REPORT ON THE NATIONAL SURVEY

OF LEAD-BASED PAINT IN HOUSING

Appendix II: Analysis

This work was conducted under
HUD Contract Number HC-5848 and
EPA Contract Numbers 68-D9-0174, 68-D2-0139, and 68-D3-0011

June, 1995

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

List of Appendices

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Figures

1. INTRODUCTION

1.1 Background

Appendix II focuses on the statistical analysis of data collected during the National Survey of Lead-Based Paint, sponsored by the Department of Housing and Urban Development (HUD).

For background information on prior related studies, surveys and their limitations; lead in the environment and pathways between paint lead and blood lead; and the effects of lead-based paint abatement and refinishing on lead contamination refer to Appendix I of this report. Detailed descriptions of the National Survey design and methodology also are discussed in Appendix I. Therefore, they are not repeated here.

1.2 Report Organization

There are four chapters in Appendix II, including the introduction. Descriptions of each section are as follows:

- Chapter 2 provides prevalence data of lead-based paint in housing, and the magnitude of dust and soil lead contamination in residential environments. It also shows calculated amounts of lead in paint and the extent of priority hazards for private housing.

- Chapter 3 examines the quality of the data collected during the national survey of lead-based paint and the resulting quality of projected national estimates. In order to analyze this, the section addresses the effects of false negatives on the data quality; it analyzes non-response rates and other potential biases in the sample. The section also addresses the MAP/XRF performance in measuring lead content in paint; the laboratory measurement error inherent to all analytical techniques; and the effect of small dust sample weights on the findings.

- Lastly, Chapter 4 identifies the lessons learned from executing a large-scale national survey and gives recommendations for future field operations.

1.3 Reports Based on the National Survey

Analyses of data collected during the National Survey of Lead-Based Paint in Housing appear in three reports: the *Comprehensive and Workable Plan for the Abatement of Lead-Based Paint in Privately Owned Housing: Report to Congress*, prepared by HUD; the *Data Analysis of Lead in Soil*, prepared by EPA; and the *Analysis of Soil and Dust Samples for Lead (Pb)*, prepared for EPA.

1.4 Objectives of This Appendix

Using data collected during the National Survey of Lead-Based Paint in Housing, Appendix II provides details on the statistical procedures and ancillary data reduction techniques used to generate national prevalence estimates of lead sources in private and public housing.

Appendix II is intended for a technical audience who possess a general vocabulary and understanding of statistical analysis. However, every attempt is made to clearly define and discuss the analytical procedures used to analyze the national survey data. Readers are referred to the base report of this document for a summary of the significant findings stemming from these analyses.

2. NATIONAL ESTIMATES OF PREVALENCE

This chapter presents data on the prevalence of lead-based paint in private and public housing units built in the United States before 1980. Also included are the amount of surfaces covered with lead-based paint and the distributions of lead levels in residential paint, dust, and soil. Unless otherwise specified, all statistics presented in this chapter are weighted estimates calculated from the survey data using statistical sampling weights, as described in *Appendix I, Design and Methodology,* Section 6.6.

Throughout this chapter, the concepts of lead loading and lead concentration are used. Loading applies to the amounts of lead in paint and dust while concentration applies to lead amounts in soil. For paint, the loading concept refers to milligrams of lead per square centimeter of surface (mg/cm^2), for dust it refers to micrograms of lead per square foot ($\mu g/ft^2$). Soil is reported in parts per million (ppm), or the equivalent micrograms of lead per gram of soil (µg/g). In all three media, appropriate health-based standards have not yet been developed for housing. For ease of presentation we have used the current Federal action levels for paint, dust and soil. It is important to note that these action levels are not health-based. Because of this, the information presented below does not denote dwelling units with "safe" and "unsafe" lead conditions. More will be said on the origins of the action levels in the appropriate sections.

2.1 Private Housing

The private housing statistics presented below include revisions of data published in the *Comprehensive and Workable Plan for the Abatement of Lead-Based Paint in Privately Owned Housing: Report to Congress* (CWP Report). The statistics, both point and interval estimates, presented in the CWP did not incorporate corrections for the effects of paint, dust and soil measurement bias and variation, and the effects of incomplete testing of surfaces within housing units. The reasons for and the impact of the corrections are described in more detail in Chapter 3 of this volume. The findings presented in this section are all corrected for sources of error, and adjustments are incorporated into the results of the analyses in both their point and interval estimates.

2.1.1 Prevalence of Lead-Contaminated Paint, Dust, and Soil in Private Housing

Table 2-1: Estimated Number of Privately Owned Housing Units with Lead-Based Paint by Selected Characteristics - An estimated 64 million (77 million to 81 million)[1] or 83 percent of all privately occupied housing units in the United States built before 1980 have lead-based paint on some surface in or around the building. Housing units included in the above statistic have lead-based paint "somewhere" in one or more of the following locations: the interior; the exterior; or the common areas of multi-family structures (i.e., hallways, lobbies, mailrooms, laundry rooms, and playgrounds). A surface with lead contamination is defined here, and by HUD, as having a measured paint lead loading of 1.0 milligram of lead per square centimeter of surface (mg/cm^2) or greater.

The data collected during the National Survey suggests that older homes are more likely to have lead-based paint than newer homes. As displayed in Table 2-1, an estimated 76 percent (64 percent to 88 percent) of the housing units built *after* 1960 have lead-contaminated paint on their surfaces, but the percentage increases to 92 percent (84 percent to 100 percent) for houses built between 1940 and 1959. A slight counter-intuitive decrease in prevalence is evident in homes built between 1940 and 1959 compared to pre-1940 homes. Eighty-eight percent (79 - 97 percent) were contaminated. This decrease

[1]The numbers in parentheses are 95% confidence intervals. See Appendix II, Section 3.4.3 for the methodology used to compute the confidence intervals.

TABLE 2-1

ESTIMATED NUMBER OF PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT, BY SELECTED CHARACTERISTICS

(Paint Lead Concentration >= 1.0 mg/sq cm)

Characteristic	Total Occupied Housing Units (000) (1)	Housing Units With Lead-Based Paint Somewhere in Building		Number of Housing Units in Sample (2)
		Percent	Number (000)	
Total Occupied Housing Units Built Before 1980	77,177	83%	64,443	284
	100%	(9%)	(6,946)	
Construction Year:				
1960-1979	35,681	76%	27,275	120
	46%	(12%)	(4,282)	
1940-1959	20,476	92%	18,742	87
	27%	(8%)	(1,638)	
Before 1940	21,018	88%	18,424	77
	27%	(9%)	(1,892)	
Housing Type				
Single Family	66,418	85%	56,130	227
	86%	(9%)	(5,978)	
Multifamily	10,759	77%	8,308	57
	14%	(17%)	(1,829)	
One or More Children Under Age 7	13,912	89%	12,425	90
	18%	(9%)	(1,252)	
Census Region				
Northeast	16,963	86%	14,605	53
	22%	(13%)	(2,205)	
Midwest	19,848	91%	18,115	69
	26%	(10%)	(1,985)	
South	24,967	82%	20,393	116
	32%	(10%)	(2,497)	
West	15,399	73%	11,298	46
	20%	(18%)	(2,772)	
Owner-Occupied (2)	50,554	84%	42,516	179
		(9%)	(4,550)	
Market Value of Home				
Less than $40,000	11,885	92%	10,888	39
	24%	(12%)	(1,426)	
$40,000 to $79,999	19,401	90%	17,550	46
	38%	(10%)	(1,976)	
$80,000 to $149,999	11,863	68%	8,093	45
	23%	(18%)	(2,135)	
$150,000 and up	7,405	85%	6,276	42
	15%	(15%)	(1,111)	
Renter-Occupied (2)	24,734	82%	20,329	105
		(11%)	(2,721)	
Monthly Rent Payment				
Less than $400	16,339	85%	13,811	59
	66%	(14%)	(2,287)	
$400 and up	8,395	81%	6,822	40
	34%	(15%)	(1,259)	
Household Income (2)				
Less than $30,000	46,126	85%	39,032	156
	60%	(10%)	(4,613)	
$30,000 and up	31,048	81%	25,121	127
	40%	(11%)	(3,415)	

(1) Total units data are from the 1987 American Housing Survey.

