EXECUTIVE SUMMARY

In 1995 the U.S. Department of Housing and Urban Development (HUD) issued *Guidelines for the Evaluation and Control of Lead-Based Paint in Housing*. The *Guidelines* were written to provide detailed, comprehensive technical information on how to identify lead-based paint hazards in housing and how to control hazards safely and efficiently. Two protocols were defined in the *Guidelines* to facilitate the identification of housing that needed to be evaluated and possibly treated. These were the risk assessment protocol and the lead hazard screen protocol. When the *Guidelines* were released, there was a strong consensus among professionals that these protocols represented the best expert judgment available but there was also recognition that further research to validate the protocols was necessary.

In July 1995, one month after the *Guidelines* were published, the Federal Task Force on Lead-Based Paint Hazard Reduction and Financing released a report that identified research into the utility of the protocols as being a key topic for investigation. On November 27, 1996, the HUD Office of Lead Hazard Control issued a Notice of Funding Availability (NOFA) indicating that such research was a priority for its research program. In 1997, HUD awarded the National Center for Healthy Housing ((NCHH) formerly, the National Center for Lead-Safe Housing) a grant to assess HUD's risk assessment and lead hazard screen protocols found in the 1995 *Guidelines*. In 2001, EPA released regulations that changed the numeric standards for dust lead and soil lead hazards. The field test of the protocols presented in this report (National Risk Assessment Study) address the effectiveness of risk assessment and screening protocols using both the 1995 and 2001 standards.

Purpose

The primary purpose of this study was to assess under what conditions HUD's risk assessment and lead hazard screening protocols are accurate predictors of children's lead exposure. The study attempted to identify ways to improve the accuracy of the protocols. NCHH conducted a detailed, multi-media environmental assessment of residential lead in a variety of housing and linked those results to children's blood lead levels. The resulting data set served as a test bed for a number of statistical analyses that address many of the key issues regarding the identification of housing that contributes to childhood lead poisoning.

The study had the following goals:

1. To assess the ability of the current and original HUD/EPA risk assessment protocols to predict dwelling units that are likely to house children having elevated blood lead levels, and assess the effect of modifying the protocols.
2. To assess the ability of the current and original HUD/EPA lead hazard screening protocols to predict the need for risk assessments, to predict dwelling units that are likely to house children having elevated blood lead levels, and assess the effect of modifying the protocols.

3. To describe the contribution of friction and impact surfaces to floor dust lead loadings.
4. To assess the ability of the current HUD paint film quality classification system to predict rooms and dwellings that are likely to have elevated dust lead loadings.
5. To estimate the effect of dust lead measurement error on dust lead loadings.

Study Design

The National Risk Assessment Study was conducted in three locations: Baltimore County, Maryland; Milwaukee, Wisconsin; and New York City, New York. Housing in the latter two locations was chosen to represent older housing (pre-1950) while housing in Baltimore County was limited to that constructed between 1950 and 1978. The study population consisted of dwelling units housing a child, one to three years of age, who lived at the residence for at least six months prior to enrollment.

In Milwaukee and New York City, children's blood lead results, as reported by local blood lead registries and participating clinics, were used to identify dwellings eligible for the study. A case-control methodology was used in which dwellings were stratified by outcomes: half of the dwellings enrolled housed an EBL child (≥10 ug/dL) and the other half housed a child with a non-elevated blood lead level. Baltimore County did not have an accessible blood lead registry nor a blood lead screening rate sufficient to identify enough children for a four to six- month case-control study. NCHH elected to use a cross-sectional study design at this site. Potential study subjects were identified by the County based on a match of birth records and age of housing data from the tax assessor's office. Blood lead samples were drawn concurrently with the environmental sampling, so the population could not be selected on the basis of blood lead level. A local Institutional Review Board at each site approved the study design, protocols and forms.

A comprehensive set of environmental tests were taken in each home, including a visual inspection, XRF inspection, dust wipes, paint chips, soil and water samples. The environmental testing was completed in the home soon after the blood tests were reported so that it occurred either prior to or concurrent with the family receiving information on the benefits of lead-specific cleaning, to reduce the likelihood of cleaning prior to the environmental testing. Blood lead levels were collected or reported from one eligible child in the family, and a family interview was administered. The tests occurred within three weeks of each other and all the data were collected within one five-month summer "season" to reduce confounding factors. The original enrollment plan targeted 75 pre-1950 units in both New York City and Milwaukee, and 100 units built between 1950 and 1978 in Baltimore County.

