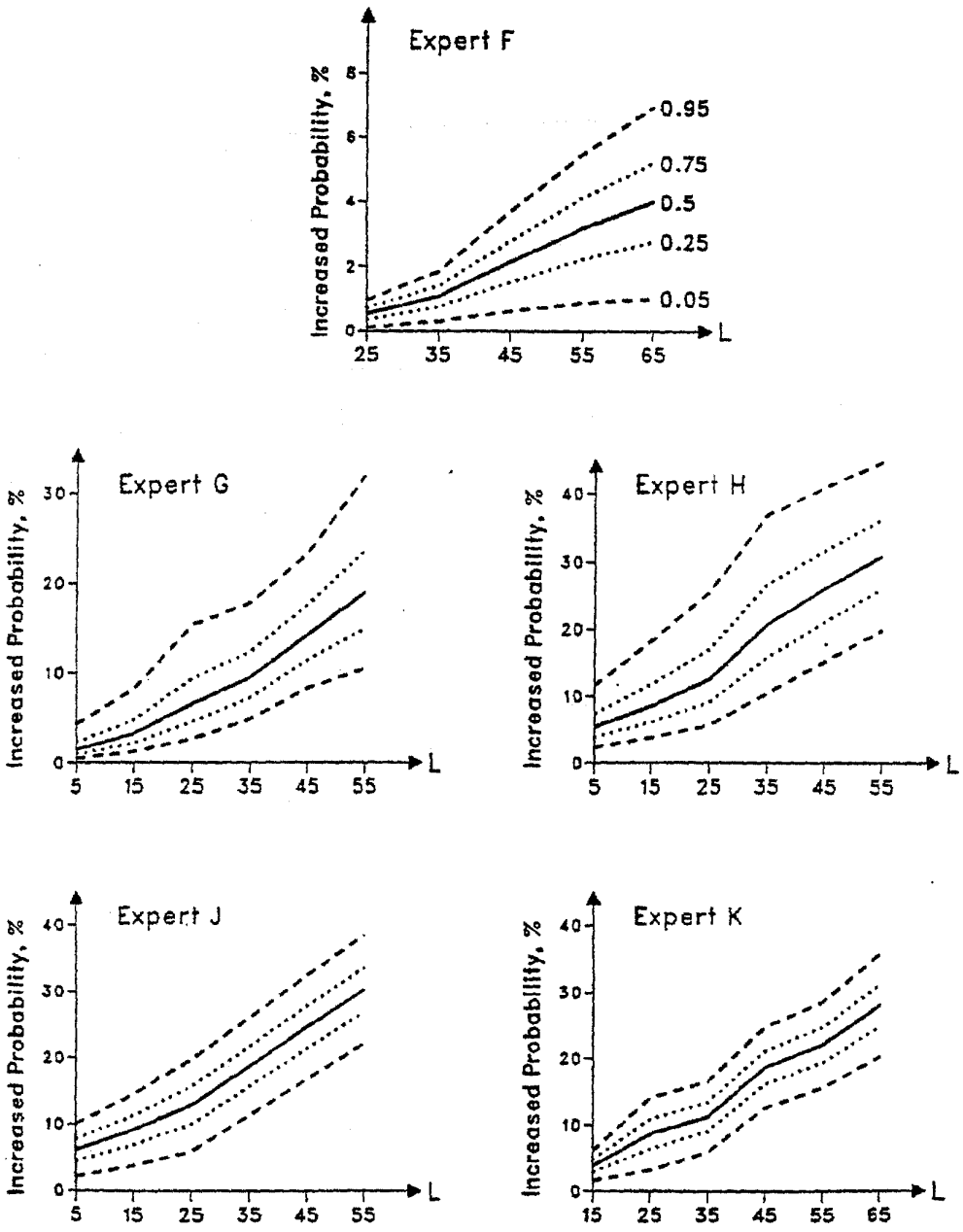

FIGURE 4.7 Increased Probability of Having IQ ≤ 85 for a Lower SES Population

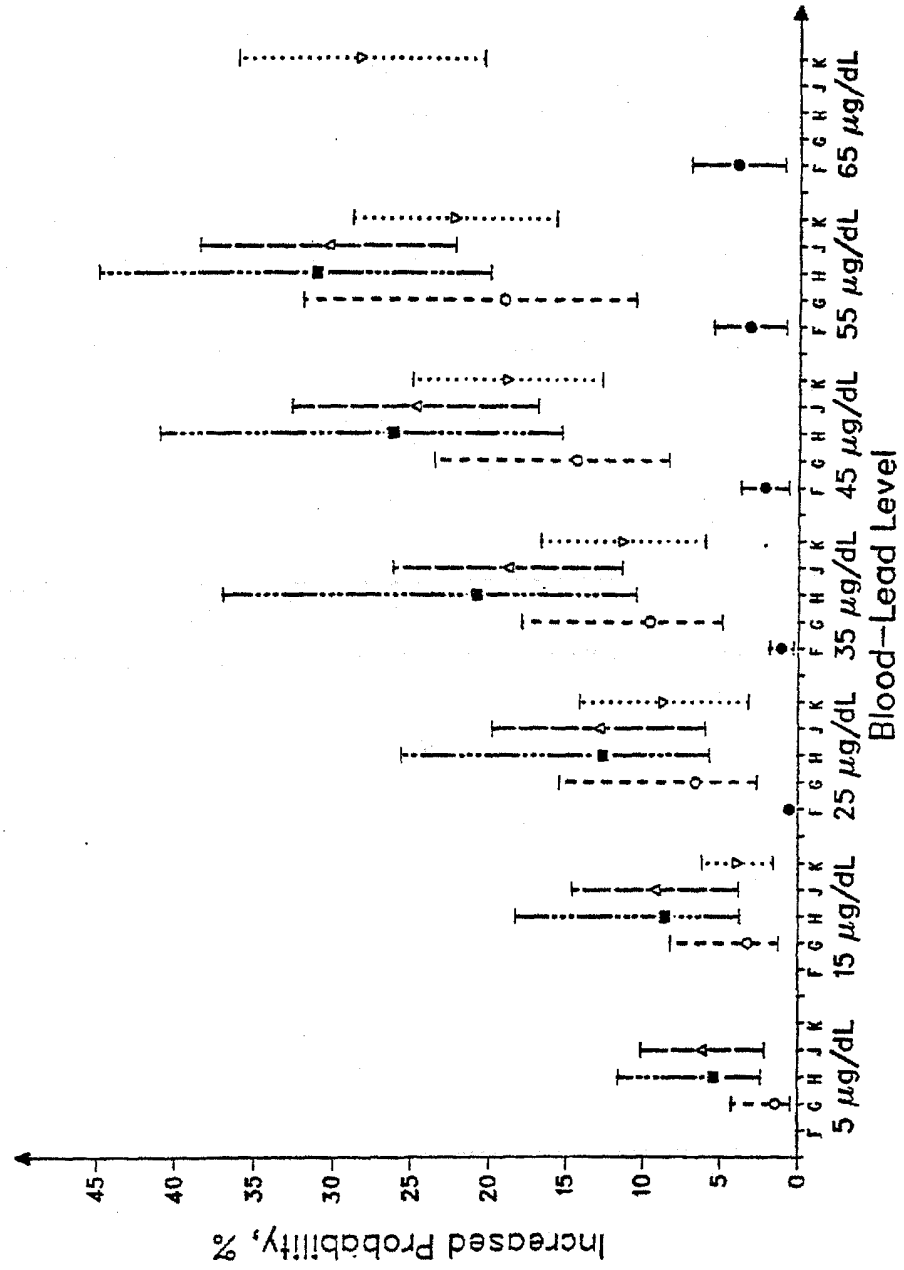

FIGURE 4.8 Blood-Lead Level Comparison of Increased Probabilities of Having IQ ≤ 85 for a Lower SES Population

4.8 DISCUSSION

Considering the extensive debate concerning the IQ effects of lead, the degree of consensus reflected in the present results is notable. This is particularly so, since the experts were selected so as to span the full range of opinion. As occurred with the hemoglobin results, asking the experts to encode their subjective probabilities about specific, well-defined scientific outcomes eliminated disagreements about definitions, policy, and other matters. The remaining disagreements, evidenced in the preceding results, are due to differing interpretations of, and extrapolations from, evidence.

Generally, Experts H and J provided similar judgments, as did Experts G, I, and K. Furthermore, these two sets of judgments overlapped to a great extent. Expert F consistently estimated smaller effects than did anyone else, and indicated less uncertainty about them.

Uncertainty (as indicated by variance) generally increased with blood lead within the range of interest, as would be expected. In addition, individual subjective probabilities cover broader ranges for the low SES than for the high SES group. This result may reflect more data at the high SES level, or more uncertainty about the influence of the covariates at the low SES level.

The judgments regarding both mean IQ decrements and dose-response functions are of interest in their own right. However, in Sec. 5 they will be combined with exposure estimates to yield risk distributions.

5 ESTIMATED RISKS OF ADVERSE HEALTH EFFECTS UNDER ALTERNATIVE AIR-LEAD SCENARIOS

Risks of adverse health effects under alternative air-lead standards can be estimated by combining dose-response functions with blood-lead distributions appropriate to the alternative standards. The estimates must incorporate the uncertainty associated with dose-response functions as well as that associated with blood-lead distributions. The former has been discussed in Secs. 2-4 for EP, hemoglobin, and IQ, respectively; the latter will be covered in this section.

The main focus of this section is to estimate the risks of adverse health effects under alternative air-lead scenarios for young children in the vicinity of several point sources of lead located in the U.S. Results were obtained by mathematically combining the PbB and dose-response information into a risk model. The exposure information (i.e., PbB levels for alternative air-lead scenarios) was developed by PEI Associates, Inc., under contract to EPA, using the integrated lead-uptake/ biokinetic model presented in the EPA staff paper. Since methodological development is not complete, these results are preliminary.

This section is organized into five main subsections. The first covers alternative air-lead scenarios and estimated resultant PbB distributions among children residing in the vicinity of selected point sources of lead. The next three subsections present estimates of the risks of adverse health effects involving EP, hemoglobin, and IQ, respectively.

5.1 EXPOSURE ESTIMATES

EPA's OAQPS used an integrated lead-uptake/biokinetic model to estimate the relationships between air quality and blood-lead levels among U.S. children residing in the vicinity of point sources of lead emissions. The model incorporates a biokinetic-metabolic model developed at New York University, which takes into account the

absorption of lead into the bloodstream through inhalation and ingestion, and its subsequent movement among body compartments (i.e., blood, bone, and soft tissue). Blood lead is the index of lead exposure generated by the model for present purposes. Estimates of exposure take into account changes in activity patterns, physiology, and metabolism of children through various stages in their development. For example, children six months and six years of age differ in terms of numerous physiological indicators, as well as concentration of lead in the immediate environment, degree of hand-to-mouth activity, time spent outdoors vs. indoors, differences in lead absorption in the gastro-intestinal tract at different ages, etc. For this analysis, only children that live in housing that is free of lead-based paint hazard, and who are not exposed indirectly to occupational sources of lead from household members, were modeled.

Blood-lead levels for children living in the vicinity of five U.S. point sources of lead emissions were included in our analysis. For the area surrounding these point sources, which will remain nameless at this time, the total number of children (0-6 years), the number of children in the 84-month cohort, and the percentage of those children living in households having incomes below poverty level are listed in Table 5.1. This level approximately corresponds to the lower SES level (the lowest 15%, based on family income) used in the IQ encodings.

Seven alternative air-lead scenarios were considered for each of the point sources. These are grouped in three categories: baseline, precontrol, and postcontrol. For illustrative purposes, air-quality estimates based on dispersion modeling used in the exposure analysis are presented in Table 5.2 for areas surrounding the five point sources under different air-lead scenarios. It is important to note that these air-lead values represent estimates at the geographic centers of the census tracts surrounding the different point sources. Air-lead concentrations at other receptor sites closer to the point sources are generally higher than the ranges shown. The baseline scenario represents modeled air quality as it existed around the particular point source for the

TABLE 5.1 1980 Population Data[a] for Selected U.S. Point Sources of Lead Emissions

Point Source	Type	Total Number of Children Aged 0-6 Years	Number of Children in 84-Month Cohort[b]	Percent below Poverty Level[c]
1	Secondary Smelter	15,835	2,273	14
2	Secondary Smelter	4,614	671	24
3	Battery Plant	23,422	3,305	19
4	Battery Plant	12,956	1,994	11
5	Primary Smelter	12,576	1,718	6

[a]Numbers calculated based on 1980 U.S. Bureau of Census data and aggregated across census tracts around each point source. Population data represent the number of children living in housing free of lead-based paint hazards and exposure to occupational sources of lead.

[b]In this case, the 84-month cohort is a group of children born in January 1974 and already 7 years old at the end of the exposure period.

[c]Data are from Ferdo (1986).

TABLE 5.2 Comparative Air Quality at Census Tract Centroids Surrounding Modeled Point Sources

Point Sources	Range of Monthly Air-Lead Concentrations[a]			
	1974-1980	1983-1989	1990-1996 (0.25 $\mu g/m^3$)[b]	1990-1996 (1.5 $\mu g/m^3$)[b]
1	0.53-1.83	0.01-0.92	0.01-0.03	0.01-0.13
2	0.54-1.83	0.02-0.92	0.01-0.03	0.01-0.13
3	0.53-1.83	0.01-0.92	0.01-0.03	0.01-0.13
4	0.53-1.83	0.01-0.92	0.01-0.03	0.01-0.13
5	0.01-1.27	0-0.64	0-0.04	0-0.23

[a]Monthly air-lead concentrations expressed in micrograms per cubic meter and rounded to two significant digits were estimated by another contractor for EPA using ISC dispersion-model and gasoline-lead-consumption data (based on previous, current, or future emission rates) at geographic centroids of census tracts surrounding each point source under different scenarios (Johnson and Paul, 1986). Ranges reflect spatial variability around point sources (decreasing air-lead levels with increased distances) and temporal trends (reduction in gasoline lead emissions since the 1970s).

[b]The 1990-1996 scenarios assume that the modeled point source is in compliance with whatever lead NAAQS is being analyzed. For example, the highest air-lead concentration projected at an air quality receptor under the 0.25 $\mu g/m^3$ lead NAAQS is no higher than 0.25 $\mu g/m^3$, with all other receptor points just at or below that level. Air quality values at geographic centroids are considerably below other receptor values because of distance from the point source and an assumption that air-lead levels in an area decline linearly with reductions in emission rates.

1974-1980 time period. Similarly, the precontrol period covers the years 1983-89. Blood-lead levels are estimated to be lower in the precontrol vs. baseline periods primarily because of the beginning of the gasoline lead phasedown, the implementation of the ambient air quality standard for lead, and reductions of lead levels in canned foods. Lead exposure in the U.S. also declined in this era due to additional factors such as increased awareness of lead-based paint hazards and the gradual reduction in the old, lead-based painted housing stock. The postcontrol period spans 1990-96, and includes five alternative air-lead-level scenarios (lead NAAQS of 0.25, 0.5, 0.75, 1, and 1.5 $\mu g/m^3$).

For each alternative air-lead scenario considered in the postcontrol period, it is assumed that the point source is in attainment with whatever standard is analyzed. In other words, the highest air-quality receptor site would be at or below whatever standard level is considered. If the dispersion modeling indicated that a standard level would be exceeded in the postcontrol scenario, the highest receptor site was adjusted down to the standard level, with all other receptor sites reduced proportionately. The reader is referred to a lead exposure report (Johnson and Paul, 1986) that details the method used to model air quality for each scenario. Blood-lead levels in the postcontrol period are estimated to be further reduced primarily because of the continued reductions in lead levels in canned foods and drinking water.

For each air-lead scenario, upper and lower bound estimates of the geometric mean blood leads were calculated by the exposure model. These bounds were obtained by considering a range of "high" and "low" settings of parameter values that significantly affect final PbB estimates. Many of the alternative postcontrol scenarios resulted in mean PbB levels that were essentially equivalent and relatively low (< 5 $\mu g/dL$). This suggests that the small incremental changes in ambient air quality levels under the alternative NAAQS analyzed may not be sufficient to significantly influence PbB distributions among populations of children in the future. The result may be misleading,

however, because major air-quality impacts are confined principally to the immediate vicinity (< 1 km) of the point sources, while the modeled blood leads were obtained by averaging over relatively large areas around the point sources (resulting in air-lead levels down to urban or rural background concentrations at the boundaries of the point source impact areas). A complete tabulation of the scenarios used in the risk calculations is provided in App. D.

Each scenario contains upper- and lower-bound estimates of the geometric means of modeled blood leads for the 84-month cohort for four age periods: 0-3 years, 4-6 years, 0-6 years, and the 36th month of life. Estimates for the first two periods are needed in order to combine expert judgments about hemoglobin decrements, which were obtained for subpopulations of children of those ages. The third is needed to estimate the risks of elevated EP levels, and the last is needed to estimate the risk of IQ effects.

All risk estimates presented are for the 84-month cohorts modeled in each air-lead scenario. The 84-month cohort is the cohort of children born in the beginning of each exposure period, and who reach their seventh birthdays at the end of the exposure period. The 84-month cohorts constitute only a small fraction of the total number or children aged 0-6 years surrounding each point source modeled. In addition, risk estimates for lead-induced IQ effects are presented solely for children from families of children from lower SES status. This group of children constitutes 6-24% of the total poulation of children around each of the five point sources (Table 5.1). As a result, any headcount risk numbers that would be calculated based on these risk estimates would underestimate the total risk in the population of young children modeled. At this time, headcount risk estimates are not presented.

It is assumed that blood leads are described by lognormal distributions with the geometric mean (GM) values tabulated, and a single geometric standard deviation (GSD) of 1.42. Properties of the lognormal distribution are described in App. B. The geometric mean (GM) blood leads for all 84-month cohorts for the three scenarios that are

significantly different from one another (for Sources 2 and 5) are listed in Table 5.3. All of the postcontrol scenarios result in PbB levels that are nearly identical, so only one postcontrol scenario (having upper and lower GM PbB values) is listed. Geometric mean blood-lead values are lowest around Source 5, which is located in a rural area, and highest around Source 2, an urban area. The others (also urban areas) are slightly less than those for Source 2. To illustrate, the upper-bound estimate of GM = 27 μg/dL for the baseline exposure scenario for Source 2 means that 55% of the population would have had PbB levels above the current Centers for Disease Control definition of an elevated PbB level of 25 μg/dL. The corresponding lower bound of GM = 19.4 μg/dL means that 33% of the population would have exceeded that level.

As a point of reference, over the years 1976-80, the second National Health and Nutrition Survey (NHANES II) (Annest et al., 1982) was conducted. It included representative examinees from urban and rural populations aged six months to five years. Urban examinees had a GM PbB level of approximately 17 μg/dL. This value is about 0.2 μg/dL lower than the lower bound GM PbB estimate for children aged 0-6 years at Source 2. Rural children had an average PbB level of 13.9 μg/dL, which is about 0.5 μg/dL higher than the upper bound for Source 5 for the baseline scenario.

5.2 RISK RESULTS FOR EP

A complete listing of results and descriptions of the models used to estimate the risk are provided in App. D. An overview is presented here. As discussed in Sec. 2 and App. A, EP dose-response uncertainty was calculated from the data published by Piomelli et al. (1982). Recall that two levels of EP were considered to be adverse health effects, 33 μg/dL and 53 μg/dL, which are one and two standard deviations above the "reference" EP level (21.7 μg/dL) determined for the group of children having the lowest PbB levels.