(2) Some respondents did not respond to the questions on economic variables. Therefore, counts for disaggregation may not add to corresponding

Note: Numbers in parentheses are approximate half-widths of 95% confidence intervals for the estimated percents and numbers. For example, the confidence interval for the percent of housing units with some lead-based paint is 83% + / -9% or 74% to 92%.

is not statistically significant,[2] however, and is probably due to a combination of sampling and measurement variation.

Childhood lead poisoning is one of the most common and preventable public health concerns in our country today.[3] For this reason, EPA also examined the potential exposure of children to lead-based paint. An estimated 12 million (11 - 14 million) of the homes projected to have lead-contaminated paint are also occupied by families with children under the age of seven.

Finally, the differences among lead-based paint prevalence by type of housing, market value of the home, amount of monthly rent payment, and household income are not statistically significant.[1]

Table 2-2: Number and Percentage of Privately Owned Housing Units with Lead-Based Paint by Concentration, Construction Year, and Sample Location - Because there are no health-based standards for determining lead paint loadings in housing, the data was examined at varying cut-off thresholds used by different states and the Federal government. The lowest (which is to say most rigorous) loading standard, 0.7 mg/cm^2, corresponds to the definition of lead-based paint employed by the State of Maryland. This data is presented in the first column of Table 2-2. The next most rigorous loading (as used in Table 2-1) is 1.0 mg/cm^2, and corresponds to the Federal definition set by HUD. The third loading, 1.2 mg/cm^2, is the threshold used by the State of Massachusetts. Lastly, 2.0 mg/cm^2 is used. Although 2.0 mg/cm^2 is not a current standard, it is included as an upper-end threshold for informational purposes. Expanded tables (similar to Table 2-1) for 0.7, 1.2 and 2.0 mg/cm^2, by selected characteristics, are presented in Appendix A of this volume (Tables A-2, A-3, and A-4).

The analysis reveals that as the lead loading threshold increases, the number of newer homes meeting the criteria decreases faster than the older homes. In fact, the prevalence of newer homes (1960-1979) with lead-based paint "somewhere" in the building falls substantially from 82 percent to 48 percent as the threshold increases from 0.7 to 2.0 mg/cm^2. In contrast, the prevalence decrease for older homes (pre-1940) is virtually unchanged. Thus, the data suggests that older homes have higher paint lead loadings than newer homes. Since the amount of lead added to manufactured commercial residential paint declined from 1940 to 1980, this observation is not unreasonable. More discussion on the levels of lead in paint is presented in Section 2.1.3.

Table 2-3: Prevalence of Nonintact Lead-Based Paint by Location - Fourteen million or 19 percent of the private dwelling units in the United States are estimated to have *nonintact* (damaged) lead-based paint somewhere in the building. Nonintact lead-based paint is defined as at least five square feet of defective lead-based paint per unit. As is evident by the data presented in the table, it is estimated there is about twice the prevalence of nonintact lead-based paint on the exterior of private housing units (13 percent) than on the interior (7 percent).

---

[2]Test of equality of two independent proportions, in the presence of a complex sample design and measurement errors. The test statistic is $Z = (p_1 - p_2)/(\text{var } p_1 + \text{var } p_2)^{1/2}$, where var $p = (0.5 * \text{confidence interval width}/1.96)^2$. The proportions are significantly different from each other if $Z > 1.96$ (two-sided test); p_1 is significantly greater than p_2 if $Z > 1.645$ (one-sided test).

[3]CDC [1991]. *Preventing Lead Poisoning in Young Children*. U.S. Department of Health and Human Services, Public Health Service, Center for Disease Control.

TABLE 2-2

NUMBER AND PERCENTAGE OF PRIVATELY OWNED OCCUPIED HOUSING UNITS WITH LEAD IN PAINT BY LEAD CONCENTRATION, YEAR OF CONSTRUCTION, AND LOCATION OF LEAD-BASED PAINT

Location and Construction Year	Percentage of Homes Paint Lead Concentration (mg/sq cm)			
	>=0.7	>=1.0	>=1.2	>=2.0
Unit Interior	70%	63%	60%	45%
1960-1979	58%	49%	43%	23%
1940-1959	76%	69%	66%	54%
Built before 1940	84%	83%	82%	72%
Interior Common Areas	7%	5%	4%	3%
1960-1979	5%	4%	3%	1%
1940-1959	5%	5%	5%	5%
Built before 1940	11%	6%	5%	5%
Building Exterior	77%	73%	69%	59%
1960-1979	68%	61%	54%	36%
1940-1959	84%	81%	80%	76%
Built before 1940	86%	86%	86%	83%
Somewhere in Building	87%	83%	80%	68%
1960-1979	82%	76%	69%	48%
1940-1959	94%	92%	89%	83%
Built before 1940	88%	88%	88%	88%

Location and Construction Year	Number of Homes (000) Paint Lead Concentration (mg/sq cm)			
	>=0.7	>=1.0	>=1.2	>=2.0
Unit Interior	53,856	48,986	45,960	34,499
1960-1979	20,669	17,483	15,234	8,340
1940-1959	15,537	14,114	13,434	10,961
Built before 1940	17,653	17,392	17,289	15,200
Interior Common Areas	5,225	3,597	3,103	2,508
1960-1979	1,913	1,317	1,021	425
1940-1959	1,018	1,018	1,018	1,018
Built before 1940	2,300	1,261	1,066	1,066
Building Exterior	59,258	56,495	53,585	45,644
1960-1979	24,099	21,804	19,113	12,718
1940-1959	17,146	16,675	16,454	15,553
Built before 1940	18,018	18,018	18,018	17,373
Somewhere in Building	66,829	64,059	61,473	52,690
1960-1979	29,195	27,278	24,770	17,218
1940-1959	19,210	18,739	18,280	17,046
Built before 1940	18,426	18,426	18,426	18,426

TABLE 2-3

PREVALENCE OF NON-INTACT LEAD-BASED PAINT (LBP) BY LOCATION IN THE BUILDING - PRIVATELY OWNED OCCUPIED HOUSING UNITS

Location of Non-Intact LBP	**Occupied Housing Units With Non-Intact Lead-Based Paint**	
	Number (000)	**Percent (1)**
Building Interior (2)	5,596	7%
Interior Common Areas	249	0%
Building Exterior	9,657	13%
Both Interior and Exterior	1,083	1%
Somewhere in Building (3)	14,354	19%

(1) Base equals all 77,177,000 housing units built before 1980.

(2) "Building Interior" means that the only non-intact LBP is in the interior; there may be intact LBP on the exterior. "Building Exterior" has a similar meaning.

(3) A housing unit has some non-intact LBP if there are more than 5 sq. feet of damaged LBP somewhere. Similar definitions apply for interior and exterior LBP. It is therefore possible for a housing unit to have non-intact LBP somewhere in the building without having either non-intact exterior LBP or non-intact interior LBP (for example, a house with 3 sq. ft. of damaged interior LBP and 3 sq. ft. of damaged exterior LBP).

Note: There was no non-intact LBP in sampled playgrounds.

Table 2-4 shows the estimated number of housing units with dust lead loadings below and above the HUD Interim Guidelines are presented. Specifically, HUD's Interim Guidelines for dust lead levels are 200 µg/ft^2 for floors, 500 µg/ft^2 for window sills,[4] and 800 µg/ft^2 for window wells.[5] All of these loading standards are set for clearance purposes only (i.e., declaring an abated residence ready for re-occupancy after lead paint abatement). Even if HUD's dust standards were health-based, readers should be cautioned when comparing the National Survey dust lead results to the HUD guidelines because the former were collected with a vacuum sampling technique and the latter were developed for a wipe sampling technique. The sampling recoveries, i.e., percent of dust lead the sampler picks up from the surface, for both techniques are unknown, and probably differ substantially. There also is preliminary evidence that wipe samples tend to attain higher dust lead loadings than do vacuum samples.