From June to October 1998, certified risk assessors conducted comprehensive risk assessments/paint inspections in two hundred fifty-four dwellings. In Milwaukee and New York City, the recruitment goal of the case-control study design was achieved with 153 enrolled dwellings housing approximately equal numbers of non-EBL children (<10 μg/dL) and EBL children (≥10 μg/dL).

Table ES-1: Number of Dwellings Enrolled by Site and Elevated Blood Status

Site	Non-EBL Child Present (< 10 μg/dL)	EBL Child Present (≥ 10 μg/dL)	All Dwellings
Baltimore County	99	1	100
Milwaukee	42	37	80*
New York City	35	39	74
Total	176	77	254

*One unit was enrolled and tested based on the verbal report of a blood lead test, but the blood lead result was never confirmed by electronic report.

In Baltimore County, where the limited screening data precluded the use of a case-control design, the cross-sectional approach resulted in a sample dominated by dwellings that housed children with blood lead levels below 10 μg/dL. Since the study population was made up solely of dwellings that were built after 1950, and had a largely White population, the results are consistent with the CDC's National Health and Nutrition Examination survey (NHANES) for 1991-1994. The NHANES survey estimated that 1.4% of White children living in housing built between 1946 and 1973 would have a blood lead level above 10 μg/dL.

Summary of Analyses and Results

Goals 1 & 2: To assess the ability of the HUD/EPA risk assessment and screening protocols to predict dwelling units that are likely to house children having elevated blood lead levels, and assess the effect of modifying the protocols.

A standard method to assess the accuracy of a diagnostic test is to examine the performance characteristics of the test, using four probability measures: sensitivity, specificity, positive predictive value, and negative predictive value. All four terms are defined in the report; for this summary, sensitivity and specificity are defined below:

- *Sensitivity (or True Positive Rate)*: Probability that a dwelling unit fails an environmental assessment given that there is a resident child with an elevated blood lead concentration.
- *Specificity (or True Negative Rate)*: Probability that a dwelling unit passes an environmental assessment given that a resident child does not have an elevated blood lead concentration.

The analyses that were conducted also included a statistical test of independence between the environmental assessment result (pass/fail) and the presence or absence of a child with an elevated blood lead level. A result with a p-value less than 0.05 indicated that the environmental assessment result did not predict the child’s blood lead status.

Although the original intent of the study was to combine information from the three study sites, it proved to be inappropriate to do so. Substantial differences in both blood and environmental lead levels were found across all sites. Because only one of the 100 children enrolled in Baltimore County had an elevated blood lead level, these results could not be used to assess the effectiveness of the environmental testing protocols. As presented below, even though Milwaukee and New York City had similar aged (pre-1950) housing in the study, the environmental lead levels were very different at the two sites.

Results:

- **This study supports the premise that environmental lead results can be used to identify homes where children are likely to have elevated blood lead levels.** Analyses described below suggest that an environmental lead test can be a fairly predictive tool by maintaining the current standards but dropping window sill dust tests and assessments of paint. Further study may conclude that changes to the current standards could further improve the risk assessment protocols.
- However, **neither the current risk assessment protocols nor the screening protocols were significant predictors of the blood lead status** ($< \text{or} \geq 10$ μg/dl) of a child in the dwelling.
- **Housing units in New York and Milwaukee had significantly different environmental lead levels,** although the sites had a similar distribution of children with and without elevated blood lead levels. Only water lead levels were similar at the two sites (Table ES-2).
    - **The arithmetic mean and maximum floor dust lead loading and the perimeter soil lead concentration for a dwelling were significantly different in the dwellings with and without an enrolled child with an elevated blood lead level *in Milwaukee*.** Surprisingly, neither window sill nor trough lead loadings were significantly related to blood lead status. Given the observed relationships between certain environmental lead media and blood lead levels, the home environment was assumed to be a primary source of lead exposure in Milwaukee.
    - **No environmental lead measures were significant predictors of blood lead status *in New York*.** In fact, window sill and window trough dust lead loadings, number of surfaces with non-intact interior lead-based paint and play area soil lead went in the “wrong” direction in New York, with lead levels lower in homes with children with elevated blood lead levels. Although the children and homes in New York City were enrolled under the same study design as in Milwaukee, the home environment did not appear to be the primary source of lead exposure in New York. Further analysis of data collected from household questionnaires failed to identify likely sources of the children’s elevated blood lead status.