The probabilistic dose-response functions were combined with the PbB distributions (for children aged 0-6 years) described in Sec. 5.1 to produce the overall risk

TABLE 5.3 Geometric Mean Blood-Lead Levels (μg/dL) for 84-Month Cohorts[a]

	Source 5		Source 2[b]	
Age Group	Lower	Upper	Lower	Upper
Baseline Scenario (1974-80)				
0-3 Years	8.6	13.9	19.4	27.0
4-6 Years	6.4	11.5	14.3	21.2
0-6 Years	7.7	13.4	17.2	24.5
36th Month	7.7	12.9	19.2	27.1
Precontrol Scenario (1983-89)				
0-3 Years	3.3	6.7	5.2	9.9
4-6 Years	2.0	5.5	2.4	5.9
0-6 Years	2.7	6.2	4.0	8.2
36th Month	2.9	6.5	4.4	9.7
Postcontrol Scenarios (1990-96)				
0-3 Years	1.4	3.7	1.4	3.6
4-6 Years	1.1	3.6	1.1	3.4
0-6 Years	1.3	3.7	1.3	3.7
36th Month	1.3	3.7	1.3	3.7

[a]Estimated PbB levels are averages based on air-quality estimates for geographic centers of all census tracts within the impact area surrounding a given point source.

[b]Values for other sources are about 5-10% lower than those for Source 2.

results for EP. These results are summarized in Figs. 5.1 and 5.2 for the 33-μg/dL and 53-μg/dL EP levels, respectively. Each figure shows results for six air-lead scenarios (lower- and upper-bound GM values for baseline, precontrol, and postcontrol time periods) for Sources 2, [illegible] and 5. The full risk estimate for each scenario is a probability distribution over a number or fraction of the population, but only the median and 90% credible intervals are shown. The median is indicated by an X, and the 90% credible interval is indicated by bars extending both above and below the median value.

For the precontrol and postcontrol scenarios for areas surrounding all sources but Source 5, EP risk estimates are approximately at background levels. Specifically, the median values are about 10.7% for the 33-μg/dL EP level, and 2.1% for the 53-μg/dL EP level. The 90% credible intervals around these median values are about 6-15% (for 33 μg/dL) and 0.1-5% (for 53 μg/dL).

For the baseline scenarios and Sources 1-4, the median response rates are about 21% (for 33 μg/dL) and 8% (for 53 μg/dL) for the lower-bound GM PbB values and are about twice those rates for the upper-bound GM PbB values. The uncertainty about these values is quite small, indicated by 90% credible intervals that are about equal to the median values plus or minus 1-2%. The narrow 90% credible intervals are due to the large numbers of children in the higher PbB groups examined by Piomelli et al. (1982), which dominate the risk calculations for the baseline scenarios for Sources 1-4.

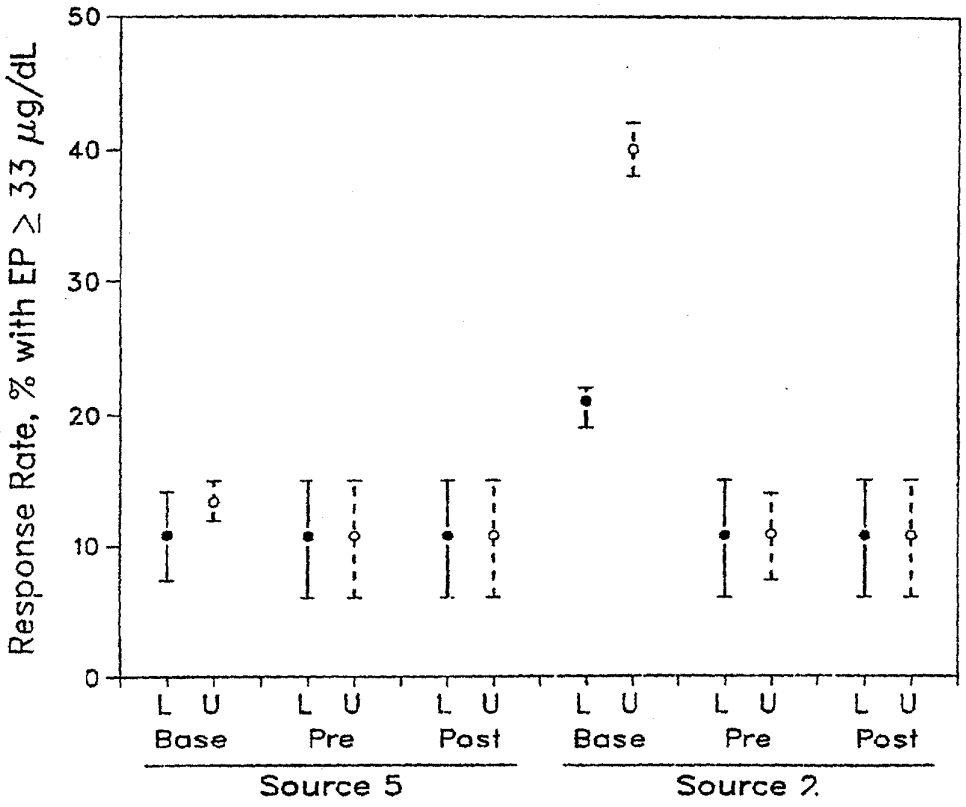

FIGURE 5.1 Response Rate, %, with EP ≥ 33 µg/dL

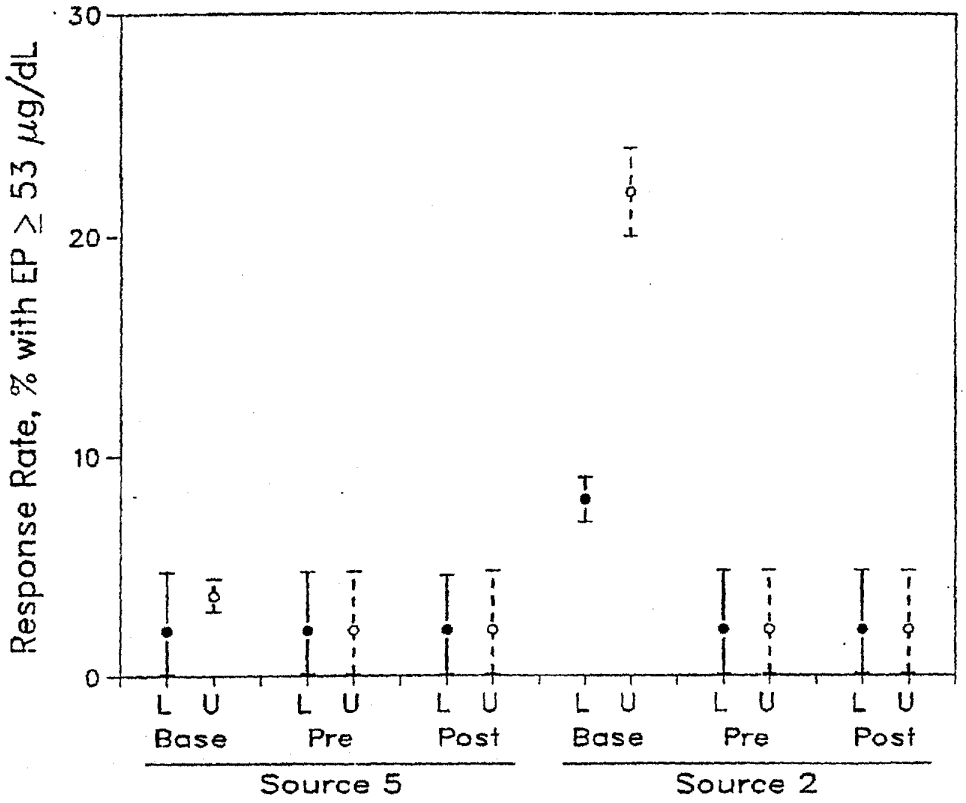

FIGURE 5.2 Response Rate, %, with EP ≥ 53 µg/dL

5.3 RISK RESULTS FOR HEMOGLOBIN

Recall from Sec. 3 that the probabilistic judgments of four experts regarding dose-response functions for lead-induced hemoglobin decrements were obtained for two hemoglobin levels considered to be below normal (9.5 g/dL and 11 g/dL). In a fashion exactly the same as that for calculating the risks of adverse EP effects, functions that the experts agreed represented their judgments (cf. Sec. 3) were combined with the various PbB distributions described in Sec. 5.1, to produce estimates of overall risk.

The results of the calculations (App. D) are summarized in Figs. 5.3-5.6. Results are shown only for children aged 0-3 years, who are more sensitive to lead exposure than children aged 4-6 years. Median risk estimates for children aged 4-6 years and for children aged 0-6 years are just slightly lower than those for ages 0-3 years. The upper ends of the credible intervals for ages 0-3 years are about 25% higher than those for ages 0-6 years.

Figures 5.3 and 5.4 are for Source 5, for the 11-g/dL and 9.5-g/dL hemoglobin levels, respectively. Figures 5.5 and 5.6 are analogous and for Source 2. In these figures, median values of the risk distributions and 90% credible intervals around those values are indicated for each air-lead scenario, based on the functions fit to the judgments of Experts A, C, D, and E. The median is indicated by an X, and the 90% credible interval is indicated by bars extending both above and below the median value. The bars are of various textures to help one to visually discriminate results for the individual experts.

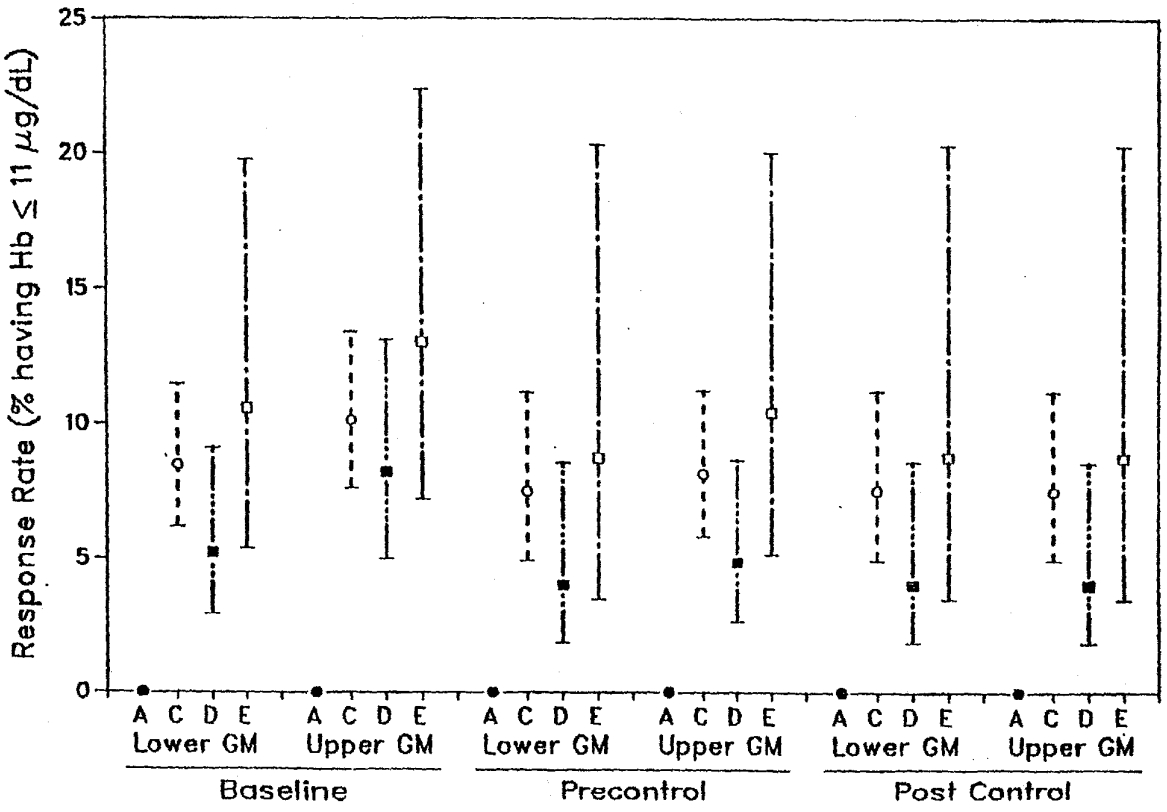

FIGURE 5.3 Point Source 5, with Hb ≤ 11 µg/dL

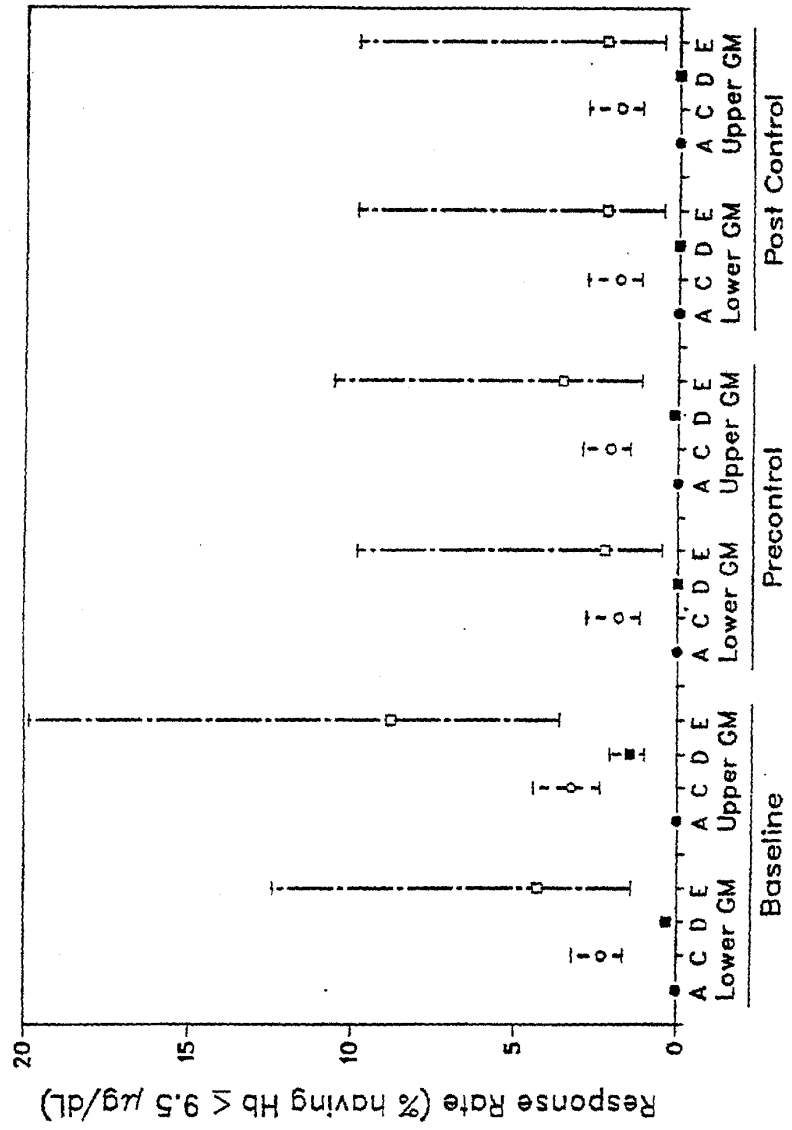

FIGURE 5.4 Point Source 5, with IIb ≤ 9.5 µg/dL

TEH 0413883

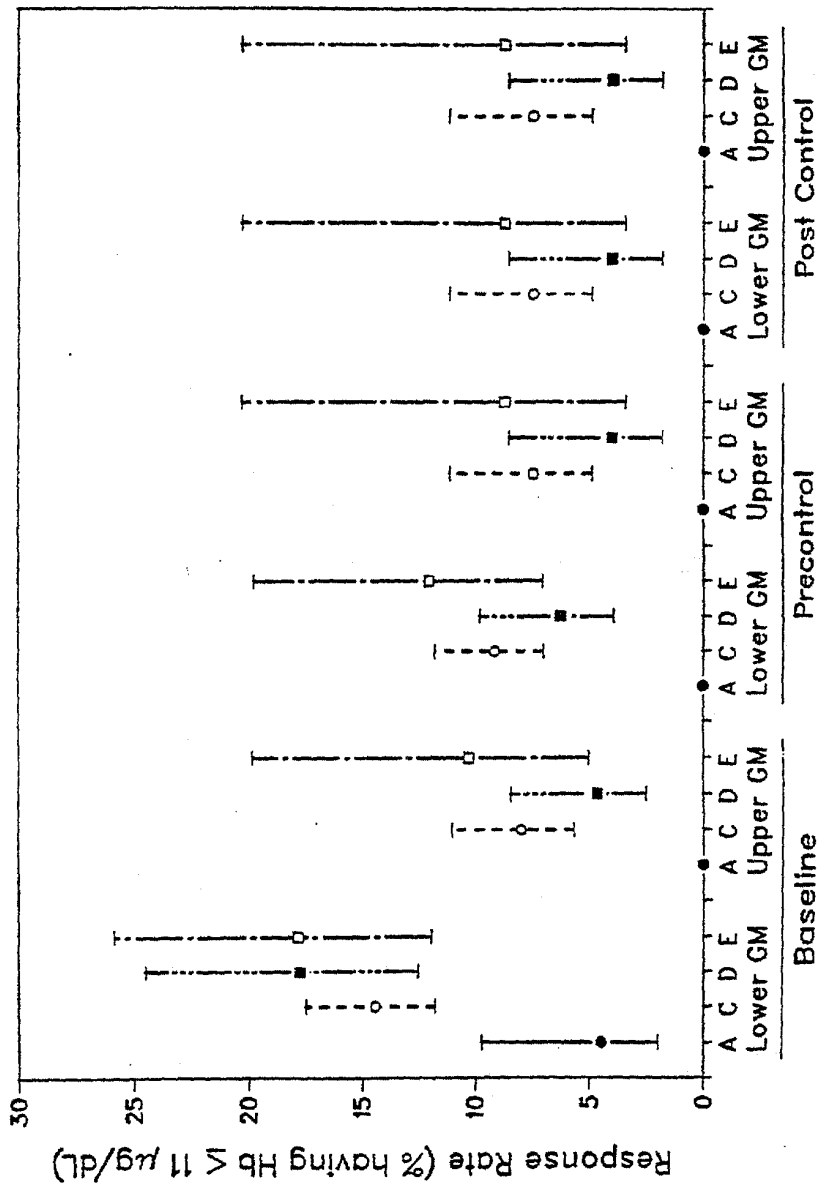

FIGURE 5.5 Point Source 2, with Hb < 11 µg/dL

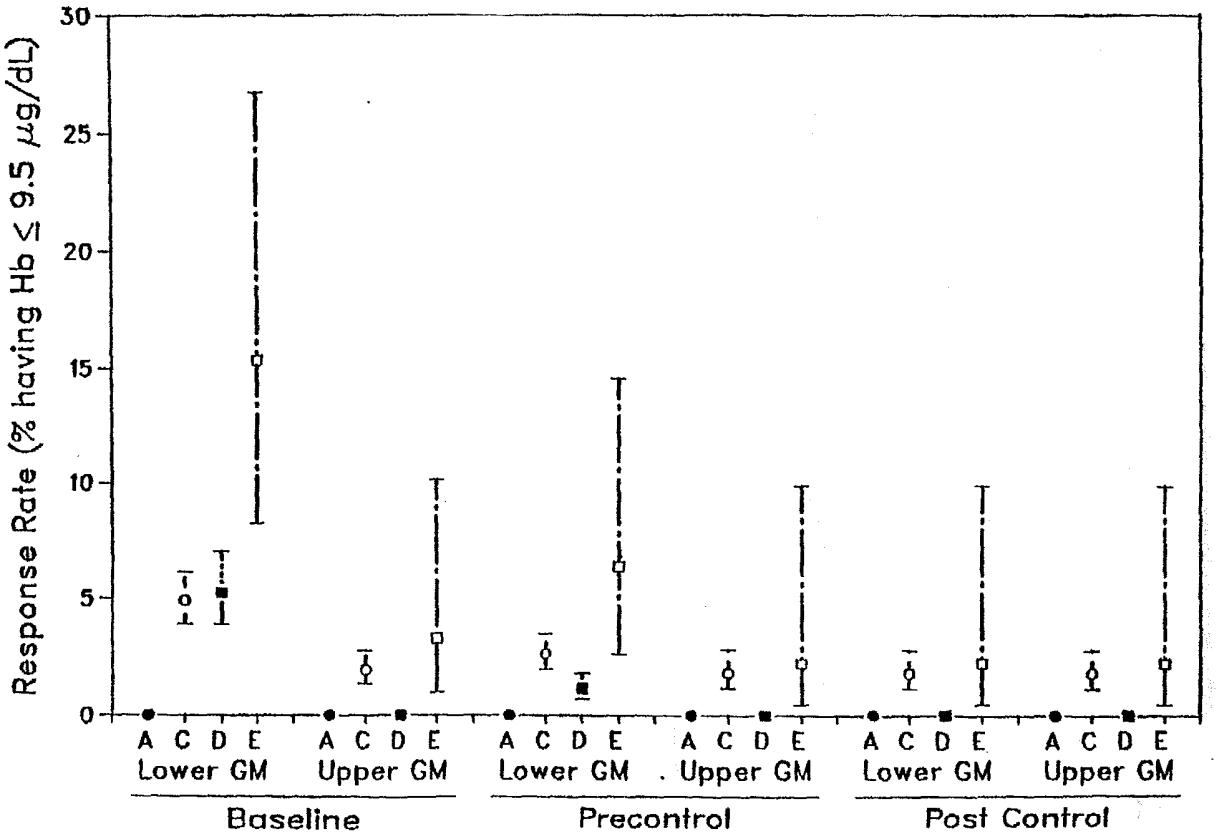

FIGURE 5.6 Point Source 2, with Hb ≤ 9.5 µg/dL

Figure 5.3 (for the 11-g/dL hemoglobin level) indicates that the Expert A risk estimates* are essentially zero for all cases and that the 90% credible intervals for Experts D and E risk estimates overlap to some degree. The differences between the risk estimates based on the lower and the upper GM PbB values are response rates of about 2% for Experts C-E in the baseline scenarios and 1% for Experts C and E and 2% for Expert D for in the precontrol scenarios. Upper and lower GM PbB values are equal in all postcontrol scenarios.