Table 2-4: Rate of Occurrence of Privately Owned Occupied Housing Units With Dust Lead In Excess of the Federal Guidelines - This table shows that an estimated 17 percent of the privately owned housing units in the United States have dust lead loadings exceeding the relevant HUD guidelines, with the highest lead contamination found in window wells. This finding supports conclusions from other dust lead studies that suggest window wells typically have the highest dust lead loadings found in a home.[6]

Table 2-5: Dust Lead Loadings by Location, With or Without Lead-Based Paint - This table examines the prevalence of dust lead loadings above HUD's Interim Guidelines in homes that have lead-based paint above 1.0 mg/cm^2. Sixteen percent of the homes with no lead-based paint are projected to have dust lead loadings exceeding HUD's guidelines. In homes with lead-based paint on both interior and exterior surfaces, it is estimated that 29 percent will exceed the dust lead loading criteria. Although it appears from the table that the presence of lead-based paint contributes to higher dust lead loadings, there are additional sources of lead in the environment to account for dust lead in homes with no lead-based paint. To explore other potential sources of lead more thoroughly, EPA has further analyzed the National Survey data to examine the associations between dust lead, soil lead, paint lead, automobile emissions, and other factors in a report titled *Data Analysis of Lead in Soil*.[7] Interested readers are referred to this document for additional information.

Table 2-6: Association Between Lead in Interior Surface Dust and Lead-Based Paint Condition - This table presents the prevalence of high interior dust lead loadings in relation to the location and condition of lead-based paint. High loadings occur more frequently in housing units with nonintact lead-based paint than on other housing units. This holds true for units with interior nonintact lead-based paint, for units with exterior nonintact lead-based paint, and for units with nonintact lead-based paint somewhere on the building. In each case, more homes with nonintact lead-based paint have dust lead loading above HUD's guidelines than do homes with intact lead-based paint.

---

[4]Window sills are defined by HUD as the lower part of the window inside the room. In common carpentry terminology this is more commonly called a window stool.

[5]Window wells are defined by HUD as the bottom of a window between the screen and the glass. In common carpentry terminology this is more commonly called a window sill.

[6]Two examples are Landrigan, P, et al.: *Epidemic Lead Absorption Near an Ore Smelter: The Role of Particulate Lead*, N Eng J of Med 292: 123-9 (1975) and Farfel, M.R. and Chisolm, J.J.: *An Evaluation of Experimental Practices for Abatement of Residential Lead-Based Paint: Report on a Pilot Project*, Env Res 55: 199-212 (1991).

[7]*Data Analysis of Lead in Soil and Dust*. U.S. Environmental Protection Agency. EPA 747-R-93-011, September, 1993.

TABLE 2-4

RATE OF OCCURRENCE OF PRIVATELY OWNED OCCUPIED HOUSING UNITS WITH DUST LEAD IN EXCESS OF THE FEDERAL GUIDELINES

Interior Surface Dust Lead			
Location	**Federal Guideline (1) (ug/sq ft)**	**Number (000) of Housing Units Above Guideline (1)**	**Percent of Housing Units Above Guideline (1)**
Unit Interiors			
Somewhere	varies	13,317	17%
Window well	800	11,340	15%
Window sill (3)	500	2,684	3%
Floor (3)	200	808	1%
Window Only (2)	varies	12,508	16%
Interior Common Areas (3)	varies	1,249	2%

(1) HUD Interim Guidelines.
(2) Window includes window sill, window well or both.
(3) Categories with small sample sizes should be interpreted with caution.

TABLE 2-5

DUST LEAD LOADINGS BY LOCATION IN PRIVATELY OWNED OCCUPIED HOUSING UNITS WITH OR WITHOUT LEAD-BASED PAINT (LBP)

	Dust Within Guidelines (1)		Dust Exceeding Guidelines (1)	
Location of LBP	**Number (000)**	**Percent**	**Number (000)**	**Percent**
No LBP at All	10,681	84%	2,055	16%
Interior LBP Only	7,074	88%	925	12%
Both Interior and Exterior LBP	16,267	71%	6,529	29%
Some Interior LBP	38,855	80%	9,682	20%
Some Common Area LBP (2)	3,354	93%	242	7%
Some Exterior LBP	45,898	81%	10,598	19%
Some LBP	53,181	83%	11,262	17%

(1) HUD Interim Guidelines.
(2) Categories with small sample sizes should be interpreted with caution.

TABLE 2-6

ASSOCIATION BETWEEN LEAD IN INTERIOR SURFACE DUST AND LEAD-BASED PAINT (LBP) CONDITION FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS

Location of LBP	Condition of LBP	Dust Lead Within Guidelines (1)		Dust Lead Exceeds Guidelines (1)		Total	
		Number (000)	Percent	Number (000)	Percent	Number (000)	Percent
No LBP		10,681	83%	2,055	16%	12,736	100%
Interior LBP	Intact	35,592	82%	7,801	18%	43,393	100%
	Non-Intact (2)	3,715	66%	1,881	34%	5,596	100%
Exterior LBP	Intact	41,369	88%	5,470	12%	46,839	100%
	Non-Intact	4,529	47%	5,128	53%	9,657	100%
Somewhere	Intact	45,862	91%	4,679	9%	50,541	100%
	Non-Intact	7,319	53%	6,583	47%	13,902	100%

(1) "Within Guidelines" means that surface lead dust does not exceed 200 ug/sq ft on floors, 500 ug/sq ft on window sills, and 800 ug/sq ft on window wells. See HUD Interim Guidelines.

(2) A housing unit has non-intact interior LBP if there are more than 5 sq. ft. of damaged interior LBP. Similar definitions apply for "exterior" and "somewhere".

(3) Categories with small sample sizes should be interpreted with caution.

Table 2-7: Association Between Lead in Soil and Exterior Lead-Based Paint Condition - Presented are the estimated number of privately owned housing units with soil lead concentrations in excess of 500 ppm by the presence and condition of exterior lead-based paint. In order to examine associations between varying characteristics measured in this data set, 500 parts per million (ppm) of lead in soil was used to designate a threshold.[8]

The table suggests that housing units with exterior nonintact lead-based paint are more likely to have high (above 500 ppm) soil lead concentrations than are other housing units.

Table 2-8: Estimated Number of Housing Units in Selected Lead Hazard Categories - Chapter 2 of the *Comprehensive and Workable Plan for the Abatement of Lead-Based Paint in Privately Owned Housing: Report to Congress* discussed the risks to children exposed to lead-based paint, nonintact lead-based paint, dust lead, and soil lead. Table 2-8 displays the estimated prevalence of these hazards. However, as noted above, there are currently no health-based standards to make accurate hazard judgments. The standards used here are feasibility based interim guidelines which may not directly apply to the types of samples collected during the National Survey. Therefore, the priority hazard numbers presented in Table 2-8 should be treated qualitatively and should not be viewed as definitive values.

The table shows that of the 64 million homes with lead-based paint, an estimated 12 million are occupied with families who have children under the age of seven years old (also presented in Table 2-1). Of the 12 million, an estimated 4 million live in homes with dust lead loadings above HUD's guidelines or in homes with nonintact lead-based paint present. Both of these characteristics may pose significant hazards to children.

2.1.2 Amounts of Lead Paint in Private Housing

The previous section detailed the estimated prevalence of private dwelling units in the United States that have lead-based paint somewhere on their surfaces. This section presents national estimates of the square footage (interior and exterior) of surfaces covered with lead-based paint. Painted surfaces were sampled and quantified in the National Survey using a number of methods, depending on the component. The methodology is described in detail in *Appendix I, Design and Methodology,* Section 3.7.