Table ES-2: Descriptive Statistics of Environmental Lead Media by Blood Lead Outcome (EBL/Non-EBL) and Site

Statistic	Site	Number of Units	Number w/EBLs	Lead Levels (Geometric Mean)[1]	
				EBL Homes	Non-EBL Homes
Floor Dust Lead (max) (μg/ft^2)	ML	64	31	45	23
	NY	69	36	8	7
Floor Dust Lead (mean) (μg/ft^2)	ML	64	31	24	12
	NY	69	36	4	4
Sill Dust Lead (max) (μg/ft^2)	ML	62	33	459	355
	NY	63	32	36	52
Sill Dust Lead (mean) (μg/ft^2)	ML	62	30	299	247
	NY	63	32	28	43
Trough Dust Lead (max) (μg/ft^2)	ML	59	27	6,749	5,171
	NY	55	28	239	422
Trough Dust Lead (mean) (μg/ft^2)	ML	59	31	9,601	6,895
	NY	55	28	282	483
Perimeter Soil Lead (ppm)	ML	56	30	2,918	1,298
	NY	17	32	965	457
Play Area Soil Lead (ppm)	ML	25	14	287	261
	NY	4	3	773	948
Water Lead (first draw) (ppb)	NY	64	31	3	3
	ML	69	36	4	3
Number of LBP Surfaces-Non-Intact (Exterior)	ML	64	31	6	8
	NY	69	36	1	1
Number of LBP Surfaces-Non-Intact (Interior)	ML	64	31	18	14
	NY	69	36	4	7

[1]For Number of LBP Surfaces-Non-Intact, the arithmetic mean values are presented and tested instead of the geometric mean values.

Table ES-3: Environmental Lead Media and Standards Examined

Media	Standards Examined					
Floor Dust Lead (mean) (μg/ft^2)	None	10	15	25	40	100
Sill Dust Lead (mean) (μg/ft^2)	None	125	250	500		
Trough Dust Lead (mean) (μg/ft^2)	None	800	5,000	10,000		
Perimeter Soil Lead (ppm)	None	400	1,200	2,000	5,000	
Play Area Soil Lead (ppm)	None	400	(1,200 was tested but no sample was above this level)			
Water Lead (first draw) (ppb)	None	5	10	15		
Number of LBP Surfaces-Non-Intact (Exterior)	None	1	5	10		
Number of LBP Surfaces-Non-Intact (Interior)	None	1	5	10		

Table ES-4: Standards for Optimal Protocols in Milwaukee

Protocol Media and Standards				Sensitivity (%)	Specificity (%)	P-value
Mean Floor Dust Pb ($\mu g/ft^2$)	Perimeter Soil Pb (ppm)	Play Area Soil Pb (ppm)	Water Pb (ppb)			
10	2,000	400	-	*100*	*36*	*<0.001*
10	2,000	-	-	*97*	*39*	*0.001*
10	-	400	-	*94*	*45*	*<0.001*
10	-	-	-	*90*	*48*	*0.001*
15	5,000	400	-	*77*	*58*	*0.006*
15	-	400	-	*77*	*58*	*0.006*
15	5,000	-	-	*84*	*55*	*0.002*
100	2,000	400	10	*77*	*58*	*0.006*
-	2,000	400	10	*74*	*61*	*0.006*
100	2,000	400	-	*74*	*61*	*0.006*
-	2,000	400	-	*71*	*64*	*0.007*
100	2,000	-	10	*71*	*64*	*0.007*
-	2,000	-	10	*68*	*67*	*0.012*
100	2,000	-	-	*68*	*67*	*0.012*
-	2,000	-	-	*65*	*70*	*0.012*
25	5,000	-	10	*61*	*73*	*0.011*
25	5,000	-	-	*58*	*76*	*0.010*
25	-	-	10	*48*	*79*	*0.035*
25	-	-	-	*45*	*82*	*0.030*
40	5,000	-	-	*39*	*85*	*0.048*

Alternative risk assessment protocols were tested using the data from Milwaukee. All permutations of the environmental lead media and standards listed in Table ES-3 were used as possible predictors of blood lead status (< or ≥ 10 μg/dl). Of the 92,190 protocols examined, 20 protocols were identified that were significant predictors of the blood lead status and optimized the performance characteristics (Table ES-4). Certain factors emerged from the results:

- **Floor dust lead loadings and perimeter soil lead concentrations were the two exposure sources most likely to be included** in the alternative protocols. These findings reinforce the earlier findings that these media were most predictive of the presence or absence of a child with an elevated blood lead level.
- **The optimal protocols included the complete range of mean floor dust lead loading standards tested (10-100 $\mu g/ft^2$).** They also included the higher levels of perimeter soil lead concentrations tested (2,000 and 5,000 ppm).
- **Some of the optimal protocols included play area soil lead (400 ppm) and water lead (10 ppb).** While the play area level matches the current standard, the water lead level is 5 ppb lower than the current action level.
- **Window sill and window trough dust lead and frequency of interior and exterior non-intact lead-based paint were not elements of the alternative protocols.** These results match the earlier findings that these media were not

predictive of homes in this study with or without a child with an elevated blood lead level.