In the precontrol scenario, the median response rates for Expert C risk estimates are about twice those for Expert D and about 15% less than those for Expert E. The Expert D and E risk estimates demonstrate the largest amounts of uncertainty, having upper 90% credible interval (CI) values about twice the median values, and lower 90% CI values about half the median values. For Expert C, these differences are about 25-75% and 50-75%, respectively. This difference is evident from the wider 90% credible intervals indicated in Fig. 5.3 for the Expert D and E risk estimates.

The highest risk estimates are obtained for the baseline, upper-bound scenario, for which the age 0-3 GM blood lead is 13.9 μg/dL; the lowest risk estimates are obtained for the precontrol and postcontrol (lower GM scenarios) for which the age 0-3 GM blood lead is 1.4 μg/dL. The ratios of the highest and lowest median risk estimates are about 1.3, 1.9, and 1.5 for Experts C, D, and E, respectively. Thus, an order of magnitude reduction in blood lead results in at most a 50% reduction in the median value of estimated risk. Similarly, the largest upper bounds and the smallest lower bounds of the 90% credible intervals on risk do not change much in going from the highest PbB scenario to the lowest PbB scenario. Thus, the range of estimated risks that are plausible is not very sensitive to the mean PbB levels associated with the scenarios considered in this study.

*When we use the phrase "expert A risk estimates," we mean a longer statement, namely, "risk estimates based on the judgments of Expert A."

In nearly all of the scenarios for Source 5, there is virtually no chance that any children will have hemoglobin levels less than or equal to 9.5 g/dL, based on the judgments of both Experts A and D. The Expert C and E risk estimates are about half those for the 11 g/dL hemoglobin level. The Expert C risk estimates again show the smallest degree of uncertainty. As for the 11 g/dL hemoglobin level, estimated risk varies little as GM blood lead is changed. The conclusion for Source 5 is that the probability of adverse hemoglobin levels occurring decreases as air lead (and hence blood lead) decreases, but the percent reduction in risk is considerably less than the percent reduction in exposure.

For Source 2, modeled mean PbB levels are about double those for Source 5 for the baseline and precontrol scenarios, and about the same for the postcontrol scenarios. Subsequently, the risks (response rates) are higher -- about 10-25% higher for the 11 g/dL level and about 10-100% higher for the 9.5 g/dL level. In addition, nonzero risk estimates result for Expert D for all scenarios except the postcontrol 9.5-g/dL hemoglobin level scenario. All of these results are consistent with the observation for Source 5 that a plot of risk vs. GM blood lead would be relatively flat, especially at the 11 g/dL hemoglobin level.

The conclusions for Source 2 are that: estimated risks are higher than those for Source 5 in the baseline and precontrol scenarios (because of higher estimated exposure to lead), and the same for the postcontrol scenarios. For Source 5, the probability of adverse hemoglobin levels occurring decreases as air lead decreases, but the percent reduction in risk is considerably less than the percent reduction in exposure. For the other point sources, the GM PbB values are about 5-10% less than those for Source 2.

5.4 RISK RESULTS FOR IQ

Judgments regarding IQ effects were obtained in a manner that allowed consideration from two perspectives that roughly correspond to dose-effect and dose-

response functions of two types of health endpoints. The first of these is expected IQ decrement. The second is more complicated, involving the lead-induced increase in fraction of children having IQs below a specified critical level denoted by IQ*. We show results in this section for IQ* values of 85 and 70, which correspond to one and two standard deviations below 100, although other values of IQ* could have been used. Further details on calculating these quantities are provided in App. C. We discuss the risk results for these two health endpoints in the next two sections.

Both health effects require combining the judgments about the dose-response functions with the exposure estimates from Sec. 5.1 to produce probabilistic risk estimates. Results are illustrated here for Sources 2 and 5, for the lower SES group (the most sensitive population). Risk estimates calculated for Source 5 are smaller than those for any other point source because the modeled blood-lead levels are lower for this point source, and because the percentage of households having incomes below poverty level is the smallest. Modeled blood-lead levels and the percentage of households having incomes below the poverty level are highest around Source 2. For the remaining sources, modeled PbB levels are about 5-10% lower than those for Source 2.

For the IQ health effects, the blood lead of interest is that at the 36th month. IQ, as described in Sec. 4, is measured on the seventh birthday.

5.4.1 Expected IQ Decrement

Expected IQ decrement is calculated by combining the judgments about IQ decrements (relative to a control group of children sheltered from lead) at each blood-lead level with a blood-lead distribution. Results are shown in Figs. 5.7 (for Source 5) and 5.8 (for Source 2). Results for the other sources are slightly less than those for Source 2. In these figures, median values of the expected IQ distributions and 90% credible intervals around those values are indicated for baseline, precontrol, and

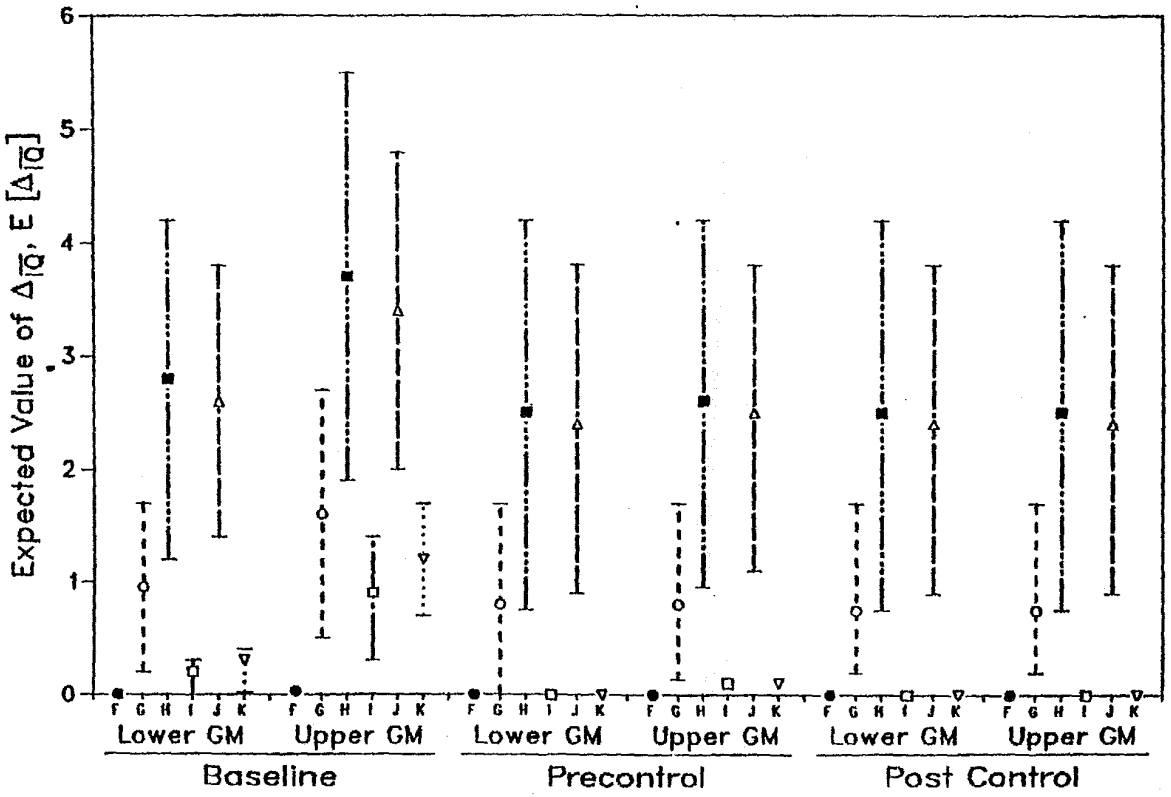

FIGURE 5.7 Expected IQ Decrement for Lower SES Children Living around Source 5

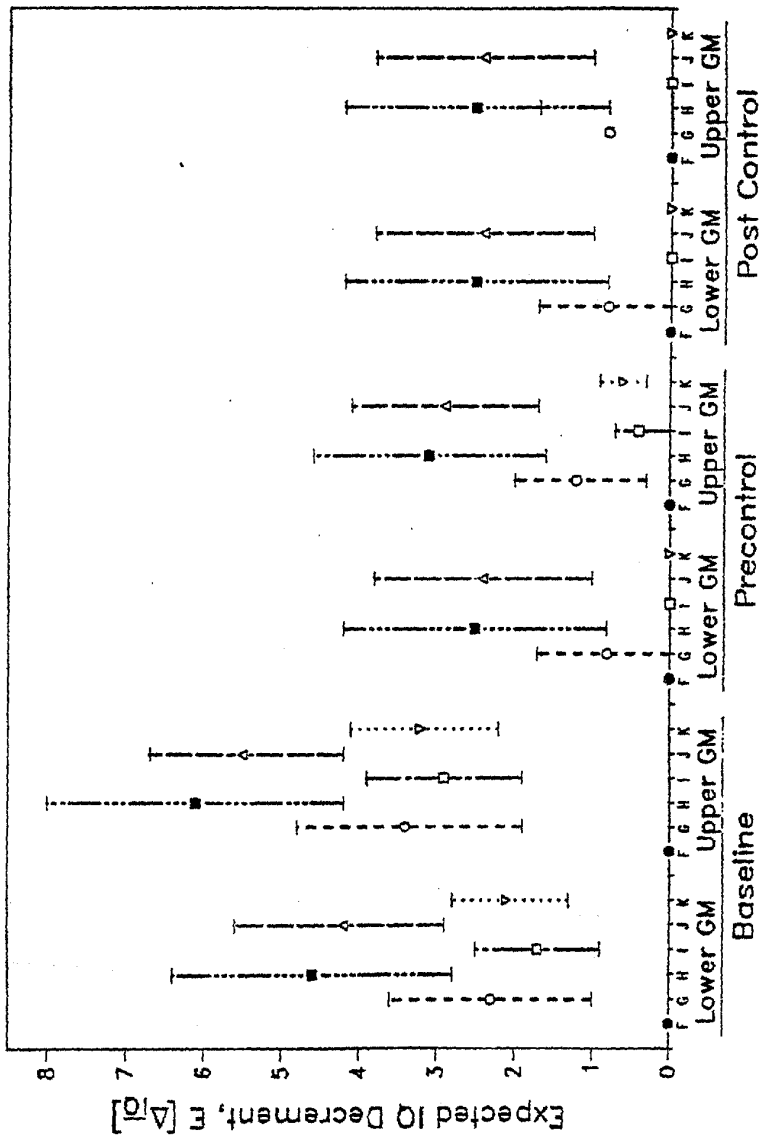

FIGURE 5.8 Expected IQ Decrement for Lower SES Children Living around Source 2

postcontrol scenarios (upper- and lower-bound estimates; postcontrol scenario results are essentially equivalent for all NAAQS levels analyzed), based on the functions fit to the judgments of Experts F through K. The median is indicated by a unique symbol for each expert, and the 90% credible interval is indicated by bars extending both above and below the median value. The bars are of various textures to help to visually discriminate results for the individual experts.

For Source 5, the risk estimates exhibit the following patterns. Expert G, H, and J risk estimates are quite similar to one another, as are Expert F, I, and K risk estimates. The 90% credible intervals for Experts G, H, and J overlap. The distributions for H and J are nearly identical, and their median estimates are about double that for G. The upper ends of the 90% credible intervals for H and J are around four IQ points. The upper GM, baseline scenario risk estimates (for all but Expert F) are about one IQ point higher than the other estimates. (These median estimates of IQ decrement are about 2.7 for H and J, 1 for G, and about 0-0.3 for F, I, and K.) Estimates of IQ decrement for I and K are essentially zero for all other scenarios.

For Source 2, the estimates of IQ decrement are about two points higher in the baseline scenarios for all experts except F. (Estimates of IQ decrement for F are always equal to zero.) This is attributable to the higher blood-lead levels estimated around Source 2. Estimates for the precontrol and postcontrol periods are about equal to those for Source 5. As for EP and hemoglobin, the general pattern for estimates of IQ decrement is that they do not differ substantially (within an expert) in the precontrol and postcontrol scenarios.

5.4.2 Increased Probability of Lead-Induced IQ Levels being $\leq$ IQ*

To obtain functions analogous to dose-response functions, uncertainty in mean IQ of children sheltered from lead, within-group IQ standard deviation, and mean IQ decrement given different blood-lead levels were combined to calculate the increased

probability of having IQs less than or equal to IQ*, as a function of PbB for IQ* values of 70 and 85. These results were then combined with the geometric mean blood leads summarized in Table 5.3 to calculate overall risk estimates.

Results are shown in Figs. 5.9-5.12. The first two figures are for Source 5, and the next two are for Source 2. Figure 5.9 (for IQ* = 70) shows that the increased probability of having IQ levels less than or equal to IQ* for these modeled point sources are very small (median probability values around 0.01 to 0.02 for Experts G, H, and J) or essentially zero (for Experts F and K). There are no estimates for Expert I because that individual did not provide the needed judgments about mean IQ levels for children sheltered from lead exposure or about IQ standard deviations.

The largest risk estimates are calculated in the upper GM, baseline scenario (median estimates range from 0.014-0.024 for the increased probability of having IQs $\leq$ 70, as one would expect. For this scenario, the risk estimates based on the judgments of all experts but F are comparable. Estimates for K of this magnitude result because Expert K judges that mean IQ levels for lower SES children sheltered from lead will be quite low (around 85). For the scenarios with lower blood-lead levels, Expert K risk estimates become much smaller relative to the others because Expert K judges that there is a threshold for lead-induced IQ effects around 5-15 μg/dL.

For IQ* = 85, the results in Fig. 5.10 indicate a five-fold higher degree of risk than just discussed. In the precontrol and postcontrol scenarios, credible intervals for Expert G, H, and J risk estimates overlap (median values of 0.014-0.061, while median estimates for F and K are zero and 0.004, respectively. Again, the distributions for H and J are quite similar, having upper 90% credible interval values around 0.1. Precontrol and postcontrol risk estimates for increased probability are about equal, and about 0.01 lower in probability than those for the lower GM, baseline scenario. In the baseline scenario, a 60% decrease in GM (from the upper bound to the lower bound) is

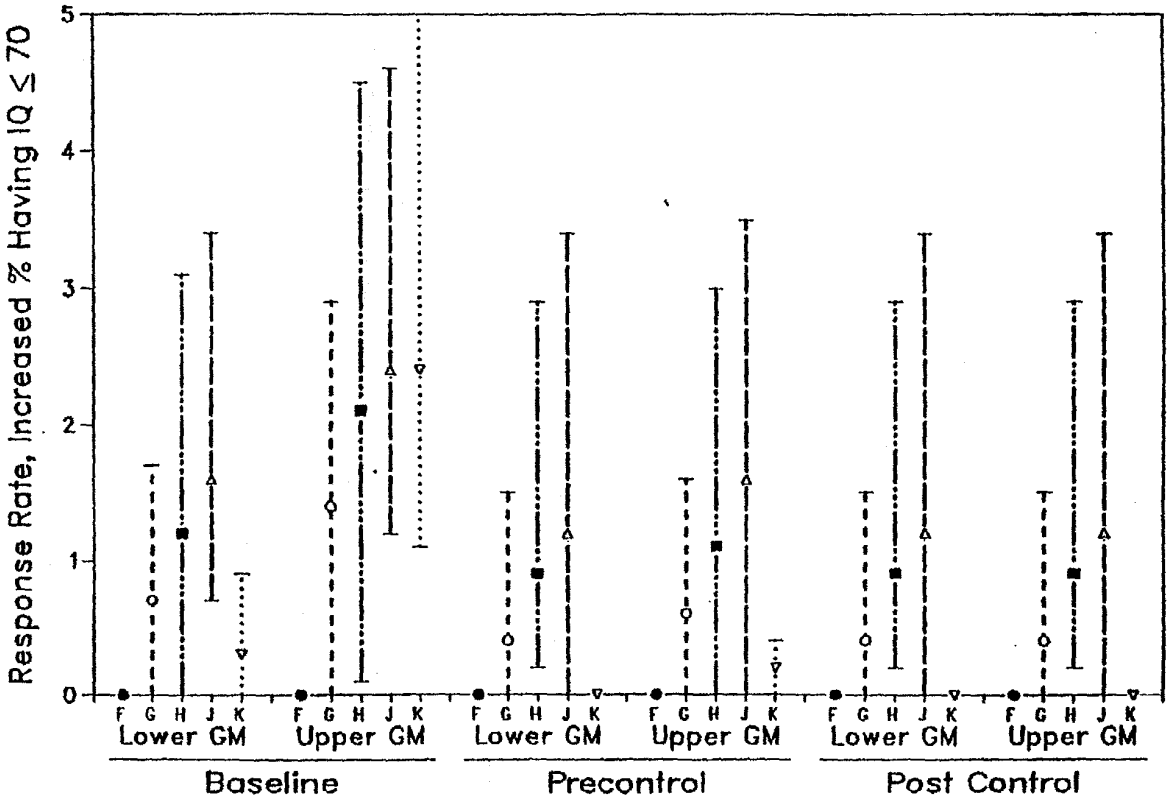

FIGURE 5.9 Increased Probability of Children Living around Source 5 Having IQ Levels ≤ 70