Table 2-9: Amounts of Lead-Based Paint on Interior Surfaces by Component/Substrate - An estimated 29 billion square feet of painted interior surfaces are covered with lead-based paint. This represents 12 percent of the area of painted interior surfaces in pre-1980 privately-owned homes. On average, each home with interim lead-based paint has approximately 601 square feet of interior lead-based paint. Although only 9 percent of the paint on walls, ceilings, and floors is lead-based, those components account for 62 percent of all interior, surface area lead-based paint. Conversely, paint on "shelves/other" (shelves, cabinets, fireplaces, and closets) is much more likely to be lead-based, even though the total surface areas are much less. The separate breakdown by material substrate shows the Wood and Drywall with the largest amount of lead-based paint. This would be expected since Walls, Ceilings, and Floors are typically made from these substrate materials.

---

[8]This value was derived from the EPA's *Interim Guidance on Establishing Soil Lead Cleanup Levels at Superfund Sites,* September 7, 1989 (OSWER Directive # 9355.4-02).

TABLE 2-7

ASSOCIATION BETWEEN LEAD IN SOIL AND EXTERIOR LEAD-BASED PAINT CONDITION FOR PRIVATELY OWNED HOUSING UNITS (Numbers Represent Thousands of Occupied Housing Units)

Presence and Condition of Exterior Lead-Based Paint	Lead in Soil Somewhere			
	Within Guidelines (1)		Exceeding Guidelines (1)	
	Number	Percent	Number	Percent
No LBP	17,719	91%	1,660	9%
LBP Present, Intact	35,914	80%	9,215	20%
LBP Present, Non-Intact	3,815	44%	4,821	56%
Some Exterior LBP	39,729	74%	14,035	26%
Total	57,448	79%	15,695	21%

(1) The guideline is 500 ppm. See EPA, Interim Guidance.

TABLE 2-8

ESTIMATED NUMBER OF PRIVATELY OWNED OCCUPIED HOUSING UNITS IN SELECTED LEAD HAZARD CATEGORIES (Numbers Represent Thousands of Housing Units)

Lead Hazard Categories	All Occupied Housing Units	Housing Units With Children
Lead-based Paint Present (1)	64,443	12,427
Lead-based Paint Present and Paint Non-Intact (2)	14,354	3,321
Lead-based Paint Present and Lead Dust Present (3)	11,262	1,676
Lead-based Paint Present : Paint Non-Intact, OR Lead Dust Present	19,030	4,025

(1) Lead-based paint concentration of at least 1.0 mg/sq cm.

(2) At least 5 square feet of defective lead-based paint.

(3) Lead in dust exceeds 200 ug/ sq ft for floors, or 500 ug/sq ft for window sills, or 800 ug/sq ft for window wells.

TABLE 2-9

AMOUNTS OF LEAD-BASED PAINT (LBP) ON INTERIOR SURFACES
3Y PAINTED COMPONENT FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS
(LBP Concentration >= 1.0 mg/sq cm)

Component/Substrate	National Total Amount of LBP (millions of sq ft)	National Total Amount of LBP (percent of all paint on component)	Amount LBP Per Housing Unit With LBP (1) (square feet)
Components:			
Walls/ceiling/floor	18,148	9%	371
Metal component (2)	107	4%	2
Non-metal component (3)	7,172	24%	146
Shelves/other (4)	4,011	36%	82
Totals	29,437	12%	601
Substrates:			
Wood	11,672	26%	238
Metal (5)	141	4%	3
Drywall or plaster	17,113	9%	350
Concrete	225	3%	5
Undetermined	287	5%	6
Totals	29,437	12%	601

(1) Base equals the estimated 48,986,000 units with lead-based paint on interior surfaces.
(2) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(3) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(4) Includes shelves, cabinets, fireplace, and closets, on any substrate.
(5) Metal substrate refers to any architectural component on metal substrate.

Note: Because of rounding, totals may not be exactly the same as the sum of the numbers.

Table 2-10: Amounts of Lead-Based Paint on Exterior Surfaces by Component/Substrate - Table 2-10 presents data on the prevalence of exterior lead-based paint by architectural component and material substrate categories. An estimated 49 billion square feet of painted exterior surfaces are covered with lead-based paint. This represents 44 percent of the area of painted exterior surfaces in pre-1980 privately-owned homes. On average, each home with exterior lead-based paint has approximately 869 square feet of exterior lead-based paint. The data indicates there is more exterior surface area painted with lead-based paint than interior surface area (see Table 2-9 above). As expected, the component breakdown shows that the walls account for 78 percent of all exterior lead-based paint. Similarly, the breakdown by exterior material substrate shows wood with the largest amount of lead-based paint (63 percent).

Tables 2-11: Amounts of Lead-Based Paint on Interior Surfaces by Selected Characteristics - The table clearly shows that older homes, on average, have more lead-based paint on interior surfaces than do newer homes. The table also indicates that the distribution of lead-based paint in single and multifamily dwellings is approximately the same (11 percent - 12 percent).

Tables 2-12: Amounts of Lead-Based Paint on Exterior Surfaces by Selected Characteristics - The trend seen on interior surfaces (Table 2-11) is repeated on the exterior surfaces, but is even more pronounced. On average, 70 percent of all exterior painted surfaces on pre-1940 housing have lead-based paint.

2.1.3 Levels of Lead in Paint, Dust and Soil in Private Housing

Table 2-13: Arithmetic Mean Paint Lead Loadings by Characteristics - For both interior surfaces and exterior surfaces, a clear trend is apparent in paint lead loadings (mg/cm^2) from newer to older homes. Old lead-paint has more lead in it than newer lead-based paint. This is consistent with the paint manufacturing trends, where the amount of lead added to paint has dropped since the 1940's.

Two additional observations from the table are that Northeastern homes contain statistically significantly higher lead loadings on interior surfaces than on the rest of the nation's housing stock and exterior lead-based paint contains statistically significantly more lead than interior paint. Because these numbers are arithmetic means, however, they may be influenced by large values in the data and may give misleading results. For this reason, tables with geometric means (which approximate the median) are given in Appendix A of this volume and should be consulted for supplemental information. Table A-8 (geometric means) in Appendix A of this document shows the same general trend as does Table 2-13, but the differences between characteristics are not as dramatic.

Table 2-14: Arithmetic Mean Paint Lead Loadings by Component/Substrate and Construction Year - This table further breaks down Table 2-13's construction year category by component and characteristic. The data shows the same trend, i.e., an increase in paint lead area concentration from newest to oldest, especially on exterior walls. Geometric means of this data are also presented in Appendix A, Table A-9, which provides additional useful information. Again, the geometric means show the same general trends as does the arithmetic means.

Table 2-15: Unweighted Percentiles and Mean for Paint Lead Measurements by Sample Location - Arithmetic means, standard deviations, medians, and other selected percentiles are provided for the actual Scitec Metals Analysis Probe X-ray fluorescence device (MAP/XRF) measurements taken at privately owned housing units. The descriptive statistics (all unweighted) are grouped by sample location (interior of unit, exterior of unit, and all common areas). Two important findings outlined by the table are that a substantial number of samples had no detectable lead, and the highest measurements were recorded from exterior painted surfaces.

TABLE 2-10

AMOUNTS OF LEAD-BASED PAINT (LBP) ON EXTERIOR SURFACES BY PAINTED COMPONENT FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS (LBP Concentration >= 1.0 mg/sq cm)

Component/Substrate	**National Total Amount of LBP** (millions of sq ft)	**National Total Amount of LBP** (percent of all paint on component)	**Amount LBP Per Housing Unit With LBP (1) (square feet)**
Components:			
Walls	38,447	49%	681
Metal component (2)	403	8%	7
Non-metal component (3)	9,530	41%	169
Porches/other (4)	726	12%	13
Totals	49,106	44%	869
Substrates:			
Wood	30,930	46%	547
Metal (5)	5,486	33%	97
Drywall or plaster	1,969	46%	35
Concrete	7,426	40%	131
Undetermined	3,296	60%	58
Totals	49,106	44%	869

(1) Base equals the estimated 56,495,000 units with lead-based paint on exterior surfaces.
(2) Includes only metal windows, doors, soffit and facia, columns, and railings.
(3) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(4) Includes porches, balconies, stairs, etc., on any substrate.
(5) Metal substrate refers to any architectural component on a metal substrate including aluminum siding on exterior walls.