Further analyses of data from Milwaukee explored optimal floor sampling locations:

- **The choice of floor sampling locations (Living Room, Kitchen, Bedroom, Bath and Unit Entry) and combination of locations had little difference on the ability to assess risk.** Almost all combinations of floor sampling locations were highly associated with the blood lead outcomes.
- **Floor samples taken from either *the room entry* or *central part of the floor* were generally more predictive of blood lead status** than those taken from under the window or a perimeter location.
- Although the HUD Guidelines recommend that risk assessors interview families to identify a child's "play room", **there was little difference between the predictive power of floor dust lead loadings from the "play room" versus the living room on blood lead status.** In fact, the p-values for the living room floor samples were equal or better to the play area floor samples suggesting that identifying the "play room" may not be necessary.
- **Although the choice of floor sampling locations do not appear to make a difference on the predictive power of the mean floor dust lead loadings, they may have an impact on selecting an optimal standard.** For example, the Unit Entry floor dust lead loadings were about twice as high as the interior floor dust lead loadings, so a mean floor sample result including the Unit Entry would perform differently against a given standard than a floor sample result without the Unit Entry.

Goal 3: To describe the contribution of friction and impact surfaces to floor dust lead loadings.

Risk assessors observed and recorded rubbing and/or binding on all painted doors and windows in the study. This information was included in statistical models to assess the influence of friction and impact surfaces on floor and window sill dust lead loadings.

The possible pathways of lead that are accounted for in the model included:

1. Window friction, window paint condition and window paint lead
2. Door friction, door paint condition and door paint lead
3. Lead paint (and condition) of the room
4. Exterior Lead Sources (Soil lead, other point sources)
5. Blow-in from the exterior
6. Track-in from the exterior

Results:

- **Assuming that window or door friction does produce dust lead, the results indicate that floor sampling would not be a good measure of rubbing or binding.** The interaction between the observation of rubbing/binding on doors and door paint lead and the interaction between the observation of rubbing/binding on windows and window paint lead were *not* significantly related to the floor dust lead loadings.

- **The analysis offers support to the hypothesis that window friction is a significant source of window sill dust lead even when window paint is intact.**
- **Dust lead loadings were higher on window sills where rubbing or binding was identified or window paint was not intact and dust lead loadings on those windows increased with the levels of paint lead.**

Goal 4: To assess the ability of the current HUD paint film quality classification system to predict rooms and dwellings that are likely to have elevated dust lead loadings.

When the grant for this study was awarded, HUD defined paint lead hazards as any lead-based paint in poor condition (Table ES-5). Since then, HUD and EPA issued regulations stating that all non-intact lead-based paint is a hazard. The findings for both definitions of paint deterioration are presented in the report.

Table ES-5: Categories of Paint Film Quality (HUD Guidelines Table 5.3)

	Total Area of Deteriorated Paint on Each Component		
Type of Building Component	Intact	Fair	Poor
Exterior components with large surface areas.	Entire surface is intact	Less than or equal to 10 square feet	More than 10 square feet
Interior components with large surface areas (walls, ceilings, floors, doors)	Entire surface is intact	Less than or equal to 2 square feet	More than 2 square feet
Interior and exterior components with small surface areas (window sills, baseboards, soffits, trim)	Entire surface is intact	Less than or equal to 10 percent of the total surface area of the component	More than 10 percent of the total surface area of the component

Results:

In Milwaukee,

- ***Non-intact* lead-based paint (LBP), but not *poor* LBP was a significant predictor of *floor dust lead loading*.**
- However, **the presence of *poor* LBP was a significant predictor of *blood lead status*, but not *non-intact* LBP.**
- When alternative numbers of LBP surfaces in poor condition were considered (i.e., 1, 2, 5, 10, 20 or 30), one or more LBP surfaces in poor condition had the greatest effect on the odds of having an elevated blood lead level. **A dwelling with at least one surface with poor LBP was 126% more likely to house an EBL child than a dwelling with no poor LBP.**

In Baltimore County and New York City,

- No measure of deteriorated LBP was a significant predictor of floor dust lead loading or blood lead status.