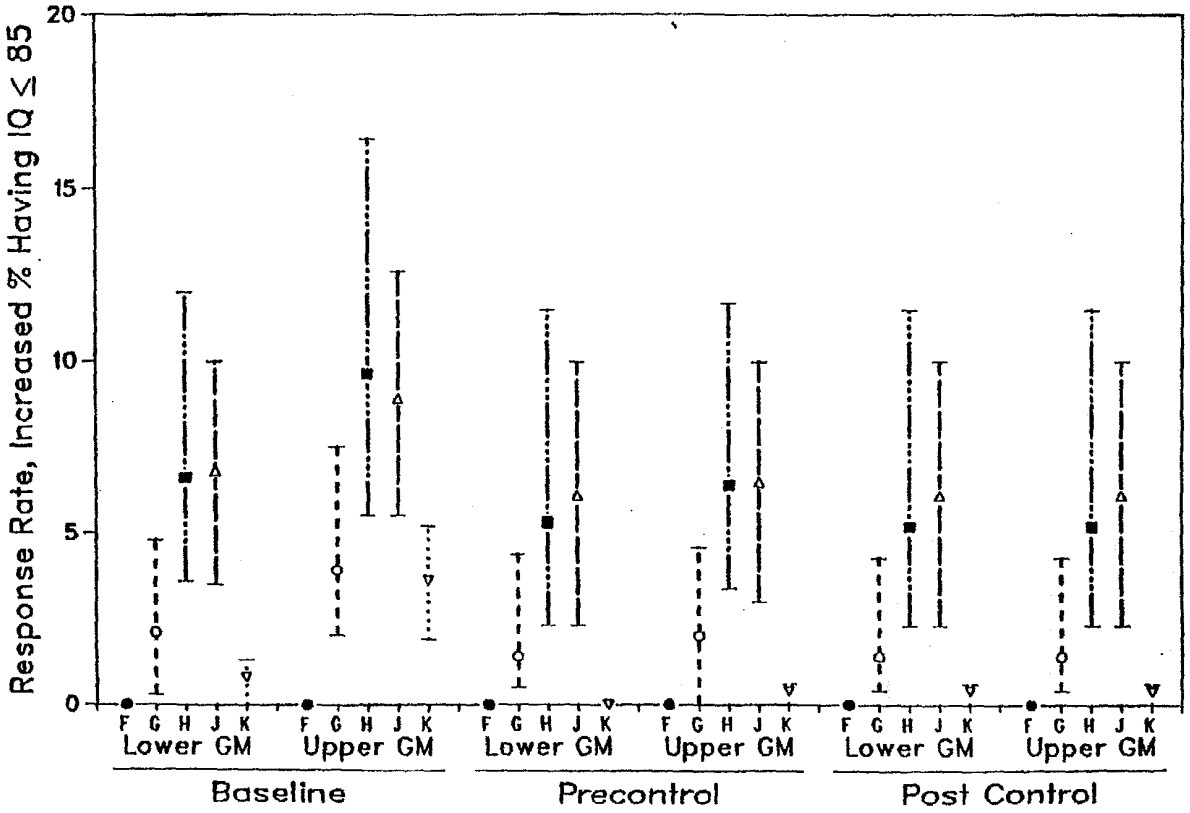

FIGURE 5.10 Increased Probability of Children Living around Source 5 Having IQ Levels ≤ 85

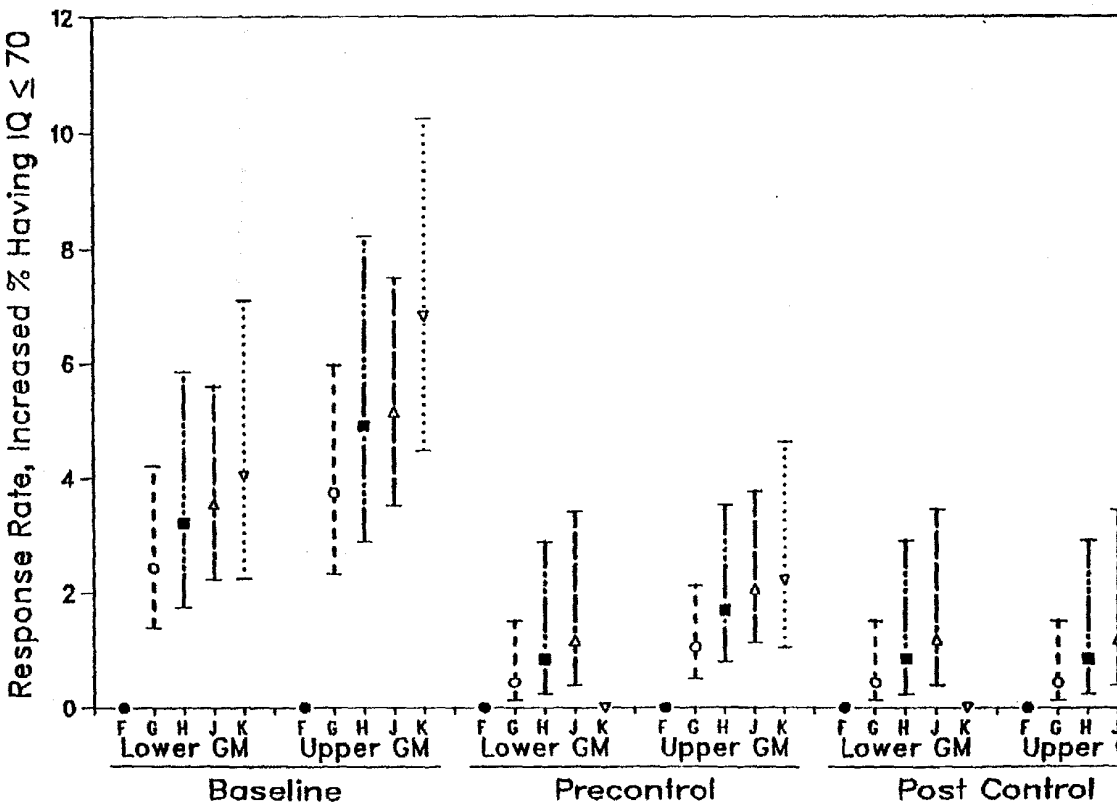

FIGURE 5.11 Increased Probability of Children Living around Source 2 Having IQ Levels ≤ 70

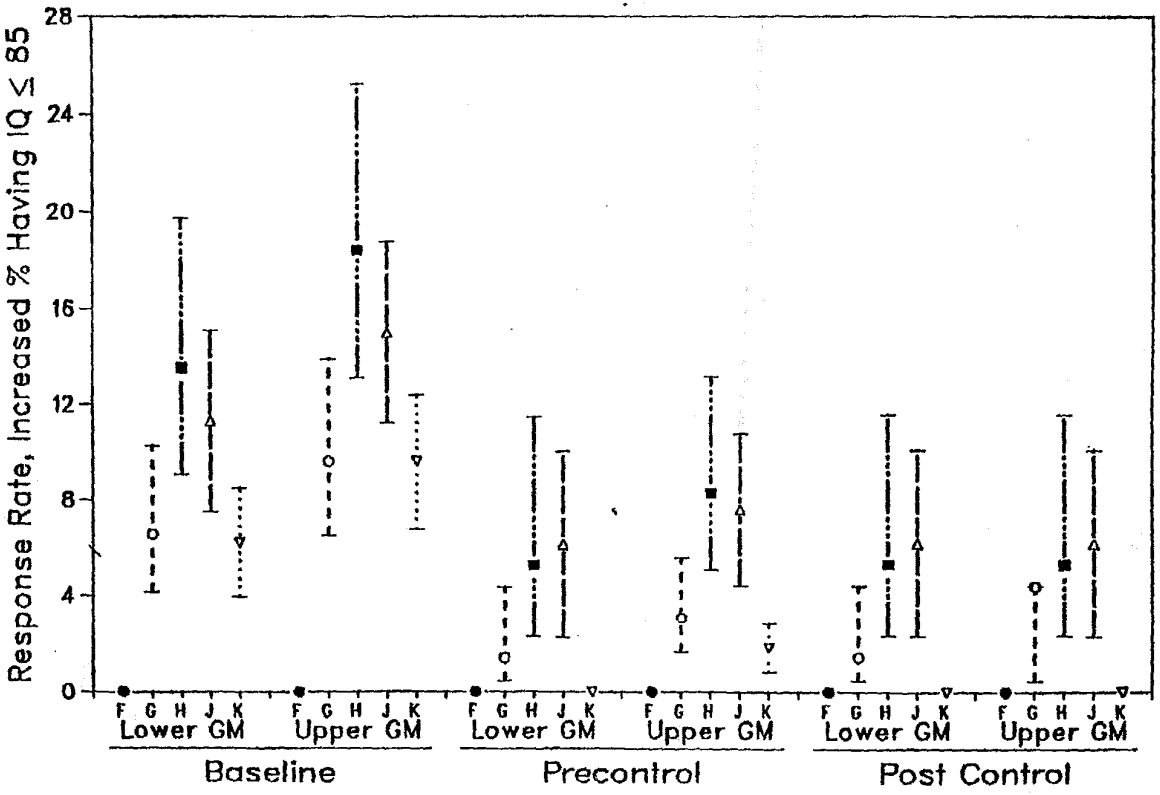

FIGURE 5.12 Increased Probability of Children Living around Source 2 Having IQ Levels ≤ 85

accompanied by a 25-50% decrease in the risk estimates for Experts G, H, and J, and an 80% decrease for Expert K. Further decreases in blood-lead levels do not substantially lower the risk estimates for any of the experts. Thus, the relative decrease in the risk measure vs. a decrease in exposure depends on the level of exposure.

Results for Source 2 are shown in Figs. 5.11 (for IQ* = 70) and 5.12 (for IQ* = 85). For IQ* = 70, risk estimates based on the judgments of Experts G, H, J, and K are similar for the baseline scenarios. The 90% credible intervals overlap in the lower and upper GM cases. In fact, they are nearly overlapping across GM values. Median values of increased probability range from 0.024 to 0.04 for the lower GM values, and from 0.037 to 0.068 for the upper GM values. For the upper GM, precontrol scenario, median risk estimates range from 0.01 to 0.022 among these experts. Expert K risk estimates are zero for the remaining scenarios. The median values for the other experts are about 0.004-0.01, and upper 90% confidence limit values are 0.015-0.034. Probability distributions determined for Experts H and J in each scenario are quite similar.

The risk estimates calculated for IQ* = 85 are generally about three to four times larger than those for IQ* = 70 for Experts G, H, and J, and about 50% higher for K. As for Source 5, there is agreement among the experts, as indicated by the overlapping 90% credible intervals. For a specific expert, the credible intervals for the lower and upper GM values overlap for the baseline scenario. Expert F risk estimates again are small (median values of 0.003 and 0.007 for the lower and upper GM cases). In the postcontrol scenarios, the Expert G, H, and J risk estimates have median values around 0.014-0.062 and upper 90% credible interval values around 0.04-0.1. The probability distributions for Expert H and J risk estimates again are quite similar, and Expert G risk estimates are about one-third those for H and J. As for the other health end points, the general pattern for estimates of the increased probability of having children with IQs less than or equal to IQ* is that those for the precontrol and postcontrol periods are about one-third to one-fourth those for the baseline period, and that the 90% credible intervals for the lower and upper GM blood-lead cases overlap.

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue, Argonne, Illinois 60439

DRAFT — DO NOT CITE

ESTIMATING THE RISKS OF LEAD-INDUCED HEALTH EFFECTS

by

Thomas S. Wallsten* and Ronald G. Whitfield

Energy and Environmental Systems Division
Decision and System Sciences

January 1986

work sponsored by

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air Quality Planning and Standards

*L.L. Thurstone Psychometric Laboratory, University of North Carolina, Chapel Hill

CONTENTS

CONTENTS (Cont'd)

TABLES

TABLES (Cont'd)

FIGURES

FIGURES (Cont'd)

ESTIMATING THE RISKS OF LEAD-INDUCED HEALTH EFFECTS

by

Thomas S. Wallsten and Ronald G. Whitfield

ABSTRACT

to be completed.

1 INTRODUCTION

The Clean Air Act charges the U.S. Environmental Protection Agency (EPA) with setting and reviewing both primary and secondary National Ambient Air Quality Standards (NAAQS) for selected pollutants. Each primary standard must be set at a level sufficient to protect public health with an adequate margin of safety. This report presents the results of a risk assessment performed to assist in the review of the primary NAAQS for lead.

For each review, the scientific basis for revising the primary lead NAAQS is presented in an updated document entitled *Air Quality Criteria for Lead*, hereafter referred to as the criteria document (CD). It summarizes and evaluates available scientific evidence about the adverse health effects of lead. An EPA staff paper is then written, which considers the information in the CD in specifying the elements EPA staff members believe to be critical in revising the lead NAAQS. The staff paper is intended to bridge the gap between the science presented in the CD and the information needed by the EPA Administrator in setting a new ambient standard for lead. Particular attention is paid to those areas where judgments must be based on careful interpretation of imperfect evidence. Indeed, to provide the required adequate margin of safety, uncertainty must be taken into consideration at each step of the process.

Scientific uncertainty was handled in a highly simplified manner when the 1978 air-lead standard was set. Instead of quantifying uncertainty, deterministic assumptions were made that were considered conservative given qualitative assessments of the uncertainty associated with available evidence. Calculations based on these assumptions resulted in a primary lead NAAQS of 1.5 μg Pb/m^3 of air.

Soon thereafter EPA's Office of Air Quality Planning and Standards (OAQPS) began exploring methods of risk assessment that incorporated uncertainty in the standard-setting process in a formal, defensible, and open manner. This report presents the outcome of formal risk assessments regarding three adverse health effects of lead to aid the EPA Administrator in determining an adequate margin of safety for the current review of the primary lead NAAQS. This work was funded by EPA under an Interagency Agreement between EPA and the U.S. Department of Energy.

Of the many potential adverse health effects associated with exposure to lead, EPA's OAQPS selected three for inclusion in this risk assessment: hemoglobin decrement, elevated erythrocyte protoporphyrin (EP) levels, and intelligence quotient (IQ) decrement. It was decided to limit the population at risk for this risk assessment to all U.S. children from birth through their seventh birthdays. Pregnant women (fetus), who are particularly sensitive to lead, are not explicitly included in this analysis.

1.1 OVERVIEW

To the extent possible, details of the methods and results of the risk assessments are presented in appendixes rather than in the body of this report. Nevertheless, readers will obtain a good understanding of the risk assessments from the main text. Section 1.2 presents the motivation for the particular form of the risk assessments that were carried out. Section 1.3 discusses the role of judgmental probability encoding in the risk assessment. Section 1.4 summarizes the overall risk assessment strategy, while Sec. 1.5 discusses dose-response uncertainty.

Section 2 treats the dose-response uncertainty for lead-induced elevated EP levels. Because relatively complete data are available, judgmental probability encoding was unnecessary in considering lead-induced changes in EP level. Sections 3 and 4 focus on dose-response uncertainty for lead-induced hemoglobin and IQ decrements, respectively. Judgmental probability encoding was required in these cases; the encoding procedures are described in the respective sections prior to the summaries of the results. Section 5 discusses possible blood-lead* (PbB) distributions and their levels of uncertainty under alternative NAAQS. The uncertainty for each health effect is then combined with that for each alternative PbB distribution to yield derived risk estimates. The report concludes with four appendixes.

1.2 MOTIVATION

The purpose of a risk assessment for lead health effects is to estimate probability distributions over fractions of the population at risk that would suffer specific, well-defined health effects under alternative primary NAAQS for lead. The probabilities must be estimated in a formal, defensible manner that is open to public scrutiny. They will generally reflect two sources of uncertainty, one deriving from the characteristics of the data and the other from a lack of knowledge. The former type of uncertainty reflects measurement and sampling error, and probabilities are generally calculated by means of standard statistical procedures. Although this type of uncertainty can be altered by means of experimental manipulations, it cannot be eliminated completely in any finite study.

The latter type of uncertainty reflects the paucity, incompleteness, and indirect nature of much of the available data. For example, it is frequently necessary to draw inferences about a health effect in a specific population on the basis of epidemiological

*Blood lead is generally used as the index of lead exposure in health studies and is used in this way in this assessment.

or clinical data on different populations under varying or different exposure conditions, or from laboratory data on other species or on *in vitro* preparations. This type of uncertainty is judgmental, and no statistical techniques exist for quantifying it. It can, however, be quantified using appropriate procedures to encode subjective probabilities. Different experts will assess the uncertainty differently, depending on their faith in the implicit or explicit theories required for extrapolation from the data to the effect in question and on their perception of the distance over which the extrapolation must be made. Therefore, it is necessary that the procedures accommodate and represent a possible divergence of judgments. The level of this type of uncertainty, as well as the degree of divergence among experts, can be reduced, and conceivably even eliminated, by increasing the amount of available knowledge.

Conducting a risk assessment in which these two types of uncertainty are distinguished provides one means for responding to a challenge issued by William Ruckleshaus in a speech given at Princeton University when he was EPA Administrator. He stated:

> If I am going to propose controls that may have serious economic and social effects, I need to have some idea of how much confidence should be placed in the estimates of risk that prompted those controls (Ruckleshaus, 1984, p. 158).

In the same talk, he went on to propose some principles for reasonable discussions about risk:

> First, we must insist on risk calculations being expressed as distributions of estimates and not as magic numbers that can be manipulated without regard to what they really mean. We must try to display more realistic estimates of risk to show a range of probabilities. To help do this, we need new tools for quantifying and ordering sources of uncertainty, and for putting them in perspective.

> Second, we must expose to public scrutiny the assumptions that underlie our analysis in management of risk (Ruckleshaus, 1984, p. 158).

If there were complete information relevant to a given health effect in a population, then the statistical uncertainty could be represented by a single probability distribution for each alternative NAAQS under consideration. A hypothetical example of a risk assessment output of this type is given in Table 1.1. For example, in this hypothetical situation, the probability is 0.01 that fewer than 0.5% of the population will suffer the health effect, the probability is 0.05 that fewer than 1% of the population will suffer it, and so forth. (Equivalently, the probability is 0.99 that more than 99.5% of the population will not suffer the health effect, 0.95 that more than 99% will not suffer it, etc.) Alternative NAAQS 4 provides the greatest degree of protection.

Presumably, output of this sort would be helpful to the EPA Administrator in selecting a standard that in his or her judgment would protect the public health with an adequate margin of safety. For example, if he or she determined that the intent of the Clean Air Act is met by protecting 99.5% of the population at risk with probability 0.99, then alternative 1 would be selected. If the intent is met by protecting 99.5% of the population with probability 0.95, then alternative 2 is the appropriate one. Of course, this example is highly simplified because multiple health effects will generally be of interest, and populations can be defined in various ways. However, it illustrates the role that formal risk analysis can play in the standard-setting process.

Generally, the information relevant to a particular health effect associated with an environmental agent will be indirect and incomplete, and experts will differ with regard to the associated judgmental uncertainty. The extent of agreement across experts is a measure of how firm the probability estimates are, and is useful information in setting the standard.

One way of representing the extent of agreement is to propagate the encoded probabilities of each expert through the entire analysis, which results in a family of

TABLE 1.1 Illustrative Risk Output for a Specified Health Effect Given Full Information

Response Rate R_0	Probability (Response Rate $\leq R_0$)			
	NAAQS Alternative			
	1	2	3	4
0.5%	0.01	0.06	0.19	0.38
1.0%	0.05	0.11	0.23	0.60
1.5%	0.41	0.53	0.64	0.75
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

distributions under each alternative NAAQS. A hypothetical example in this case is shown in Table 1.2. Under alternative 1, for example, there is a 0.01 to 0.02 probability that fewer than 0.5% of the population will suffer the health effect. The closer together are the distributions in a family, the more firm are the probability estimates.

In setting the NAAQS, the degree of firmness in the estimates can be taken into account in any of a number of ways. For example, one criterion might be that 99.5% of the population be protected with a probability whose lower bound is no less than 0.01 or with a probability estimated to be 0.02 in at least, for example, 75% of the risk distributions.

On a less formal level, the probability encoding of each expert can be accompanied by a discussion of the relevant data. Summaries of these discussions can be an important supplement to the probability distributions and provide guidance on their use.

The risk assessments in this report yield outputs similar to those in Table 1.2 plus associated discussions. However, while endeavoring to make the outputs fully interpretable, we do not presume to suggest the criteria that should govern their use in the standard-setting process.

1.3 JUDGMENTAL PROBABILITY ENCODING

We wish to emphasize that subjective probability encoding is unnecessary whenever adequate, direct data are available. However, more often than not, such data are not available, and the only alternatives to encoding judgmental probabilities will be to treat the added uncertainty qualitatively, which currently cannot be done in a rigorous manner, or to ignore it altogether, which is indefensible.

For encoded judgmental probabilities to be useful, they must meet two criteria. The first is subtle, and itself judgmental, but is nevertheless very important; namely, judgments must be obtained from respected experts who span the range of respected

TABLE 1.2 Illustrative Risk Output for a Specified Health Effect Given Incomplete Information

Response Rate R_0	Probability (Response Rate $\leq R_0$)			
	NAAQS Alternative			
	1	2	3	4
0.5%	0.01-0.02	0.02-0.06	0.11-0.19	0.26-0.38
1.0%	0.01-0.05	0.04-0.11	0.15-0.23	0.49-0.60
1.5%	0.29-0.41	0.40-0.53	0.53-0.64	0.68-0.75
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

opinion. Although it would appear to be difficult to establish what the range of opinion is and who the appropriate experts are, in practice the relevant issues tend to be debated frequently and publicly enough that such decisions are possible.