Note: Because of rounding, totals may not be exactly the same as the sum of the numbers.

TABLE 2-11

AMOUNTS OF LEAD-BASED PAINT (LBP) ON INTERIOR SURFACES BY SELECTED CHARACTERISTICS FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS (LBP Concentration >= 1.0 mg/sq cm)

Characteristic	National Total Amount of LBP (millions of sq ft)	National Total Amount of LBP (percent of all paint)	Amount LBP Per Housing Unit With LBP (square feet)	Number of Housing Units With LBP (000s)
Construction Year				
1960-1979	5,279	5%	302	17,483
1940-1959	8,247	13%	584	14,113
Before 1940	15,912	22%	915	17,392
Housing Type				
Single Family	27,001	12%	645	41,884
Multi-family	2,436	11%	343	7,104
One or More Children Under Age 7	4,290	10%	471	9,112
Total pre-1980 housing	29,437	12%	601	48,986

TABLE 2-12

AMOUNTS OF LEAD-BASED PAINT (LBP) ON EXTERIOR SURFACES BY SELECTED CHARACTERISTICS FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS (LBP Concentration >= 1.0 mg/sq cm)

Characteristic	National Total Amount of LBP (millions of sq ft)	National Total Amount of LBP (percent of all paint)	Amount LBP Per Housing Unit With LBP (square feet)	Number of Housing Units With LBP (000s)
Construction Year				
1960-1979	10,502	23%	482	21,803
1940-1959	12,635	41%	758	16,675
Before 1940	25,969	70%	1,441	18,018
Housing Type				
Single Family	46,216	45%	924	50,014
Multi-family	2,890	31%	446	6,482
One or More Children Under Age 7	6,127	26%	581	10,548
Total pre-1980 housing	49,106	44%	869	56,495

TABLE 2-13

ARITHMETIC MEAN PAINT LEAD LOADINGS IN PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980, BY SELECTED CHARACTERISTICS (Paint Lead Concentration >= 1.0 mg/sq cm)

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Total Occupied Housing Units Built Before 1980	0.7 (0.4 , 0.9)	1.9 (1.3 , 2.5)
Construction Year:		
1960-1979	0.3 (0.2 , 0.4)	0.6 (0.3 , 0.8)
1940-1959	0.5 (0.3 , 0.7)	1.5 (0.9 , 2.1)
Before 1940	1.4 (0.7 , 2.2)	4.6 (2.6 , 6.5)
Housing Type		
Single Family	0.7 (0.4 , 0.9)	2.0 (1.3 , 2.7)
Multifamily	0.4 (0.2 , 0.7)	1.0 (0.3 , 1.6)
One or More Children Under Age 7	0.7 (0.3 , 1.0)	1.6 (0.5 , 2.7)
Census Region		
Northeast	1.5 (0.6 , 2.5)	2.4 (1.4 , 3.4)
Midwest	0.5 (0.2 , 0.8)	2.1 (1.2 , 3.0)
South	0.3 (0.2 , 0.5)	1.7 (0.4 , 2.9)
West	0.4 (0.2 , 0.6)	1.2 (0.4 , 2.0)

Note: Numbers in parentheses are 95% confidence intervals for the respective arithmetic means.

TABLE 2-14

ARITHMETIC MEAN PAINT LEAD LOADINGS IN PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980, BY ARCHITECTURAL COMPONENT AND CONSTRUCTION YEAR (Paint Lead Concentration >= 1.0 mg/sq cm)

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Walls/ceilings/floor		
1960-1979	0.3 (0.2 , 0.4)	0.5 (0.2 , 0.8)
1940-1959	0.5 (0.2 , 0.7)	1.4 (0.6 , 2.3)
Before 1940	1.3 (0.5 , 2.1)	6.2 (2.6 , 9.9)
Metal (1)		
1960-1979	0.2 (0.1 , 0.4)	0.2 (0.0 , 0.3)
1940-1959	0.2 (0.1 , 0.4)	0.4 (0.1 , 0.7)
Before 1940	0.3 (0.2 , 0.4)	1.1 (0.1 , 2.2)
Non-metal (2)		
1960-1979	0.4 (0.0 , 0.9)	0.8 (0.4 , 1.3)
1940-1959	0.8 (0.4 , 1.3)	2.6 (1.2 , 3.9)
Before 1940	2.7 (1.7 , 3.7)	5.0 (2.7 , 7.2)
Other (3)		
1960-1979	0.1 (0.1 , 0.2)	0.2 (0.0 , 0.3)
1940-1959	0.2 (0.1 , 0.3)	0.8 (0.1 , 1.5)
Before 1940	1.4 (0.7 , 2.1)	2.1 (0.2 , 4.0)

Note: Numbers in parentheses are 95% confidence intervals for the respective arithmetic means.

Interior:
(1) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(2) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(3) Includes shelves, cabinets, fireplace, and closets, on any substrate.

Exterior:
(1) Includes only metal windows, doors, soffit and facia, columns, and railings.
(2) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(3) Includes porches, balconies, stairs, etc., on any substrate.

TABLE 2-15

PERCENTILES AND MEAN FOR PAINT LEAD (XRF) MEASUREMENTS FOR PRIVATE HOUSING UNITS BY SAMPLE LOCATION (UNWEIGHTED)

(Paint Lead Concentrations in mg/sq cm)

	Location		
	Interior	**Exterior**	**Common Areas**
Minimum	0.00	0.00	0.00
1%	0.00	0.00	0.00
5%	0.00	0.00	0.00
10%	0.00	0.00	0.04
25%	0.03	0.05	0.18
Median	0.19	0.42	0.70
75%	0.60	1.85	2.20
90%	1.66	5.81	5.54
95%	4.49	9.30	8.74
99%	10.18	27.71	19.69
Maximum	21.82	53.81	19.69
Mean	0.81	2.07	2.10
Standard Deviation	1.95	4.64	3.70
No. of Samples	4,273	1,047	218

Table 2-16: Percentiles for Lead in Soil Samples by Sample Location - Unweighted descriptive statistics, including arithmetic means, standard deviations and selected percentiles, are presented for lead concentrations in soil samples. Soil samples were collected from three locations on the property of each dwelling unit: a drip-line sample, collected at a point of about one foot from an exterior wall of the dwelling, potentially contaminated with deteriorated lead-based paint; an entrance sample collected near the most commonly used entrance, to measure the potential for tracked-in lead; and a remote sample, intended to measure background lead from sources other than lead-based paint. These locations are analyzed separately in this table, and overall statistics are presented.

The concentrations at the locations near the housing unit (drip-line and entrance) are similar to each other and are higher than the remote samples when the medians are compared. This would be expected if lead-based paint is contributing to soil lead contamination. The arithmetic means which, reflecting the skewness of the distribution, are substantially larger than the median, and reflect the same trends (higher concentrations near the structures). For more detailed analysis of the soil lead data collected during the National Survey, readers are referred to EPA's report entitled *Data Analysis of Lead in Soil*.

Figure 2-1: Box Plot Example - Dust lead loading data is presented in box plot form. For readers not familiar with boxplots, a descriptive discussion is given below.

Each boxplot shows a univariate data distribution, for example, the dust samples collected from a specific location (e.g., entrance floor). The box in the boxplot represents the middle 50 percent of the data; the bottom of the box gives the 25th percentile; the top gives the 75th percentile; and the horizontal line inside the box gives the median. The vertical lines extending from the top and bottom of the box reach to the largest and smallest observations, respectively, except for outliers. Outliers are plotted separately, as shown in Figure 2-1. Data sets approximating a normal distribution will produce a symmetrical boxplot, and fewer than 1 in 100 observations will be classified as unusual.