Across all three sites.

- **The results indicated that concerns about field implementation should not be a factor when determining the best method to identify deteriorated lead-based paint.** Pairs of risk assessors using the 3-level system (intact, fair, poor) to assess the condition of paint had a level of concurrence (67%) that was exactly the same as for the most basic test of deterioration (intact/non-intact).

Goal 5: To estimate the effect of dust lead measurement error on dust lead loadings.

In a subset of dwellings from all three sites in the study, side-by-side reliability samples were collected. Side-by-side dust samples in the home were used to estimate side-by-side variability for each sample type and site. All dust samples in the home (except additional side-by-side samples) were used to estimate between building variability and combined estimates of room/error variability for each sample type and site.

Using a combined estimate of room/error variability, observations were randomly generated from a log-normal distribution with these estimates of variability and various specified "true" average dust lead levels. This analysis was based on the assumption that there is some "true" unobservable dust lead level in a dwelling on a given surface type. Each dust sampling location was assumed to be equally representative of the true "unobservable" dust lead level in the dwelling on that surface type.

For the sample mean and maximum based on 1, 2 and 4 samples per dwelling, the following errors are evaluated:

(i) Type I (False Positive) Error = the probability that the sample statistic fails the dust lead standard given that the "true" lead level is below the standard.

(ii) Type II (False Negative) Error = the probability that the sample statistic passes the dust lead standard given that "true" lead level is above the standard.

The analyses generated Type I and Type II error estimates for each combination of site, surface type (uncarpeted and carpeted floors, window sills and window troughs), number of samples (1-5), and a prescribed set of dust lead standards. To simplify the presentation of these numerous results, a limited number of estimates are presented in Table ES-6. Estimates that represent significance levels 0.05, 0.10 and 0.20 are presented for each of the sites for floors and window sills. For comparative purposes, the effects of having only one or two samples collected in the dwelling are presented for floors in Milwaukee.

Results:

Table ES-6: Estimates of Upper and Lower Uncertainty Levels by Study Site, Surface Type, and Number of Samples

City	Surface	Standard	# of Samples	Confidence Level					
				95%		90%		80%	
				Lower	Upper	Lower	Upper	Lower	Upper
Milwaukee	Floor[1]	40	1	17	145	21	120	28	85
			2	20	99	23	83	30	67
			4	24	73	27	65	32	53
Balt. Co	Floor[1]	40	4	28	58	30	55	34	50
Milwaukee	Floor[1]	40	4	24	73	27	65	32	58
New York	Floor[1]	40	4	25	68	28	60	32	53
Balt. Co	Sill	250	4	120	630	150	540	190	440
Milwaukee	Sill	250	4	80	1630	115	1250	180	840
New York	Sill	250	4	105	940	140	740	175	560

[1]Central dust sampling location, carpets and bare floors combined

From the perspective of being most protective of a child's health, the upper uncertainty bounds in the table are of most interest. For example, the 80% upper uncertainty bound for a window sill dust sample of four rooms in New York City (see bottom row of table) was 560 μg/ft^2 when the window sill standard was 250 μg/ft^2. In practical terms, this means that if the "true" average lead level is 560 μg/ft^2, then there is a 20% chance that the sample mean will be below the standard of 250 μg/ft^2. These estimates are based on the good recovery rates achieved by the labs in this study. If the recovery rate is low, the variability effects could be compounded.

Sample variability may be just as harmful to the interests of a property owner and the affordability of housing. Using the sampling characteristics in the example above with the 80% lower uncertainty bound, approximately 20% of the time a home with a "true" lead level of 175 μg/ft^2 would fail the standard of 250 μg/ft^2. In other words, 20% of the time a dwelling with a true level 30% lower than the standard will fail the standard due to variability.

- **The results suggest that with additional samples in a dwelling, errors are less likely, but even with four samples the rate of error can be high.** For example, if 40 μg/ft^2 was established as a "health-based" standard for floors, these results suggest that that it may be appropriate to set an "action-level" below that standard to take into account the variability and be truly health protective.
- The high levels of variability for window sills (and window troughs) may help explain why these components were not predictive of blood lead outcomes. **With the level of variability, any sampling plan including window sill samples may have problems predicting risk.**