For the present risk assessments, EPA's OAQPS selected the experts whose judgments were to be encoded. Although their names are included with the assessment, their individual judgments and discussions are identified by code letter only. Thus, users of the risk assessments can decide whether the full range of opinion is represented, and the individual experts can feel free to give their best responses without worrying that they may involve themselves in endless arguments or discussions with others who may disagree.

The second criterion is that an individual's encoded probability judgments must be stable over time (barring new information), coherent in a well-defined way, and accurately represent his or her uncertainty. Research relating to these issues has been reviewed by Wallsten and Budescu (1983), and an experiment specifically addressing them in the present context was performed by Wallsten et al. (1983). The techniques used in the present work were based on this foundation but also relied on pilot work described in Whitfield and Wallsten (1984).

1.4 RISK ASSESSMENT STRATEGY

The strategy employed in this analysis is most easily understood by first assuming that no uncertainty is involved. In this case, one first specifies the dose-response curve for each selected health effect of lead. For a well-defined population, well-defined exposure conditions, and a well-defined physiological or behavioral effect, a dose-response curve plots the percentage of the population exhibiting the effect as a function of dose level. This function is assumed to be unaffected by alternative air-lead scenarios. However, the alternative scenarios do determine the distribution of dose levels in the population. Thus, one can combine the dose-response function for a health

effect with the dose distribution under each alternative scenario to yield the percentage of the population that would suffer the effect under that alternative scenario. In reality, there is uncertainty at each step, and it must be incorporated into the analysis in a manner providing final outputs of the form shown in Table 1.1 or 1.2.

1.5 DOSE-RESPONSE UNCERTAINTY

In addition to the two types of uncertainty discussed in Sec. 1.2, there is ambiguity regarding the level of an effect that should be considered adverse and for which a dose-response function would be of interest. This issue was dealt with in the present work by simply selecting two levels of the relevant variable for each health effect and estimating a dose-response function for each. The levels were selected so as to span the range that arguably delimits normal from adverse consequences.

It must be emphasized that a population dose-response function exists for a given effect under specified conditions. When relatively complete data are available regarding the dose-response function of interest, the population values are still uncertain because of sampling and measurement error. Because this was the case with elevated EP levels, a straightforward application of the appropriate probability distributions to the data yielded the desired results (see Sec. 2). However, when the data regarding the dose-response function of interest are incomplete or indirect, then there is judgmental uncertainty associated with extrapolation (see Sec. 1.3). This was the case with both the hemoglobin and IQ decrements. Because unique factors were associated with each of these instances, the procedures and models employed for quantifying the uncertainty were different.

2 PROBABILISTIC DOSE-RESPONSE RELATIONSHIPS FOR LEAD-INDUCED ELEVATED ERYTHROCYTE PROTOPORPHYRIN LEVELS

Piomelli et al. (1982) examined PbB and EP levels from 2004 New York City children ages 2 to 12 years, who had been screened to account for the effects of iron deficiency and other confounding variables. Empirical dose-response curves were presented for two levels of EP elevation; in each case, the percentage of the sample with EPs above the critical level was plotted as a function of PbB level. Our consultants felt that the population from which the sample was drawn was sufficiently close to the population of interest (children up to their seventh birthdays) that extrapolation was not necessary. The data were used directly.*

Because of the very careful procedures employed by Piomelli et al. (1982), it is necessary to consider only sampling, and not measurement, error. Each of the points on a dose-response function is based on a number of observations (N) , of which X exhibit the health effect and the remainder (N - X) do not. These data provide an estimate $\hat{P}$ of the true population proportion P suffering the health effect at the PbB level in question. Of course, uncertainty is associated with the estimate because of sampling error. In classical probability theory, this uncertainty is handled by using the binomial distribution with parameters $\hat{P}$, X, and N to calculate confidence intervals around $\hat{P}$ as an estimate of P, or more generally to calculate a probability distribution over P. Thus, the output is a probability distribution over the percentage of the population affected at each observed PbB level.

The Bayesian approach is to calculate a posterior probability distribution over P, given the data and a prior distribution over P. When the prior distribution is represented as a beta distribution with parameters a and b, then the posterior distribution is also

*Similar data were collected by Hammond et al. (1985) on a sample from a different population of children. The dose-response functions were very similar to those obtained by Piomelli et al. (1982), but certain nonindependencies in the data due to their longitudinal nature made them unsuitable for the present purpose.

beta, but with parameters a + X and b + (N - X). A conservative strategy is to assume a diffuse prior state of information characterized by a = b = 0 (Winkler, 1972). In this case, the resulting probability distribution over the fraction of the population affected at a given PbB level is very close to that calculated in the classical manner. Technically, however, the interpretation is different, and the Bayesian approach used is the one most suitable for our purposes. Further details are provided below and in App. A.

Figure 2.1 presents the dose-response data for two EP levels (Piomelli et al., 1982), and Table 2.1 presents the sample-size data. The two EP levels (33 μg/dL and 53 μg/dL) were each considered as adverse. Expressions can be determined for the mean dose-response curves that are shown. These expressions, listed in App. A, are functions of PbB level and what Piomelli et al. call "natural frequencies" for the occurrence of EP levels above 33 μg/dL and 53 μg/dL, and involve the normal probability distribution function.

As discussed above, the beta distribution is the proper distribution to use to describe uncertainty about the true population proportion suffering from elevated EP levels. However, for PbB levels above 5 μg/dL and below 50 μg/dL, sample sizes are large enough that the normal distribution can be used to closely approximate beta distributions for both EP levels. Table A.1 in App. A summarizes the probabilistic dose-response functions that were derived.

Figures 2.2 and 2.3 show median dose-response relationships and 90% credible intervals for the 33 μg/dL and 53 μg/dL EP levels, respectively. For all cases in which a normal approximation is adequate, median and mean response rates are identical, and confidence intervals are symmetrical about the median values. The median value is that response rate at which it is equally likely that the actual rate is above or below it. There is a probability of 0.05 that the population response rate is below the credible interval range and a probability of 0.05 that it is above it.

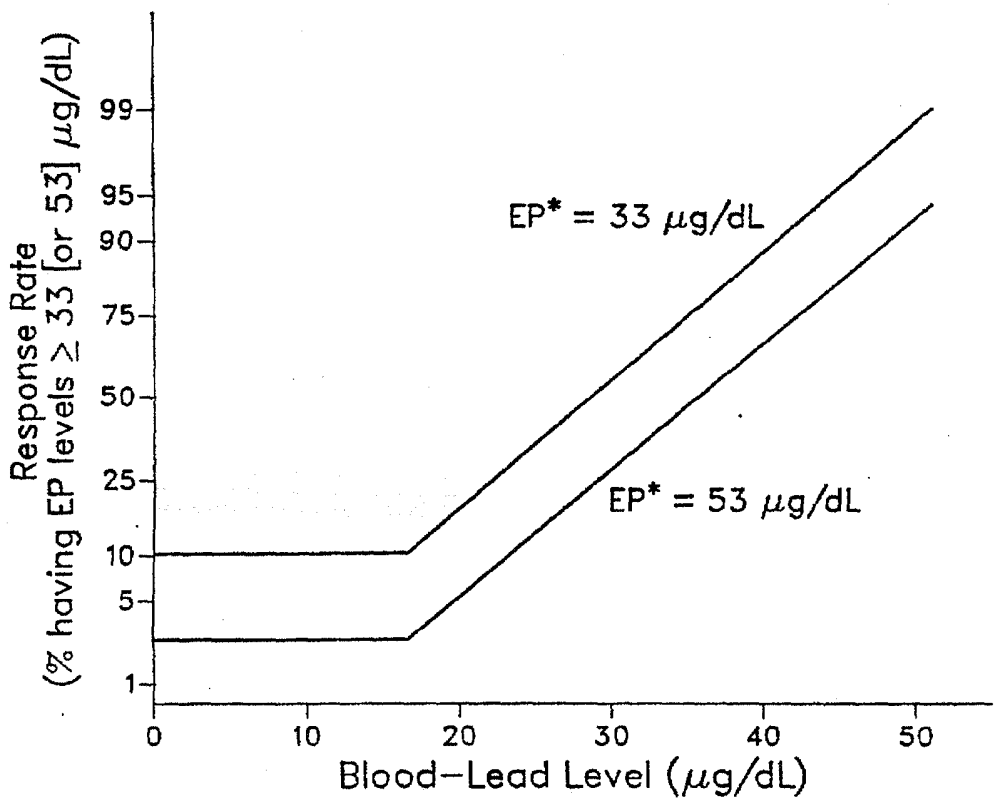

FIGURE 2.1 Mean Response Rate vs. Dose for EP Levels ≥ 33 μg/dL and 53 μg/dL (Source: Adapted from Piomelli et al., 1982)

TABLE 2.1 Sample Sizes Based on Blood-Lead Level for Piomelli's Data

PbB (μg/dL)	No. of Children
0-10	131
10-20	1177
20-30	544
30-40	109
40-98	43

Source: Seaman, personal communication (1985).

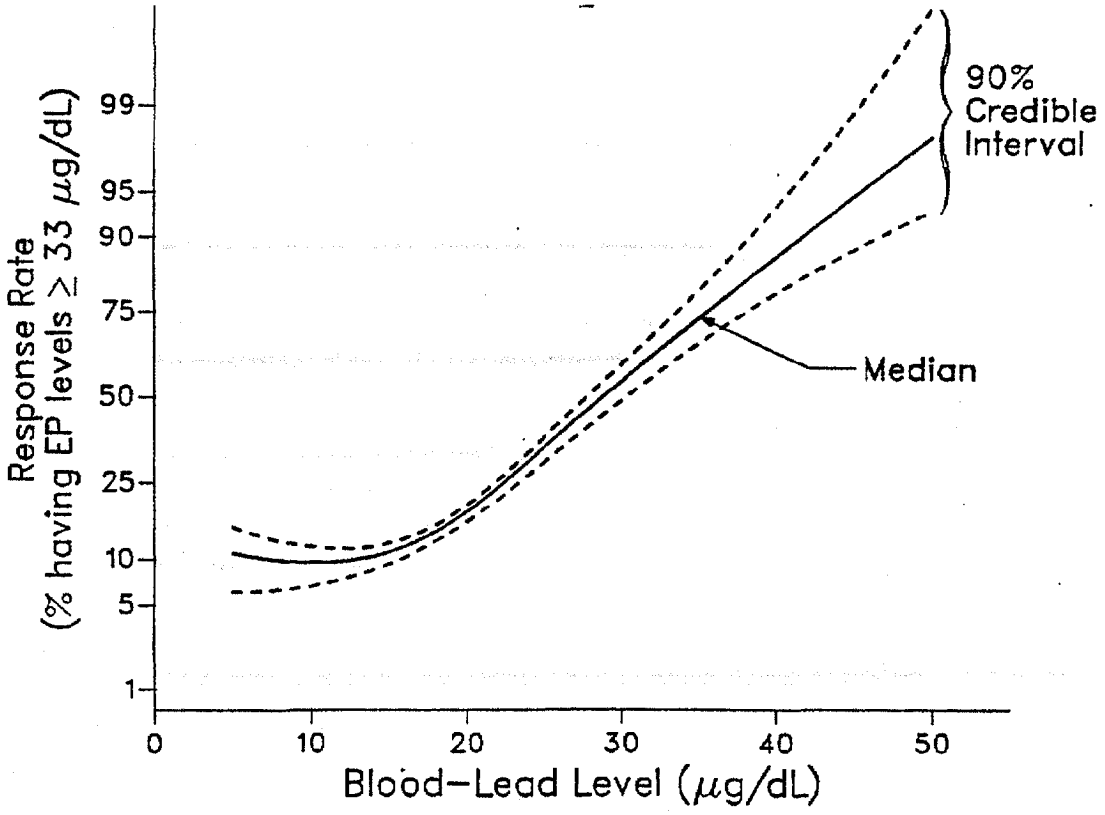

FIGURE 2.2 Median Response Rate and 90% Credible Interval vs. Dose for EP = 33 μg/dL

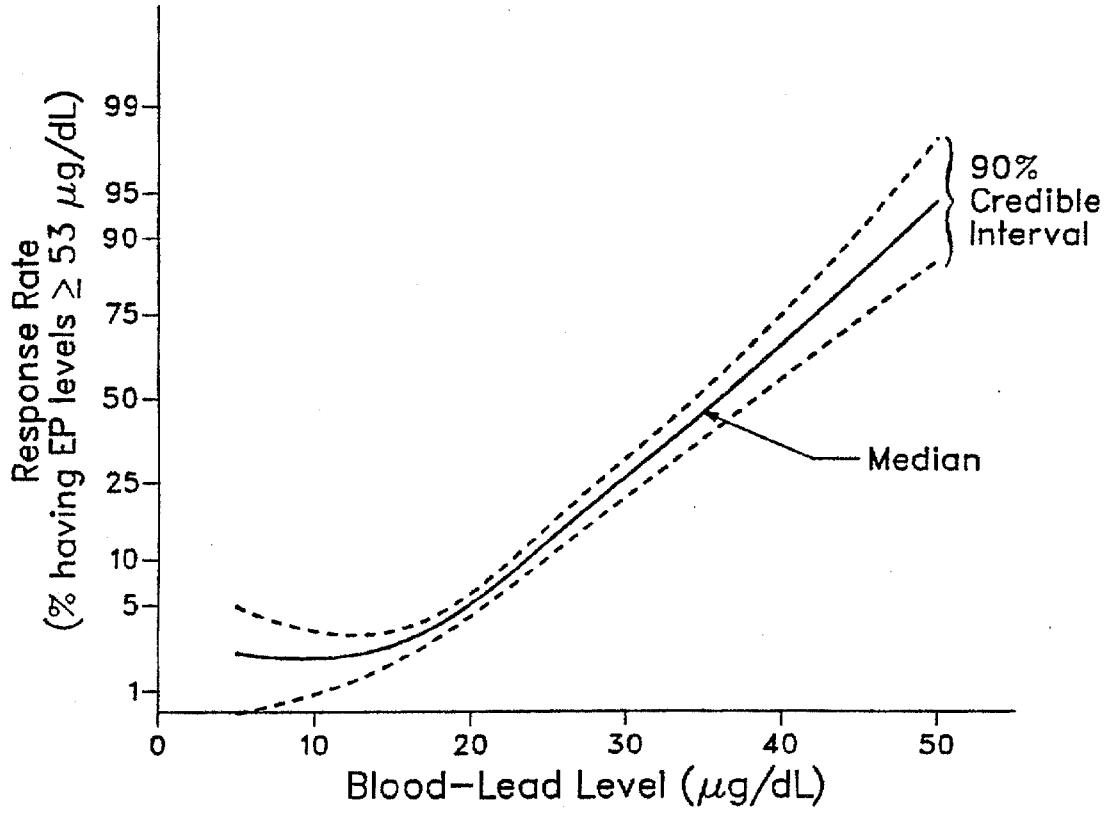

FIGURE 2.3 Median Response Rate and 90% Credible Interval vs. Dose for EP = 53 μg/dL

3 PROBABILISTIC DOSE-RESPONSE FUNCTIONS FOR LEAD-INDUCED HEMOGLOBIN DECREMENTS

Data are very incomplete regarding dose-response functions for lead-induced hemoglobin decrement in the full population of U.S. children prior to the seventh birthday. Therefore, it was necessary to encode experts' subjective probabilities about the population response rates at each of several PbB levels. Encoded probabilities have to meet two criteria. One is that they be internally consistent by specific criteria that will be specified shortly. The other is that they are satisfactory to the experts in the sense that they agree that the encodings properly represent their judgments.

Only a finite number of PbB levels can be presented for encoding; in this study the number was about six. However, it is necessary to interpolate between levels to use the judgments in the risk assessments. Such interpolations can be made by fitting a suitable probability distribution to the encoded values. The distribution must fit the judgments by a reasonable mathematical criterion, but equally importantly, the expert must agree that the function is accurate.

To meet all these requirements and to follow the recommendations developed by Wallsten et al. (1983), we developed a protocol for the probability encodings and then met with each expert on two occasions separated by about one month. Judgments were encoded during the first session, and functions were fit to them prior to the next visit. The judgments and functions were reviewed during the second session, and changes were made that were deemed necessary by the expert. When required, there was additional follow-up through the mail and by phone.

Section 3.1 describes the protocol development, and Sec. 3.2 outlines the protocol itself. The manner in which the encoding sessions were conducted, the probability encoding methods, and the procedures for fitting functions to the encoded judgments are described in Secs. 3.3, 3.4, and 3.5, respectively. Section 3.6 indicates whose judgments were encoded, Sec. 3.7 summarizes the results, and Sec. 3.8 provides some summary discussion.

3.1 PROTOCOL DEVELOPMENT

The consultants listed in Table 3.1 aided us in developing the hemoglobin probability encoding protocol. Their assistance took many forms, including discussing the literature with us, guiding us in reading key studies, and carefully commenting on numerous drafts of the protocol.

After studying relevant portions of the CD and discussing them at length with a subset of our consultants, who for the most part were the authors of those portions, we prepared a first draft of the protocol. This draft was shown to those experts as well as to others who were drawn from the population of people whose judgments might be encoded. We worked with these experts, structuring the problem in the manner we would if we were going to encode their judgments, and revised the protocol accordingly.

In this fashion, the protocol went through a number of drafts. If all the respondents viewed the problem in the same way, this way was specified in the protocol. If they differed or indicated that others might differ from them, we allowed the individual to individualize the scenario, subject only to the constraints that the obtained judgments could map into the rest of the model and be comparable across experts. The protocol was ultimately structured such that the consultant health experts considered it reasonable and indicated that they would feel comfortable providing the desired probabilistic judgments.

The protocol was tested by using it to encode the judgmental probabilities of two members of the EPA staff familiar with the effects of lead. The test was considered successful and is reported in Whitfield and Wallsten (1984).

3.2 PROTOCOL OUTLINE

The full protocol is presented in App. B; only a general outline is provided here. Section 1, the introduction, indicates what we were attempting to do and why. Sections

TABLE 3.1 Consultants for the Hemoglobin Protocol

Consultants	Affiliation
Dr. Julian Chisolm	John F. Kennedy Institute Johns Hopkins University
Mr. Jeff Cohen	OAQPS,[a] U.S. EPA
Dr. Anita Curran	Westchester County Dept. of Health
Mr. Thomas Feagans	OAQPS, U.S. EPA
Dr. Lester Grant	Environmental Criteria Assessment Office, U.S. EPA
Mr. John Haines	OAQPS, U.S. EPA
Dr. Paul Hammond	Dept. of Environmental Health University of Cincinnati
Mr. Robert Kellam	OAQPS, U.S. EPA
Dr. Paul Mushak	Department of Pathology University of North Carolina

[a]Office of Air Quality Planning and Standards.

2-5 specify the parameters of the problem, so that the dose-response functions are well defined. Section 6 raises various points in the literature for discussion. Sections 7 and 9 discuss the process of probability encoding. Section 8 deals peripherally with hemoglobin decrement and raises issues concerning possible adverse health effects associated with elevated EP levels.

More specifically, Sec. 2 defines adverse hemoglobin levels as 9.5 g/dL and 11 g/dL, thereby bracketing the range of interest. Thus, judgments were encoded for two sets of dose-response functions. Section 3 defines the population at risk for encoding as all United States children from conception through their seventh birthday. However, the experts were given the option of considering separately the age groups 0-3 and 4-6 years if they thought for any reason that the two groups would have different dose-response functions. Therefore, experts who exercised this option had their probabilistic judgments encoded for four dose-response relationships.

Section 6 elicits discussion from the expert on relevant theories and data. This step helped ensure that the problem was structured properly, that the expert consciously reviewed his or her knowledge prior to providing probability judgments, and that a possible basis for understanding any discrepant judgments was provided.