Figure 2-2: Boxplots of Dust Lead Loadings by Location - Because the data is approximately log normal (skewed to the right), it is plotted on a log scale. In doing so, the data approximates a normal distribution, reflected in the symmetry of the boxplots. From this display of the data, it is possible to visually compare lead loadings in all of the sample locations inside the dwellings simultaneously. Generally, the highest lead loadings are found in window wells, as discussed in Table 2-4.

2.1.4 Prevalence of Lead by Degree of Urbanization for Privately-Owned Housing

The objective of this section is to examine the association between the degree of urbanization and the prevalence of lead in paint, damaged paint, dust, and soil in privately owned housing.

To accomplish the objective of this analysis, the 150 census blocks in the 30 counties surveyed (see *Appendix I: Design and Methodology*) were each assigned to one of four urbanization categories based on their 1980 population as reported by the Bureau of Census. All housing units in a single census block were assigned the same urbanization category. The four urbanization categories are defined as follows:

TABLE 2-16

PERCENTILES AND MEAN FOR LEAD IN SOIL SAMPLES FROM PRIVATE HOUSING UNITS BY SAMPLE LOCATION (Soil Lead Concentration in ppm)

Statistic	All Locations (1)	Drip Line	Entrance	Remote
Minimum	1	1	3	1
1%	3	1	4	2
5%	6	6	10	5
10%	12	11	17	7
25%	23	23	30	19
Median	54	60	65	44
75%	152	201	201	119
90%	519	810	792	279
95%	1,188	1,476	1,376	545
99%	4,127	10,674	5,123	2,968
Maximum	22,974	22,974	6,828	6,951
Mean	324	448	260	204
Standard Deviation	1,207	1,766	894	691
Number of Samples	768	249	260	253

(1) Includes 6 samples taken on playgrounds which are not listed separately due to the small sample size.

FIGURE 2-1

BOXPLOT EXAMPLE

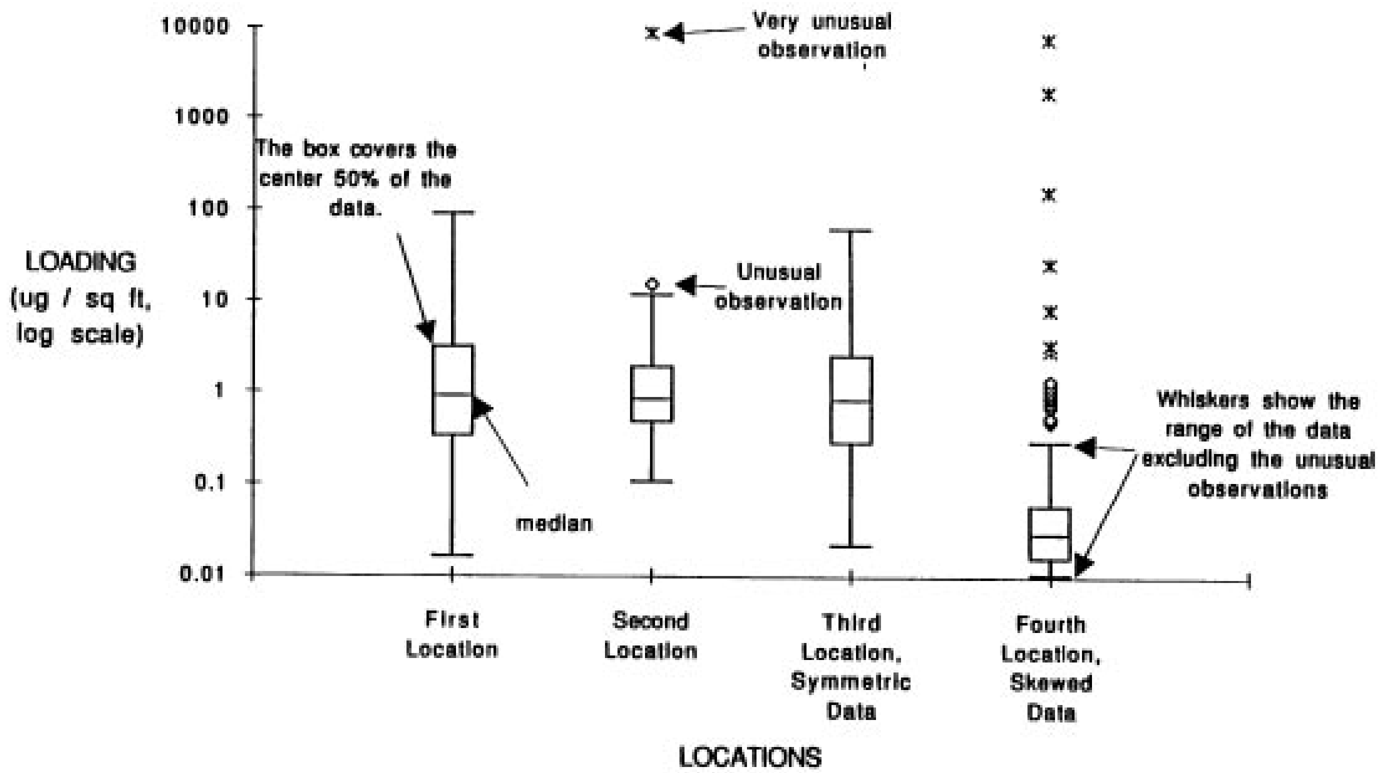
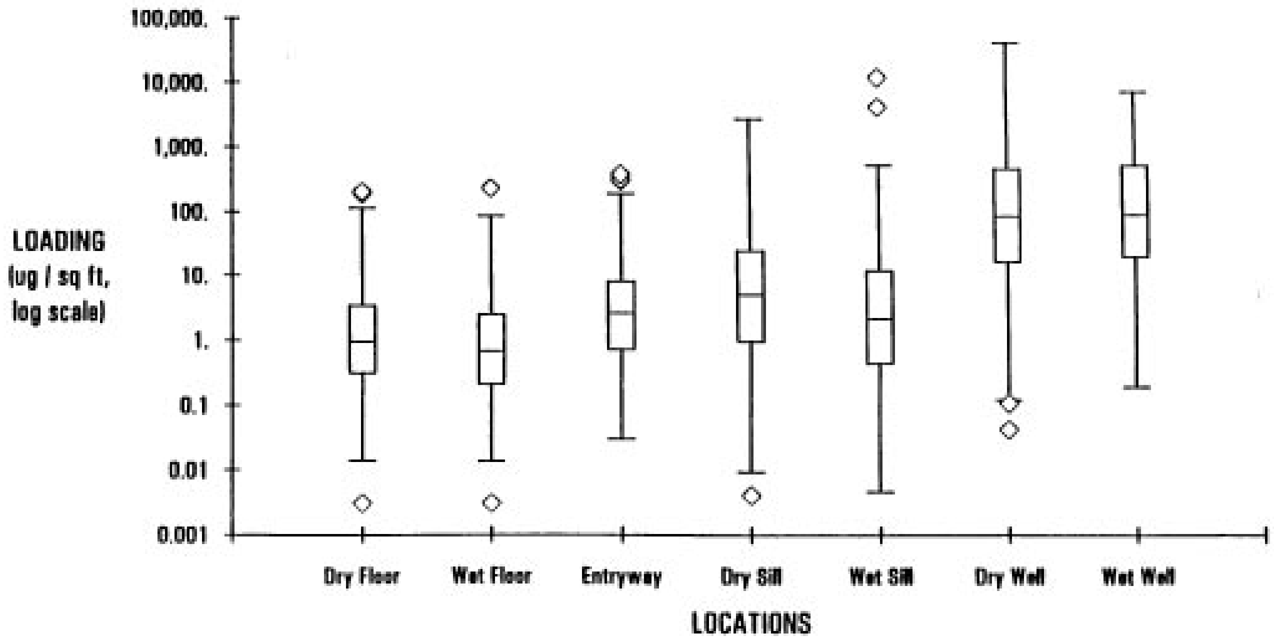


FIGURE 2-2

BOXPLOTS OF DUST LEAD LOADINGS
FOR PRIVATE HOUSING, INTERIOR LOCATIONS



- An area was considered a *large city* if it was located in a central city[9] of a Primary Metropolitan Statistical Area (PMSA)/Metropolitan Statistical Area (MSA)[10] with a population was over one million.

- An area was considered a *suburb of a large city* if it was located in a PMSA/MSA with a population of over one million, but was not located in a central city.

- An area was considered a *small city* if it was located in a PMSA/MSA with a population under one million.

- An area was considered a *rural area* if it was not located in a PMSA/MSA.

After each of the 284 sampled private housing units were assigned one of the four urbanization categories, EPA studied the relationship between the prevalence of lead in paint, damaged paint, dust, and soil and the degree of urbanization of the counties, along with region and construction year.

Tables 2-17 and 2-18 present the number of housing units in the sample, by degree of urbanization, construction year, and region. These tables show that subdividing the sample by degree of urbanization with construction year and region, the majority of cells have less than 30 housing units. This is important when interpreting the results presented in the following urbanization tables. For example, in Table 2-27, 88 percent of pre-1940 small city homes have soil lead above 500 ppm, while only 36 percent pre-1940 large city housing units have high soil lead concentrations. By examining table 2-17, it can be seen that the 88 percent was based on only 15 homes. Because 15 is a small number to scale-up results to national levels, no firm conclusions can be drawn.

Although only cells with 10 or more housing units are presented in Tables 2-19 through 2-28, caution is recommended when interpreting the results -- point estimates, confidence intervals, and tests of significance -- for cells with few housing units represented.

In many of the results discussed below, the difference between two proportions is tested for statistical significance at the .05 level. Thus the probability of a false positive (i.e., finding a significant difference when the proportions are the same) is .05 for any single comparison. Since in each table there are multiple comparisons that can be made, the probability of at least one false positive, or the probability of a false positive when comparing the largest observed proportion to the smallest, is greater than 0.05 (the exact value depends on the number of comparisons and on the correlations among the various proportions).

The analysis showed that there were no significant differences in the prevalence of housing units with nonintact paint, lead in dust, and lead in soil by the degree of urbanization. Differences were noticed in the prevalence of lead-based paint between housing units located in large cities and suburb areas versus housing units in small cities and rural areas.

---

[9]The largest city in each MSA is designated a "central city"; in addition there may be additional central cities if specified requirements are met. A more complete definition of "central city" can be obtained from the U.S. Office of Management and Budget.

[10]U.S. Office of Management and Budget current standards provide that an MSA is an area that includes at least one city with 50,000 or more inhabitants, **or** a Census Bureau-defined urbanized area of at least 50,000 inhabitants and a total MSA population of at least 100,000. OMB 1980 standards provide that within metropolitan complexes of 1 million or more population, separate component areas are defined if specified criteria are met. Such areas are designated PMSAs. More complete definitions of "MSAs" and "PMSAs" can be obtained from OMB.

TABLE 2-17

NUMBER OF PRIVATELY-OWNED HOUSING UNITS IN THE SAMPLE BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

	Construction Year			
Degree of Urbanization	**1960-1979**	**1940-1959**	**Pre-1940**	**Total**
Large City	28	33	32	93
Suburb of Large City	30	19	17	66
Small City	34	19	15	68
Rural Area	28	16	13	57
Total	120	87	77	284

TABLE 2-18

NUMBER OF PRIVATELY-OWNED HOUSING UNITS IN THE SAMPLE BY DEGREE OF URBANIZATION AND REGION

	Region				
Degree of Urbanization	**Northeast**	**Midwest**	**South**	**West**	**Total**
Large City	29	14	29	21	93
Suburb of Large City	16	13	18	19	66
Small City	8	10	44	6	68
Rural Area	0	32	25	0	57
Total	53	69	116	46	284

Note: -- "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" includes those housing units that are not located in a PMSA/MSA.

The analysis also showed that there were significant differences in the prevalence of all the lead characteristics by construction year and Census Region. In most cases, these differences were not dependent on the urbanization of the housing units.

Tables 2-19 through 2-28 present the results of the analysis. They are discussed below.

Tables 2-19 and 2-20: Estimated Number of Housing Units in the Nation by Degree of Urbanization, Construction Year, and Region - The tables show that the housing stock is relatively evenly distributed across the four urbanization categories. The percentages range from 21 percent for housing units located in rural areas to 28 percent for units located in large cities.

Tables 2-21 and 2-22: Percentage of Housing Units with Lead-Based Paint by Degree of Urbanization, Construction Year, and Region - Two sets of numbers are presented in each cell. The first number is the estimated percentage of housing units nationally with the lead characteristic. The numbers in parenthesis are the 95 percent confidence limits. Results are presented only for cells with 10 or more housing units in the sample. The tables show that the survey data indicated there is no statistical difference among the four urbanization categories in the percentage of housing units with lead-based paint.

Tables 2-23 and 2-24: Percentage of Housing Units with Damaged Lead-Based Paint by Degree of Urbanization, Construction Year, and Region - Table 2-23 shows that a higher percentage of pre-1940 housing units have damaged lead-based paint when compared to post-1940 housing units. Though a clear gradation in the percentage of housing units with damaged lead-based paint by construction year is evident by visual observation, there are no significant differences in the presence of damaged lead-based paint by urbanization.

Table 2-24 shows that a higher percentage of housing units located in large cities in the Northeast and Midwest have damaged paint than in housing units in the South and the West. This is probably due to the fact that much of the housing stock in the cities in the Northeast and Midwest is older than the housing stock in the cities in the South and West. For housing units in suburbs of large cities, there is no significant variation by region. Small sample sizes for the remaining two urbanization categories make any conclusions impossible.

Tables 2-25 and 2-26: Percentage of Housing Units with Dust Lead Loadings above Guidelines by Degree of Urbanization, Construction Year, and Region - Displayed are the results by urbanization with respect to high levels of dust lead loading (see note on HUD's Interim Dust Lead Guidelines in Section 2.1.1, Table 2-4). Analysis of the two extreme urbanization categories-large cities (25 percent) and rural areas (9 percent) shows a marginal difference at the .05 significance level.

Table 2-25 shows a clear increase in lead loadings by construction year for all four urbanization categories. Table 2-26 shows that the percentage of housing units from large cities and suburbs in the Northeast with high dust lead loadings is significantly higher than the other three regions.

Tables 2-27 and 2-28: Percentage of Housing Units with Soil Lead above Guidelines by Degree of Urbanization, Construction Year, and Region - Fewer dwelling units with soil lead levels above the Federal guidelines are found in large cities than in other areas (Table 2-27), with most of the difference evident in pre-1940 homes. It is difficult to explain this observation, especially since other

TABLE 2-19

ESTIMATED NUMBER ('000s) OF PRIVATELY-OWNED HOUSING UNITS IN THE NATION BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

Degree of Urbanization	Construction Year			Total	Percent
	1960-1979	1940-1959	Pre-1940		
Large City	7,938	6,613	7,038	21,589	28%
Suburb of Large City	9,602	5,305	4,826	19,733	26%
Small City	10,528	4,914	4,205	19,647	25%
Rural Area	7,618	3,641	4,951	16,210	21%
Total	35,686	20,473	21,020	77,179	100%

NOTE: Column totals are from the 1987 American Housing Survey. Other entries are estimates from the National Survey of Lead-Based Paint in Housing.