Sections 7 and 9, which cover the probability encoding, were derived from our prior work on encoding probability judgments about dose-response functions (Wallsten et al., 1983). Section 7 provides suggestions on how to minimize biases in probability encoding, whereas Sec. 9 discusses the intended probability encoding procedure.

3.3 CONDUCT OF THE SESSIONS

All sessions were conducted during the winter of 1984-1985. Sections 1-8 of the protocol were sent to each expert before the scheduled review. The protocol was then used to guide discussion and interaction during the sessions.

Approximately one to three hours was spent at the beginning of the first session discussing the material in Secs. 2-6 of the protocol. The concepts in Sec. 7 were then discussed and illustrated, after which the probability encoding began. (The experts were not required to read Sec. 9 of the protocol; it simply constituted a record of what we were doing.) The actual encoding took about four to six hours.

The second session took place four to six weeks after the first session and lasted from two to four hours. It included a review of the summaries of the previous session, which had been sent in the mail; additional probability encoding as necessary; and discussion of the material in Sec. 8 of the protocol.

3.4 ENCODING THE JUDGMENTS

Probability encoding began in session 1 after the relevant literature and the factors determining the dose-response function had been thoroughly discussed with the expert. The intention was to encode probability judgments for possible response rates at PbB levels of 5, 15, 25, 35, 45, and 55 μg/dL. Higher PbB levels were employed with one expert.

To determine useful points in the dose-response space for encoding probabilities, each expert was first asked to supply broad "99% confidence interval" ranges for possible response rates at each of the PbB levels. We helped the expert determine these ranges by posing suitable questions. For example, after a range was specified, we asked whether the expert could construct a plausible explanation if some day it were found that the true value fell outside that range.

For most of the experts, probabilities were encoded using a probability wheel. For this purpose, the range of possible response rates for each PbB level was divided into five equal intervals for detailed encoding at six points. For example, if the response-rate range at a particular PbB level was 0.2-0.7, subsequent encodings were recorded at 0.2,

0.3, 0.4, 0.5, 0.6, and 0.7. If during the encoding process it became apparent that other points either inside or outside the range were important, they were incorporated.

Blood-lead levels were presented for encoding in a random order; for a particular PbB level, response rates were also presented in a random sequence. Probabilities were encoded for all response rates at a given PbB level before moving on to the next PbB level.

The probability wheel used for the encoding is radially divided into two sectors of adjustable relative areas -- one orange and one blue. The wheel can be spun so that it randomly stops with either sector under a pointer. For a particular response rate R at a given PbB level L, the wheel was set at some relative area of blue, and the expert was asked to consider the following question:

> Is it more probable that the true response rate at blood-lead level L is less than R, or that a random spin the wheel would stop with the pointer on blue?

Thus, the expert was asked to weigh two probabilities and decide which was greater. The expert's answer determined whether the proportion of blue was increased or decreased, and the question was posed again. This procedure was continued until the wheel was set such that the expert considered it equally likely that the wheel would randomly land on blue as that the true response rate was less than R for PbB level L. The relative area of blue was then provisionally taken as the judged cumulative probability $F(R|L)$. The procedure was repeated for the required response rates until an entire cumulative probability distribution over R for $0 < R < 1.0$ had been encoded for a PbB of L.

If after a period of time the expert did not become comfortable with the probability wheel, an encoding method based on successive intervals and hypothetical bets was used. In this case, a cumulative subjective probability of $F(R|L) = 0.50$ was first encoded by having the expert specify a population response rate R for the PbB level under consideration such that if a bet were made such that the payoff depended on

whether the true response rate turned out to be above or below R, the expert would be indifferent as to which bet he held. In other words, the expert specified a response rate such that in his or her judgment it was equally likely that the true rate is above as below it. Frequently, it turned out that the expert was not indifferent between the two sides of the bet and that with a little more effort it was possible to find a response rate at which he or she was.

Having determined the response rate corresponding to a subjective probability of 0.50, the expert was then asked to imagine that in fact the necessary research had been done, that the true response rate was now known, and that it was greater than the value he had just indicated. (On half the occasions at this point, he or she was to imagine that the true response rate was less than that earlier specified.) Given that the true rate was greater than that which had been indicated, a new response rate R was to be specified for a new bet such that the expert would be indifferent as to whether his payoff depended on the true population response rate being above or below it. Once such a value was determined, this corresponded to a subjective probability of $F(R|L) = 0.75$. In this manner, response rates corresponding to subjective probabilities of 0.25, 0.50, and 0.75 were encoded. Response rates corresponding to subjective probabilities of 0.01 and 0.99 were also determined by imagining bets with 1:99 and 99:1 odds.

After the cumulative probability distribution over response rate was encoded for a given PbB level, regardless of the encoding method employed, the distribution was graphed and shown to the expert for discussion. Any inconsistencies were pointed out and resolved. Such inconsistencies appeared as dips in the graphs, which should rise monotonically as the response rate goes from 0 to 1 (see Wallsten et al., 1983, for details).

The implications of the encoded distributions were also discussed. For example, for $R_j > R_i$, $F(R_i|L) = 0.25$ and $F(R_j|L) = 0.75$. These probabilities imply that the expert considers the true response rate equally likely to fall inside the (R_i, R_j) interval as

outside it. If the expert disagreed with this or any of several other implications, adjustments were made. When the expert was satisfied that the encoded distribution represented his probabilistic judgment, another PbB level was selected and the process was repeated.

When distributions had been encoded for all selected PbB levels, they were plotted on a graph and discussed with the expert. If two distributions crossed at any point, this inconsistency was resolved (see Wallsten et al., 1983, for details). When the expert was satisfied that the entire set of distributions represented his or her probabilistic judgments, the encoding was finished.

Following the first session, the encoded judgments were fit by distributions in a manner described in Sec. 3.5. Summaries of these fits and the encoded judgments, along with explanatory text, were sent to the expert prior to the second visit. During the second visit, all of the materials were reviewed, additional probability encoding was carried out as necessary, and the expert was asked to indicate whether the fitted distributions correctly represented his or her judgments. In a few instances, function fitting was carried out later, and further discussion was continued by letter and telephone.

As seen in App. C, the fitted functions were very close to the encoded probabilities. When an expert finally felt that the probabilities correctly represented his or her judgments, he or she felt that the fitted functions did as well.

3.5 REPRESENTING THE JUDGMENTS

Calculating risk required fitting a distribution to the assessed points. Because the probabilistic judgments about the dose-response relationships are generally "s"-shaped over the closed [0,1] interval, as demonstrated in Wallsten et al. (1983) and Whitfield and Wallsten (1984), it can be difficult to represent them with simple mathematical functions. One particularly useful function that is relatively easy to work

with is the normal-on-log-odds distribution function, which is obtained by fitting a normal distribution to the natural log of the odds implied by the population response rate R. Thus, the following relationships can be defined:

$$X = \frac{R}{1 - R}, \text{ for } 0 \leq X \leq \infty$$

$$Y = \ln(X), \text{ for } -\infty \leq Y \leq \infty$$

where X is the odds variable and Y is the log of the odds variable. The variable Y is assumed to be normally distributed with mean μ and variance σ^2. If Y is normally distributed, then X is log-normally distributed, and the distribution induced on R is called a normal-on-log-odds distribution. Although the normal-on-log-odds distribution cannot be expressed in a closed form, all probabilities of interest can be easily derived from the normal distribution over Y, as discussed in App. B.

A separate normal-on-log-odds distribution function was fit to the elicited cumulative distribution functions (CDFs) or the probability judgments over response rate R, at each PbB level L_j, where $j = 1, \ldots, m$. This fitting was done by deriving least-squares estimates of the parameters μ_j and σ_j, denoted by $\hat{\mu}_j$ and $\hat{\sigma}_j$, for the normal distribution applied to the log-odds transformed variable Y, as described in App. B. These distributions, each with separate least-squares estimates for μ_j and σ_j, are referred to as the best-fitting normal-on-log-odds distributions. The standard regression r^2 statistic, which compares actual judgments to those predicted by the fitted function, assesses how well the derived distribution represents the probability judgments at a given PbB level.

As presented fully in App. B, the normal-on-log-odds distribution described the encoded judgments very well, with r^2 values generally exceeding 0.95. (An r^2 value of 1.0 indicates a perfect fit.) Nevertheless, goodness of fit was sacrificed to a small

degree by fitting an expert's probability judgments with a family of equal-variance normal-on-log-odds distributions. Such a procedure ensures that the derived distributions at two different lead levels never cross one another, not even at the extremes. Crossing would imply that exceeding a specified response rate is judged more probable at blood-lead level L_1 than at L_2, where L_2 is greater than L_1.

Equal-variance normal-on-log-odds distributions were derived by (1) obtaining a least-squares estimate of the standard deviation pooled over lead levels, $\hat{\sigma}'_p$, (2) setting all the $\hat{\sigma}'_j$ equal to $\hat{\sigma}'_p$, and (3) finding new least-squares estimates of the means, $\hat{\mu}'_j$ (see App. B). The result is a set of normal-on-log-odds distributions, one for each blood-lead level L_j, that have a common variance and differ only in the mean. Their goodness of fit can also be assessed by calculating r^2 between observed and predicted judgments.

Because for each PbB level L_j, probability judgments were elicited over virtually the entire [0,1] interval, and these judgments exhibited no inconsistencies (i.e., crossings), the fit of the equal-variance normal-on-log-odds distributions was generally close to that of the best-fitting normal-on-log-odds distributions. Because the equal-variance distributions cannot give rise to inconsistencies, these representations of an expert's judgments are preferred for estimating risk. They were in fact accepted by each expert, with a slight adjustment required for Expert D.

3.6 THE EXPERTS

The individuals listed in Table 3.2 were selected in the fall of 1984 by EPA/OAQPS for participation in the study as experts. All of them agreed to participate, with the understanding that their judgments would be kept anonymous. Results are therefore presented for Experts A through E, the codes being randomly assigned to the names listed in Table 3.2.

TABLE 3.2 Experts Who Participated in the Hemoglobin Encodings

Expert	Affiliation
Dr. Julian Chisolm	John F. Kennedy Institute Johns Hopkins University
Dr. Bernard Davidow	New York City Board of Health
Dr. Paul Hammond	Dept. of Environmental Health University of Cincinnati
Dr. Sergio Piomelli	Columbia University New York City Lead Poisoning Prevention Program
Dr. John Rosen	Montefiore Hospital and Medical Center

3.7 RESULTS

The qualitative discussions about the effects of lead are summarized in App. B. The detailed quantitative results, including the encoded judgments, parameter values for the fitted functions, and goodness-of-fit measures are also given in App. B. This section uses the functions fit to each individual's probabilistic judgments to summarize the quantitative results. The functions accurately represent the underlying encoded values, both because goodness of fit was excellent and because each expert endorsed the output of the respective functions as representing his judgments.

Recall that probabilistic judgments were encoded about dose-response functions for lead-induced hemoglobin decrements among U.S. children aged 0-6 years. Hemoglobin levels of 9.5 g/dL and 11.0 g/dL were considered, and the experts were given the option of dividing the population of children into age groups of 0-3 and 4-6 years. Results are presented separately for the four conditions obtained by crossing the two hemoglobin levels with the two age groups.

Experts A, D, and E believe that dose-response relationships are different for the two age groups and therefore provided four distinct sets of probabilistic judgments. Expert C believes that a single dose-response function applies to children in the 0-6 age range and therefore provided a single set of probabilistic judgments for each hemoglobin level. Expert B felt uncomfortable with the notion of judgmental probability encoding and therefore did not supply any judgments. His qualitative comments are presented in Sec. B.4 of App. B.

3.7.1 Hemoglobin Levels $\leq$ 11 g/dL, Ages 0-3 Years

Figure 3.1 summarizes the judgments of Experts A, C, D, and E regarding the dose-response function for hemoglobin levels less than or equal to 11.0 g/dL in U. S. children age 0-3 years. (With Expert C's concurrence, his judgments for age 0-6 are reproduced in both the 0-3 and 4-6 age groups.) Each person's judgments are shown in a separate panel.

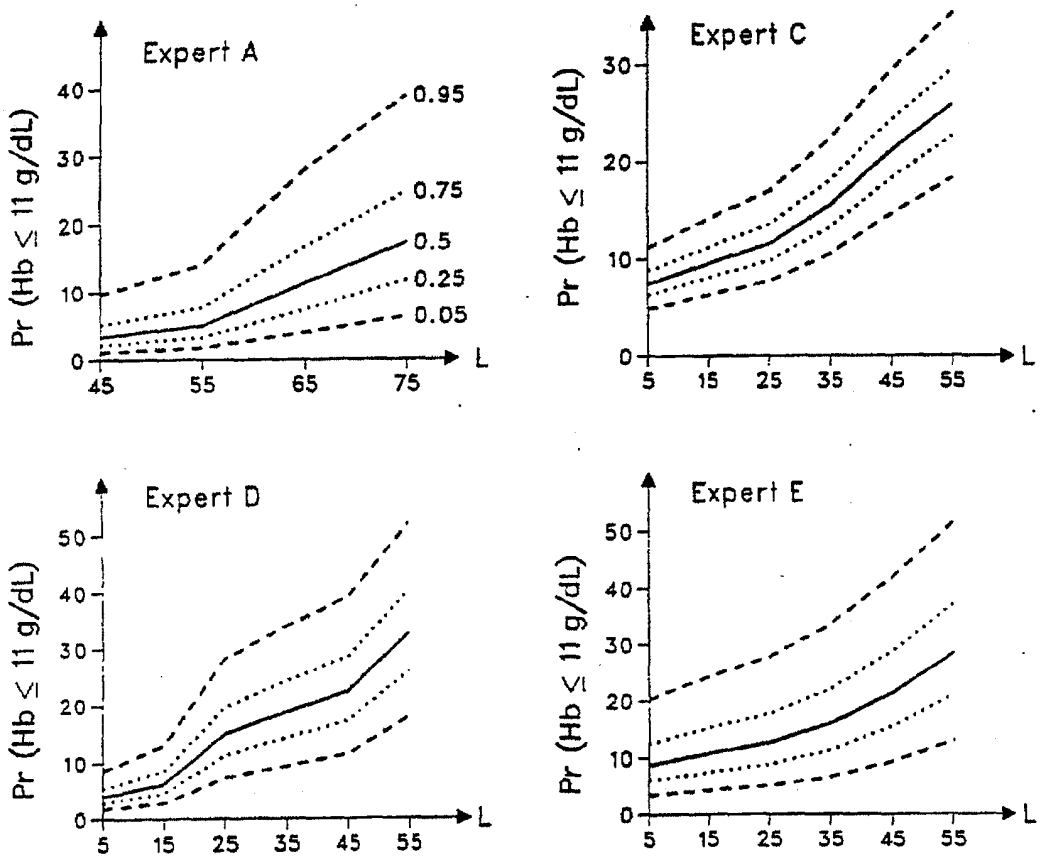

FIGURE 3.1 Dose-Response Functions, Hemoglobin Level ≤ 11 g/dL, Ages 0-3 Years

The graphs are analogous to the usual dose-response functions found in the literature, except, of course, that they are based on probabilistic judgments rather than on direct data. The dark, central curve in each panel shows the median judged dose-response curve for each expert. In other words, for a given panel, according to that expert, at each blood lead level there is a 0.50 probability that the true response rate is above the indicated value and a 0.50 probability that the true response rate is below it.

The two lighter lines on either side of the median curve contain the central 50% credible interval. Thus, according to the expert in a particular panel, at each blood lead level there is a 0.25 probability that the true response rate is below the lower light curve, a 0.50 probability that it is between the two light curves, and a 0.25 probability that it is above the upper one. In a similar manner, the dashed and dotted pairs of curves contain central 90% and 98% credible intervals, respectively.

The more tightly packed the family of functions are in a panel, the less uncertainty an expert indicates in his judgments. Thus, Figure 3.1 gives a rather complete representation of each expert's probabilistic judgments.

Figure 3.2 is less complete, but provides a convenient means of comparing judgments across experts. The axes are the same as in Figure 3.1. The vertical bars at each lead level represent each expert's central 90% credible interval for response rate, and the dot within each bar indicates the median judgment at that lead level.

Note first that there is substantial overlap in the judgments of experts C, D, and E. This is true despite the fact that expert C's judgments are for the full 0 to 6 age range, while those of the others are restricted to age 0-3. The judgments of expert A tend to differ from those of C, D, and E.

Considering the overlapping judgments of C, D, and E first, note that none suggests a threshold. According to their median judgments, the best estimate of response rate is between 4% and 9% at PbB = 5 μg/dL, rising to between 22% and 28% at PbB = 55 μg/dL.

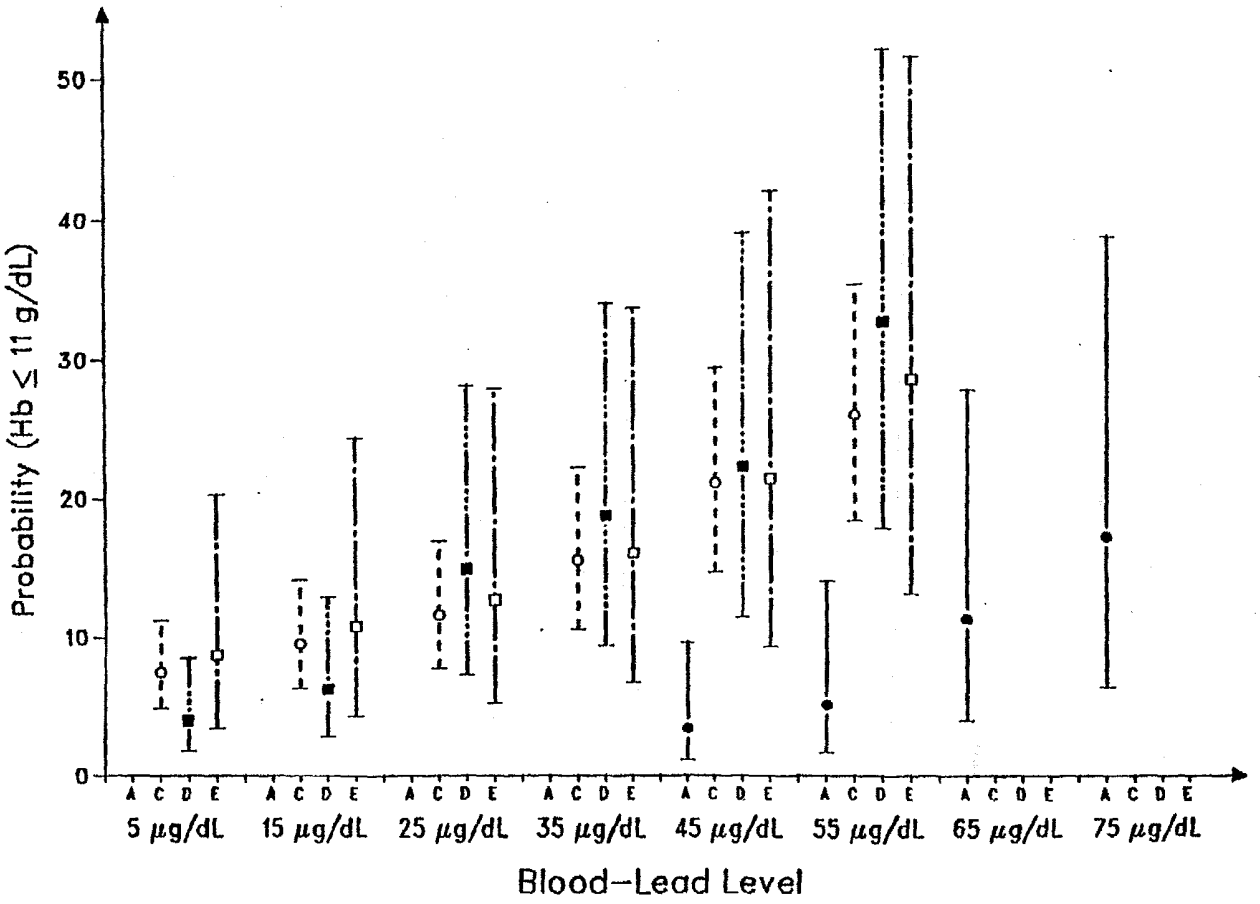

FIGURE 3.2 Comparison of Judgments, Hb ≤ 11 g/dL, Ages 0-3 Years

Expert C expresses the least uncertainty in his judgment about dose-response function, and Expert E expresses the most, although in all cases, variance in the judgments increases with PbB within the range considered (c.f. the 90% credible intervals represented by the vertical bars in Figure 3.2). Thus, considering the 90% credible intervals for experts C, D, and E, response rate is greater than a value between 12% and 28% at PbB = 5 μg/dL, with probability 0.05. Response rate then rises, and exceeds a value between 40% and 62% at PbB = 55 μg/dL, also with probability 0.05.

Expert A's judgments differ from those of Experts C, D, and E. According to his judgments, blood leads less than about 45 μg/dL do not cause hemoglobin levels to dip below 11 g/dL in this age group. At PbB = 45 μg/dL, the most likely response rate is 3%, and this rises to a most likely rate of 16% at PbB = 75 μg/dL. Furthermore, with probability 0.05, the response rate at PbB = 45 μg/dL is greater than 15%, and with the same probability it is greater than 50% at PbB = 75 μg/dL. Although our primary interest is in blood leads less than or equal to 55 μg/dL, we inquired about the higher levels here to provide a fuller picture of Expert A's judgments.

3.7.2 Hemoglobin Levels $\leq$ 11 g/dL, Ages 4-6 Years

Figures 3.3 and 3.4 display the judgments regarding dose- response functions for hemoglobin levels less than or equal to 11 g/dL for children age 4-6 years. They are read in exactly the same way as are Figures 3.1 and 3.2.

As before, the judgments of experts C, D, and E overlap considerably, and differ from those of A. According to C, D, and E, there is no threshold; the most likely response rates are between 2% and 7% at PbB = 5 μg/dL, rising to between 14% and 26% at PbB = 55 μg/dL.

In this case, expert D expresses the least uncertainty, while expert E continues to express the greatest uncertainty. With probability 0.05, response rate exceeds a value between 6% and 19% at PbB = 5 μg/dL, and with the same probability it exceeds a value between 19% and 46% at PbB = 55 μg/dL.

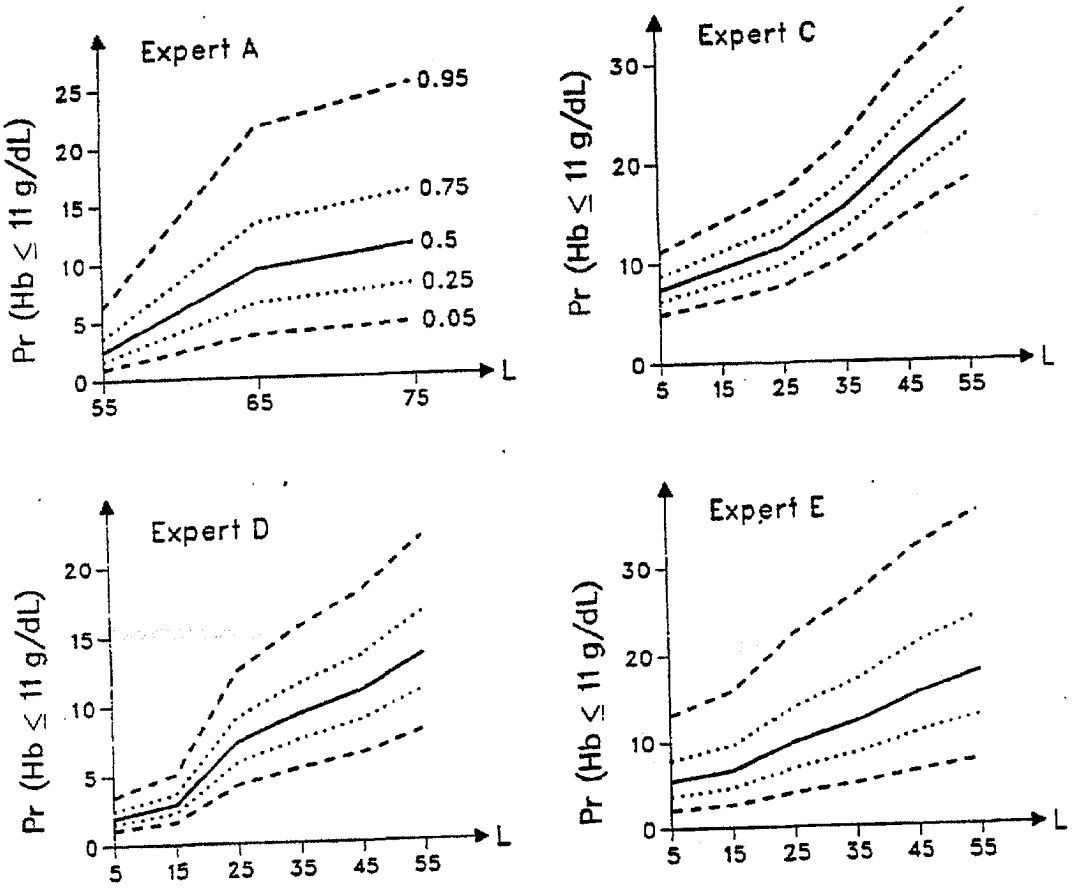

FIGURE 3.3 Dose-Response Functions, Hemoglobin Level ≤ 11 g/dL, Ages 4-6 Years

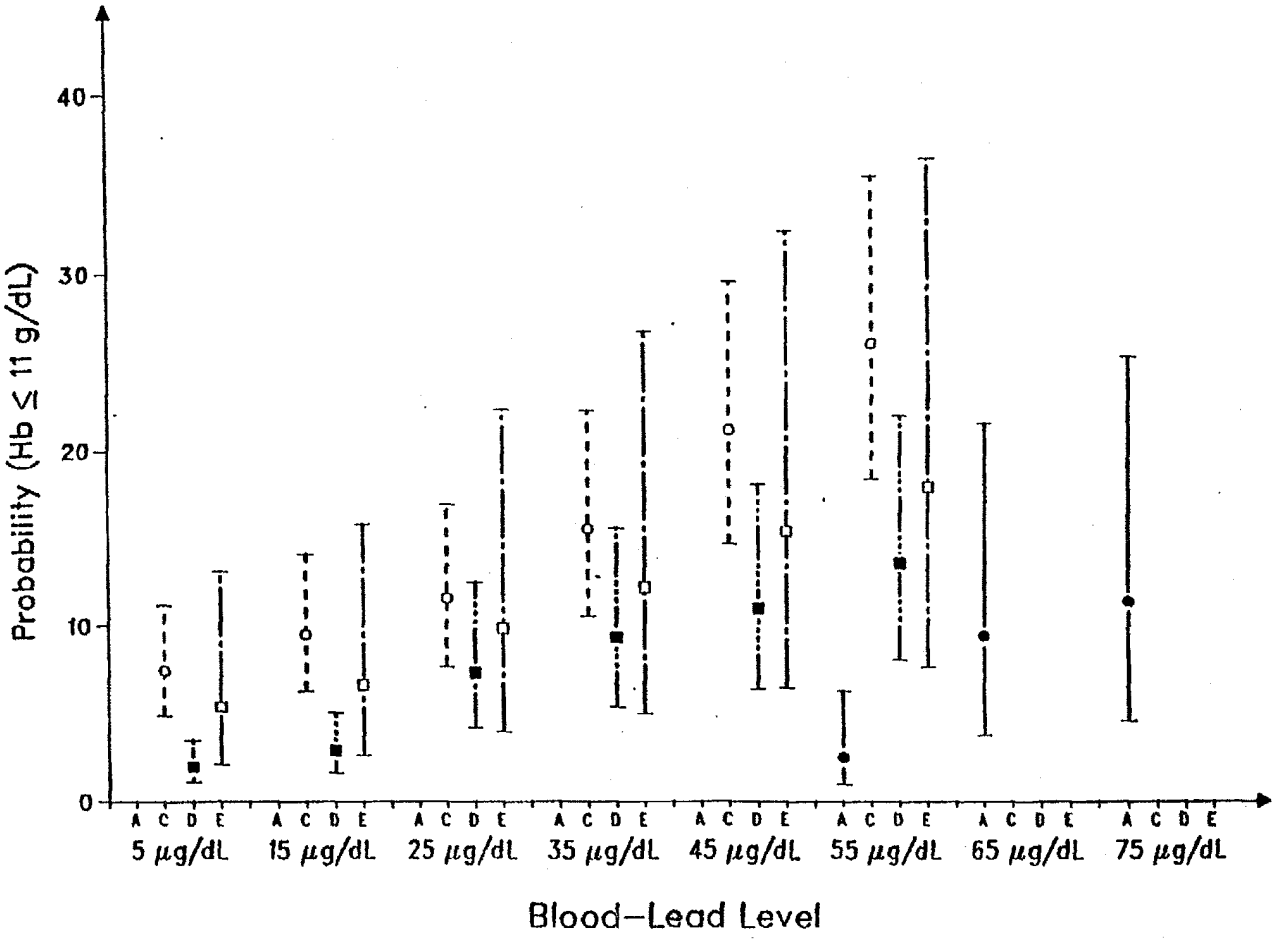

FIGURE 3.4 Comparison of Judgments, Hb ≤ 11 g/dL, Ages 4-6 Years

As before, expert A's view is that there are extremely low probabilities of lead-induced hemoglobin effects at the exposure levels considered. In his judgment, blood-lead levels less than about 55 μg/dL do not cause hemoglobin levels to go below 11 g/dL in this age group. The most likely response rate at 55 μg/dL is 2%, while at 75 μg/dL it is 11%. With probability 0.05, response rate exceeds 9% at 55 μg/dL, and with the same probability it exceeds 34% at 75 μg/dL.

3.7.3 Hemoglobin Levels ≤ 9.5 g/dL, Ages 0-3 Years

Figures 3.5 and 3.6 display the judgments regarding dose-response functions for hemoglobin levels less than or equal to 9.5 g/dL for children age 0-3 years. They are read in the same manner as are the previous pairs of figures.

In expert A's judgment, no amount of lead within the range of interest (below 55 μg/dL) will cause hemoglobin levels to drop as low as 9.5 g/dL, and therefore no results are shown for him. Also, although the judgments for C, D, and E tend to overlap, they do not display the same degree of similarity that they do at the higher hemoglobin level.

In particular, looking at the median judgments, C and E agree in estimating the most likely response rate at 2%, while D is certain that the response rate is 0%, at PbB = 5 μg/dL. Then, expert C's median response rate judgment increases relatively slowly with blood lead, while that of expert D increases relatively quickly. As a result, at PbB = 55 μg/dL D and E agree that the most likely response rate is 20%, while C considers it to be 7%.

Note further that expert E expresses considerable uncertainty in his judgments, while C and D express much less. Thus, depending on the expert, at PbB = 5 μg/dL response rate exceeds 3% or 18% with probability 0.05 (recall that according to expert D it is definitely 0%). At PbB = 55 μg/dL, it exceeds a value from 13% to 70% with probability 0.05.

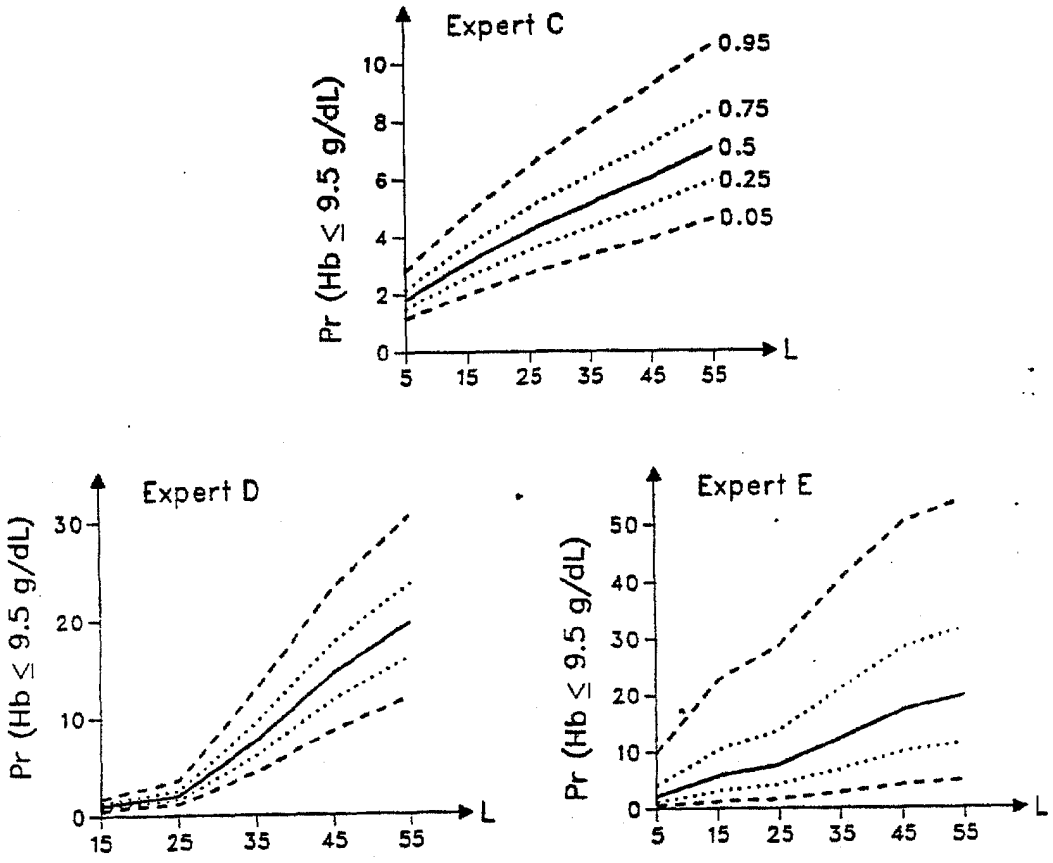

FIGURE 3.5 Dose-Response Functions, Hemoglobin Level ≤ 9.5 g/dL, Ages 0-3 Years

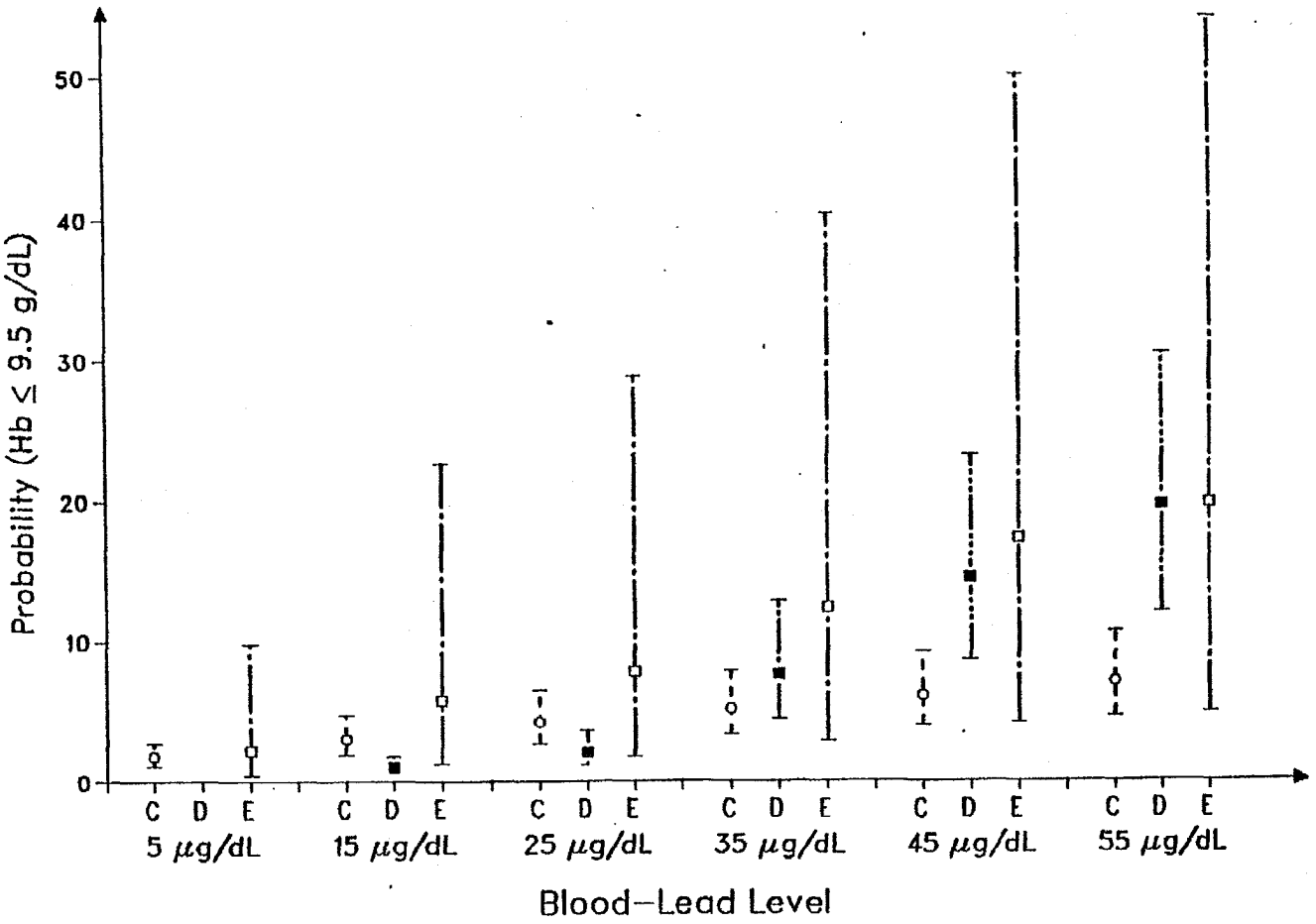

FIGURE 3.6 Comparison of Judgments, Hb ≤ 9.5, Ages 0-3 Years

3.7.4 Hemoglobin Levels ≤ 9.5 g/dL, Ages 4-6 Years

Figures 3.7 and 3.8 display the judgments regarding dose/response functions for hemoglobin levels less than or equal to 9.5 g/dL for children age 4-6 years. They are read in the same manner as are the previous pairs of figures. The pattern of judgments is very similar to that seen for the younger age group, but the response rate values are lower.

Expert A judged that blood lead at the levels under consideration would not reduce hemoglobin to 9.5 g/dL, and expert D judged that it would take at least 25 μg/dL of blood lead to do so. The median judgments of experts C and E are consistently similar, while those of expert D join them at PbB = 35 μg/dL and higher. For experts C, D, and E, the median judgments at PbB = 5 μg/dL are 0% and 2%; at PbB = 55 μg/dL, they range from 7% to 13%.

Again, expert E expresses the greatest degree of uncertainty. For C and E, at PbB = 5 μg/dL, response rate exceeds 3% and 9%, respectively, with probability 0.05. For C, D, and E, at PbB = 55 μg/dL, with probability 0.05, response rate exceeds 8% to 50%.

3.8 DISCUSSION

It must be borne in mind that these judgments are with respect to lead-induced response rate over and above any base response rate due to iron deficiency and other factors. Thus, levels of uncertainty, as well as differences or similarities of opinion reflected here, are limited solely to the effects of lead on hemoglobin level.

Although they differ in detail, it is notable that the judgments of experts C, D, and E have overriding similarities. Of particular interest, C's judgments for the 0-6 years age group tend to be just below those of D and E for ages 0-3 years, and just above their judgments for ages 4-6 years. Collectively, the judgments of C, D, and E differ from those of A, who considers that blood-lead levels considered have little or no effect on hemoglobin levels.

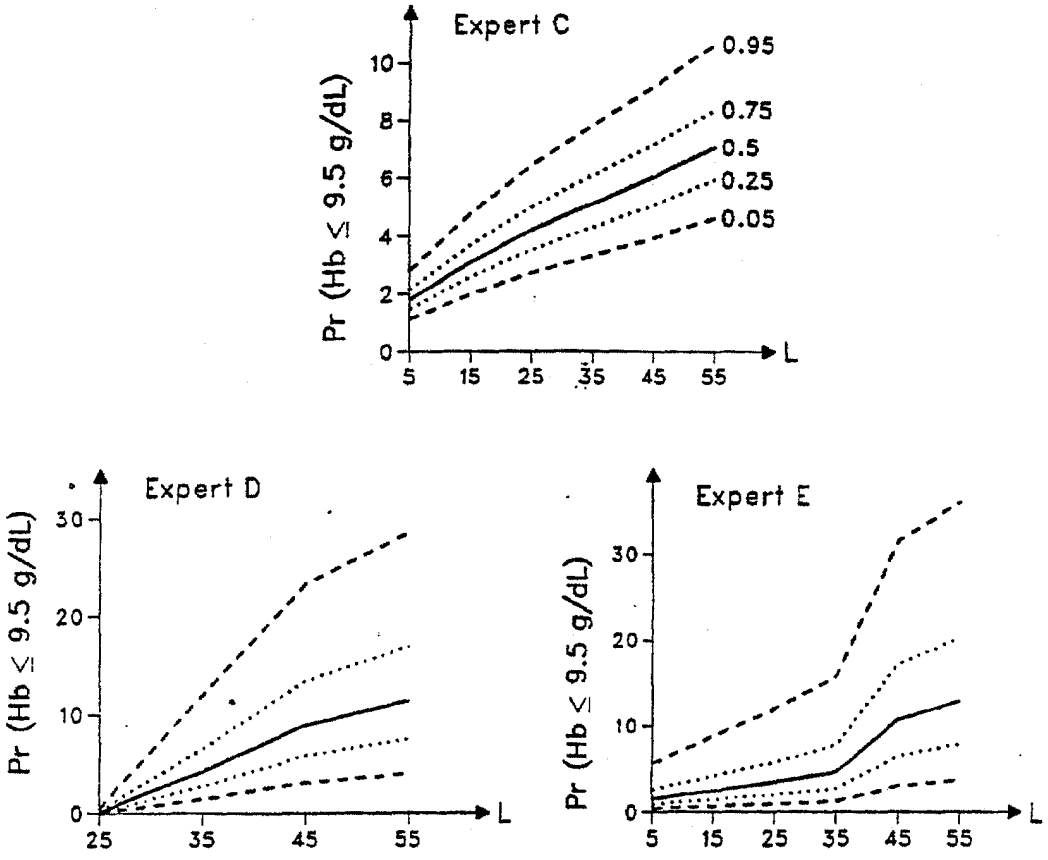

FIGURE 3.7 Dose-Response Functions, Hemoglobin Level ≤ 9.5 g/dL Ages 4-6 Years

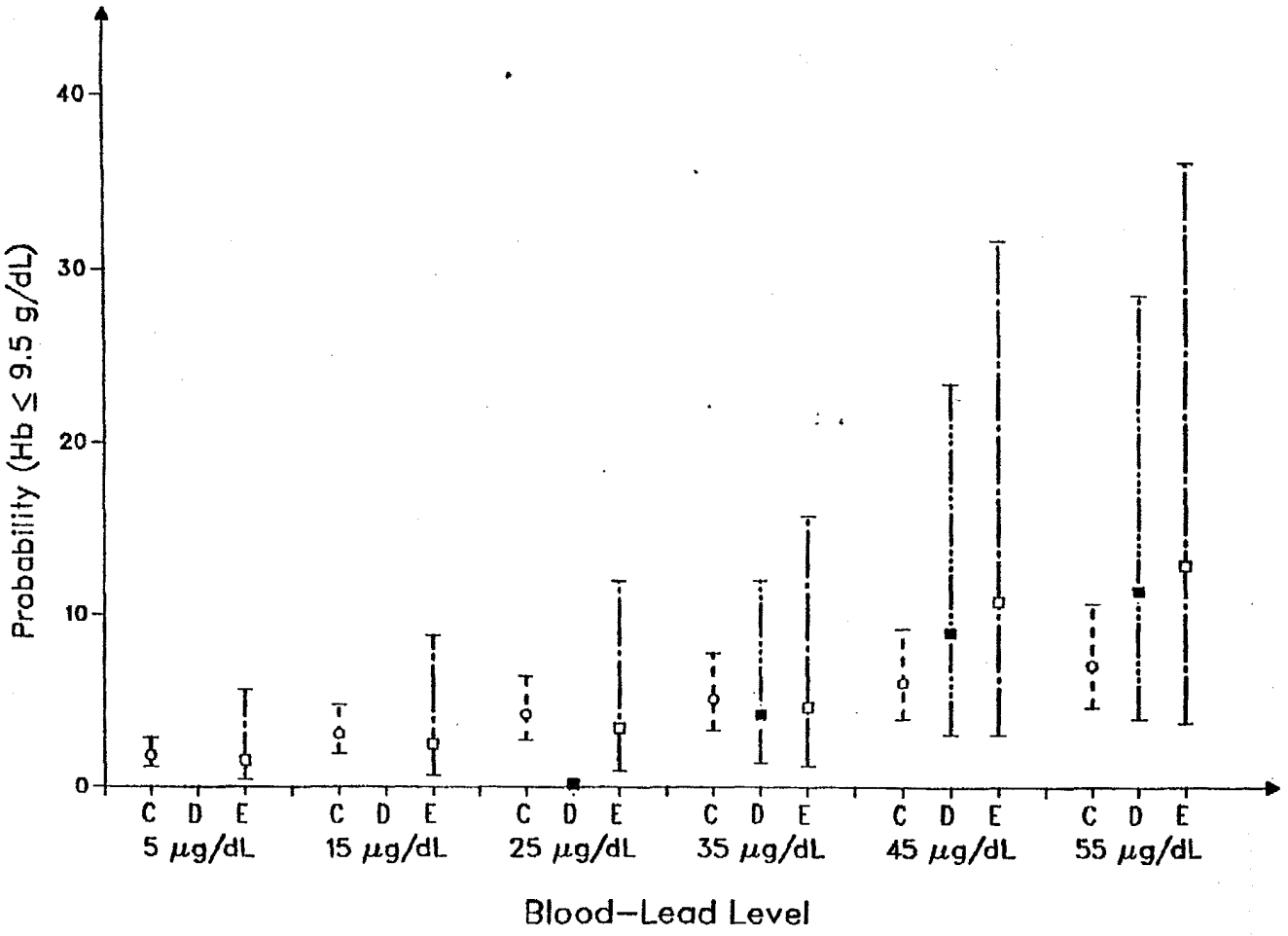

FIGURE 3.8 Comparison of Judgments, Hb ≤ 9.5, Ages 4-6 Years

In all cases, uncertainty (indicated by variance) increases with blood lead, as one would expect on rational grounds. This increase is rational because statistical uncertainty about a binomial parameter (response rate, in this case) increases as the estimated value moves toward the center of the range for that parameter.

There is also less information available at the lower hemoglobin level than at the higher one, if for no other reason than that it occurs less frequently. As a result, judgments depend on a greater extrapolation from available data, and greater disagreement among experts is to be expected. This is precisely what happened. It is interesting, however, that experts A, C, and D exhibit more certainty at 9.5 g/dL than at 11.0 g/dL, despite the greater extrapolation. Indeed, expert A is positive that there is no effect here. Apparently, A, C, and D's understanding of the effects of lead on hemoglobin level are such that each feels more confident about response rates at the lower hemoglobin level. The reverse is true for expert E.

Finally, it should be recalled that the experts were selected to represent the range of respected opinion. The fact that nevertheless, there is considerable similarity in the pattern of judgments illustrates some of the benefits of the probability encoding procedure. The differences of opinion that have appeared in the literature and in debate have been due to a mix of disagreements about definitions, about interpretations of and extrapolations from data, and about what constitutes proper public health policy. However, it is clear that the experts exhibit much less disagreement when they are required to focus on the scientific issues and to consider carefully their uncertainty about these matters.

The judgments represented here are of interest in their own right. However, in Chapter 5 they will be combined with estimated blood lead distributions under alternative NAAQS to obtain derived risk distributions.

4 PROBABILISTIC DOSE-EFFECT AND DOSE-RESPONSE RELATIONSHIPS FOR LEAD-INDUCED IQ DECREMENTS

As with hemoglobin, there are incomplete data regarding IQ effects of lead in the population of U.S. children prior to their seventh birthdays. Indeed, it is fair to say that the available data are themselves very controversial, making it doubly necessary to encode experts' subjective probabilities.

An exact parallel to the hemoglobin situation would have entailed defining dose-response functions in terms of percentage of the population with lead-induced IQ decrements exceeding a given amount, and then encoding subjective probabilities about such functions. This proved to be infeasible, however, because it did not frame the issue in the way that researchers in the area think about it, nor in terms of events that were in principle observable. Inferences about the effects of lead on IQ require complex statistics on group data, because IQ depends on multiple, correlated factors acting over long periods of time. Thus, it is impossible to observe an IQ decrement in an individual and to attribute that decrement to a particular cause.

In order to pose questions about outcomes that the experts commonly think about and to elicit probabilistic judgments that are suitable for use in the risk assessment, a hypothetical experiment was created. In this hypothetical study, very large numbers of children are randomly assigned at conception either to a control group or to one of several lead exposure groups. Exposure levels remain approximately fixed at the specified level until the children's seventh birthdays, at which time the WISC IQ test is administered. Blood lead is measured at the third birthday. The very large numbers of subjects per group eliminate the need to think about sampling error, and the random assignment of subjects to conditions eliminates the need to think about complex analyses of covariance; groups differ only in terms of exposure to lead. The experts were asked to consider this experiment and to provide subjective probabilities about expected mean IQ differences between the control group and each of the exposure groups.

Of course, such a study cannot and should not be done, but it is in principle doable. Furthermore, the hypothetical data are identical in form to data actually collected, but without the troublesome features. Indeed, the point of the necessary complex statistics on the real data is to attempt to draw inferences about what the IQ difference would be between a lead-exposed and a nonexposed group if the confounding variables were not present. Thus, researchers think about this issue, and it is precisely about it that the experts were asked to make probabilistic judgments.

In order to calculate subjective probabilities about IQ dose-response functions from judgments about mean IQ differences, it was necessary to obtain judgments about other matters as well. Thus, probabilistic judgments were encoded about the mean IQ of the unexposed control group, as well as about the IQ standard deviation within exposure groups. The IQ measure is constructed so that it is approximately normally distributed within the population. The experts were asked whether they felt that the normal distribution was reasonably applied to IQ scores within each of the exposure groups. Subsequently, subjective probabilities about dose-response functions of interest could be calculated using normal distribution theory in conjunction with the probabilistic judgments about mean IQ decrement, mean IQ of the control group, and within-group standard deviation.

Section 4.1 describes the protocol development, in which it is made clear how this hypothetical experiment was arrived at; the protocol is outlined in Sec. 4.2. Sections 4.3-4.5 discuss the conduct of the encoding sessions, the methods of probability encoding, and the mathematical representation of the judgments, respectively. Section 4.6 indicates whose judgments were encoded, Sec. 4.7 summarizes the results, and Sec 4.8 provides some discussion.

TABLE 4.1 Consultants for the IQ Protocol

Consultants	Affiliation
Dr. Vernon Benignus	HERL, U.S. EPA
Dr. Robert Bornshein	University of Cincinnati
Mr. Jeff Cohen	OAQPS, U.S. EPA
Dr. Anita Curran	Westchester County Department of Health
Dr. Gerri Dawson	University of North Carolina
Dr. Kim Dietrich	University of Cincinnati
Dr. Claire Ernhart	Cleveland Metropolitan General Hospital
Dr. Lester Grant	ECAO, U.S. EPA
Dr. Lloyd Humphreys	University of Illinois
Dr. Lyle Jones	University of North Carolina
Dr. Herbert Needleman	University of Pittsburgh
Dr. David Otto	HERL, U.S. EPA
Dr. Steven Schroeder	University of North Carolina
Dr. Bernard Weiss	University of Rochester

TABLE 4.2 Experts Who Participated in the IQ Encodings

Expert	Affiliation
Dr. Kim Dietrich	University of Cincinnati
Dr. Claire Ernhart	Cleveland Metropolitan General Hospital
Dr. Herbert Needleman	University of Pittsburgh
Dr. Michael Rutter	Institute of Psychiatry London, U.K.
Dr. Gerhard Winneke	University of Dusseldorf Dusseldorf, F.R.G.
Dr. William Yule	Institute of Psychiatry London, U.K.

4.1 PROTOCOL DEVELOPMENT

A large number of people, listed in Table 4.1, served as consultants in the development of this protocol. As already indicated, our original intention was to develop a protocol analogous to the one for hemoglobin decrement, according to which we would encode experts' probabilistic judgments about dose-response functions for lead-induced IQ decrements. In preparation, we read relevant portions of the CD and consulted with Drs. Vernon Benignus, Robert Bornshein, Kim Dietrich, David Otto, and Stephen Schroeder.

A first draft of a protocol was developed, designed to elicit probabilistic judgments about dose-response curves for adverse effects defined in terms of IQ decrements. This draft was discussed with most of the people mentioned above as well as with Drs. Claire Ernhart, Lester Grant, Lloyd Humphries, Lyle Jones, Herbert Needleman, and Bernard Weiss. Two factors became clear in the discussions.

First, although the concept of a dose-response curve is well defined in this context, it is not commonly used, for reasons already introduced in Section 4.1. The second factor was that in attempting to respond to the questions, the experts were envisioning hypothetical experiments, considering what the outcomes might be, and then extrapolating from that to the dose-response functions of interest.

Thus, it appeared that it would be far more consistent with the experts usual way of thinking and be easier for them to respond if their probabilistic judgments about outcomes of suitable hypothetical experiments were encoded directly. In addition, by properly designing these hypothetical experiments, it would be possible to eliminate the need to worry about some of the complexities that arise in actual research. The extrapolations that they were trying to do mentally could then be done mathematically. Accordingly, a new version of the protocol was developed to elicit probabilistic judgments about the outcomes of ideal hypothetical experiments. This protocol also indicated how the judgments would be used and the assumptions that would be

incorporated in working with them. The new draft was discussed with Drs. Vernon Benignus, Anita Curran, Claire Ernhart, Lloyd Humphries, Lyle Jones, Herbert Needleman, David Otto, and Stephen Schroeder. They found it much easier and more natural to respond to this version than to the previous one. However, various comments lead to a third draft of the protocol incorporating an improved ideal hypothetical experimental design. This draft was discussed with most of the people who had seen the previous one, and some very minor changes were made.

The protocol was then used with two pilot subjects, one an EPA environmental health scientist and the other a psychology graduate student doing a related dissertation. The judgmental probability encodings went very well. However, in analyzing the judgments and discussing these analyses with Dr. Lyle Jones, it became apparent that minor changes were still necessary.

Thus, a final version of the protocol was created. It was discussed with a number of the people who had seen the previous drafts, and was considered suitable for subsequent use.

4.2 PROTOCOL OUTLINE

The full protocol is provided in App. C. Only an outline is given here.

Section 1 is an introduction, describing the risk assessment project and the purpose of the probability encoding. Section 2 describes the ideal hypothetical experiment, in which large numbers of children are randomly assigned at conception to one of six exposure groups or to a control group sheltered from lead. Environmental lead exposure is roughly constant at the appropriate level for each child, such that when his or her blood lead is measured on the third birthday, it is at the designated level. The same environmental exposure continues until the seventh birthday, when the WISC-R IQ test is administered. Thus, each expert was to consider the time course of lead in the children's

systems according to his or her own theoretical understanding, subject to the constraint that blood lead was at the designated level on the third birthday.

It was explained in this section that probabilistic judgments would be encoded regarding the mean IQ decrement for each exposure group relative to the control group, the mean IQ of the control group, and within the group IQ standard deviation. Further, there was discussion regarding the shape of the within group IQ distribution. However, if an expert considered the IQ effects of lead to be different for lower socioeconomic status (SES) than for middle and upper SES subpopulations, then judgments would be encoded separately for the two groups. For this purpose, low SES children were defined as those coming from households with incomes below the fifteenth percentile.

Section 3 defined the population at risk as children up to their seventh birthday, thereby indicating why the IQ measure is taken at that time. Sections 4 and 5 specified exposure, physiological, and environmental conditions to assume for the hypothetical experiment.

As with the hemoglobin protocol, Section 6 focused on issues in the literature to elicit discussion prior to the probability encoding. The purposes were to insure that the problem was structured properly for the expert, to cause him or her to consciously review his or her knowledge prior to providing probabilistic judgments, and to provide a possible basis for understanding discrepant judgments, should they occur. Section 7 discussed factors to be aware of when encoding probabilities, and Section 8 documented the encoding procedure.

4.3 CONDUCT OF THE SESSIONS

The sessions were conducted in the Spring of 1985, and with one exception were carried out in an analogous fashion to those for hemoglobin. In lieu of providing the details again, the reader is referred to Sec. 3.3.

The single exception was that information was brought along to aid the expert, should the person have wanted it, in encoding probabilities about control group mean IQ and within group IQ standard deviation. The information concerned WISC-R IQ means and standard deviations for subpopulations stratified by various demographic variables, as compiled by Kaufman and Doppelt (1976) and subsequently summarized by Sattler (1982).*

4.4 ENCODING THE JUDGMENTS

Probability encoding began in the first session after the relevant literature and the factors affecting lead effects on IQ had been thoroughly discussed with the expert. The intention was to encode probability judgments for possible mean IQ decrements in six lead exposure groups relative to the control group. The exposures were such that at their third birthdays, members of respective groups had blood-lead levels of 5, 15, 25, 35, 45, and 55 μg/dL. However, higher levels were specified for a few of the experts.

To determine useful points in the space of mean IQ decrements for encoding probabilities, each expert was first asked to supply broad "99% confidence interval" ranges for possible mean IQ differences between the control and each of the exposure groups. We helped the expert determine these ranges by posing suitable questions (e.g., after a range was specified, we asked whether the expert could construct a plausible explanation if someday it were found that the true mean difference was outside that range). In a similar manner, "99% confidence intervals" were obtained for the mean IQ of the control group and the within group standard deviation. At this point, all of the experts agreed that within SES level the same standard deviation judgments applied to all

*It should be pointed out that these normative data could not be used in place of The experts' judgments regarding control group mean IQ and within group IQ standard deviation, because they were based on children with varying amounts of lead in their systems. Thus, each expert had to adjust these values as he or she thought appropriate given his or view on The IQ effects of lead. Each person had the option, of course, of taking the tabulated values as point estimates.

exposure groups, and that it was reasonable to assume normal IQ distributions within subgroups.

Functions were fit to the judgments following the first session. Summaries of the fits plus descriptive information were sent to those experts in the U. S. for their review prior to the second visit with them. Due to time and financial constraints, this was not possible to provide the European experts with an opportunity to review the same materials prior to the second meeting. Instead, they were brought to the second session and explained there. During the second visit with both the European and the U.S. experts, the materials were reviewed, additional probability encoding was done as necessary, and each person indicated whether the fitted distribution correctly represented his or her judgments. In a few instances, subsequent function fitting was carried out, and further discussion was continued by mail and telephone. Fortuitously, two of the Europeans were in the United States shortly after the second session, and it was possible to have a short third session with each one.

As will be seen in App. D, the fitted functions were in fact very close to the encoded probabilities. When we reached the point that an expert felt that the probabilities correctly represented his or her judgments, that person felt that the fitted function did as well.

4.5 REPRESENTING THE JUDGMENTS

As with hemoglobin, it was necessary to fit distributions to the assessed points to calculate risk. The nature of the variables being judged were different than with hemoglobin, but as before, the judgments were generally S-shaped. Judged mean IQ difference was bounded from below by 0 (theoretically, this need not have been the case, but none of the experts gave any credence to the possibility of lead enhancing IQ). Judgments about mean control group IQ and within-group IQ standard deviation were so far from any limits that for practical purposes they could be treated as unbounded. Thus, different distributions were required here than with hemoglobin.