TABLE 2-20

ESTIMATED NUMBER ('000s) OF PRIVATELY-OWNED HOUSING UNITS IN THE NATION BY DEGREE OF URBANIZATION AND REGION

Degree of Urbanization	Region				Total
	Northeast	Midwest	South	West	
Large City	7,361	2,718	5,000	6,516	21,595
Suburb of Large City	6,599	3,034	3,763	6,342	19,738
Small City	- -	3,664	10,422	- -	19,647
Rural Area	- -	10,432	5,783	- -	16,215
Total	16,963	19,848	24,967	15,399	77,179

Notes:

1. "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" inlcudes those housing units that are not located in a PMSA/MSA.
2. A "--" represents a cell with less than 10 housing units.
3. Column totals are from the 1987 American Housing Survey. Other entries are estimates from the National Survey of Lead-Based Paint in Housing.

TABLE 2-21

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD-BASED PAINT BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

	Construction Year			
Degree of Urbanization	**1960-1979**	**1940-1959**	**Pre-1940**	**Total**
Large City	84% (63% - 97%)	85% (66% - 97%)	86% (67% - 98%)	85% (73% - 94%)
Suburb of Large City	78% (56% - 94%)	98% (83% - 100%)	67% (37% - 91%)	80% (65% - 92%)
Small City	75% (54% - 91%)	91% (69% - 100%)	100% (90% - 100%)	84% (70% - 94%)
Rural Area	69% (45% - 88%)	94% (72% - 100%)	100% (89% - 100%)	84% (69% - 95%)
Total	76% (63% - 87%)	92% (82% - 98%)	88% (76% - 96%)	83% (73% - 91%)

TABLE 2-22

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD-BASED PAINT BY DEGREE OF URBANIZATION AND REGION

	Region				
Degree of Urbanization	**Northeast**	**Midwest**	**South**	**West**	**Total**
Large City	86% (66% - 98%)	80% (49% - 98%)	97% (84% - 100%)	75% (49% - 94%)	85% (73% - 94%)
Suburb of Large City	79% (50% - 97%)	100% (89%- 100%)	79% (52% - 97%)	72% (44% - 93%)	80% (65% - 92%)
Small City	-- --	84% (49% - 100%)	83% (66% - 95%)	-- --	84% (70% - 94%)
Rural Area	-- --	94% (79% - 100%)	68% (43% - 88%)	-- --	84% (69% - 95%)
Total	86% (71% - 96%)	91% (79% - 98%)	82% (79% - 98%)	73% (54% - 88%)	83% (73% - 91%)

Note: -- "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" includes those housing units that are not located in a PMSA/MSA.

-- Numbers in parentheses are 95% confidence intervals for the estimated percents. For example, the 95% confidence interval for the percent of housing units with lead-based located in a large city and constructed between 1960-1979 is 62% - 98%.

-- A "--" represents a cell with less than 10 housing units in the sample.

TABLE 2-23

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH AT LEAST 5 SQ. FT. OF DAMAGED LEAD-BASED PAINT BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

Degree of Urbanization	Construction Year			Total
	1960-1979	1940-1959	Pre-1940	
Large City	4% (0% - 8%)	13% (2% - 31%)	47% (25% - 69%)	21% (10% - 35%)
Suburb of Large City	10% (1% - 28%)	5% (0% - 24%)	15% (1% - 41%)	10% (2% - 22%)
Small City	8% (0% - 24%)	13% (1% - 37%)	51% (21% - 81%)	19% (8% - 34%)
Rural Area	8% (0% - 25%)	38% (12% - 68%)	46% (16% - 78%)	27% (13% - 44%)
Total	8% (2% - 17%)	16% (6% - 29%)	40% (25% - 56%)	19% (11% - 29%)

TABLE 2-24

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH AT LEAST 5 SQ. FT. OF DAMAGED LEAD-BASED PAINT BY DEGREE OF URBANIZATION AND REGION

Degree of Urbanization	Region				Total
	Northeast	Midwest	South	West	
Large City	41% (19% - 65%)	26% (4% - 58%)	6% (0% - 22%)	8% (0% - 28%)	21% (10% - 35%)
Suburb of Large City	11% (0% - 36%)	15% (0% - 45%)	11% (0% - 35%)	4% (0% - 22%)	10% (2% - 22%)
Small City	-- --	88% (55% - 100%)	8% (1% - 22%)	-- --	19% (8% - 34%)
Rural Area	-- --	31% (13% - 53%)	19% (4% - 42%)	-- --	27% (13% - 44%)
Total	28% (13% - 46%)	33% (18% - 50%)	10% (3% - 20%)	5% (0% - 17%)	19% (11% - 29%)

Note: -- "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" includes those housing units that are not located in a PMSA/MSA.

-- Numbers in parentheses are 95% confidence intervals for the estimated percents. For example, the 95% confidence interval for the percent of housing units with lead-based located in a large city and constructed between 1960-1979 is 0% - 19%.

-- A "--" represents a cell with less than 10 housing units in the sample.

TABLE 2-25

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD IN DUST ABOVE GUIDELINES BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

	Construction Year			
Degree of Urbanization	**1960-1979**	**1940-1959**	**Pre-1940**	**Total**
Large City	6% (0 % - 33%)	12% (2% - 30%)	60% (37% - 81%)	25% (13% - 39%)
Suburb of Large City	0% (0% - 5%)	8% (0% - 29%)	60% (31% - 86%)	17% (6% - 32%)
Small City	4% (0% - 17%)	30% (8% - 58%)	44% (16% - 75%)	18% (7% - 33%)
Rural Area	2% (0% - 14%)	6% (0% - 28%)	22% (2% - 54%)	9% (2% - 22%)
Total	3% (0% - 10%)	14% (5% - 26%)	48% (25% - 64%)	18% (10% - 28%)

TABLE 2-26

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD IN DUST ABOVE GUIDELINES BY DEGREE OF URBANIZATION AND REGION

	Region				
Degree of Urbanization	**Northeast**	**Midwest**	**South**	**West**	**Total**
Large City	63% (39% - 84%)	19% (1% - 49%)	6% (0% - 22%)	2% (0% - 16%)	25% (13% - 39%)
Suburb of Large City	47% (19% - 76%)	4% (0% - 27%)	0% (0% - 15%)	3% (0% - 20%)	17% (6% - 32%)
Small City	-- --	27% (3% - 64%)	8% (1% - 22%)	4% (0% - 41%)	18% (7% - 33%)
Rural Area	-- --	12% (25% - 30%)	4% (0% - 19%)	-- --	9% (2% - 22%)
Total	55% (37% - 73%)	15% (5% - 29%)	5% (1% - 13%)	3% (0% - 13%)	18% (10% - 28%)

Note: -- "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" includes those housing units that are not located in a PMSA/MSA.

-- Numbers in parentheses are 95% confidence intervals for the estimated percents. For example, the 95% confidence interval for the percent of housing units with lead-based located in a large city and constructed between 1960-1979 is 0% - 23%.

-- A "--" represents a cell with less than 10 housing units in the sample.

TABLE 2-27

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD IN SOIL ABOVE GUIDELINES BY DEGREE OF URBANIZATION AND CONSTRUCTION YEAR

	Construction Year			
Degree of Urbanization	**1960-1979**	**1940-1959**	**Pre-1940**	**Total**
Large City	5% (0% - 20%)	5% (0% - 19%)	36% (16% - 39%)	15% (6% - 27%)
Suburb of Large City	10% (1% - 28%)	0% (0% - 8%)	76% (47% - 96%)	23% (10% - 39%)
Small City	1% (0% - 10%)	18% (2% - 44%)	88% (61% - 100%)	24% (17% - 47%)
Rural Area	0% (0% - 11%)	12% (0% - 38%)	55% (25% - 85%)	20% (8% - 36%)
Total	4% (0% - 11%)	8% (2% - 18%)	60% (44% - 75%)	20% (12% - 30%)

TABLE 2-28

PERCENTAGE OF PRIVATELY-OWNED HOUSING UNITS WITH LEAD IN SOIL ABOVE GUIDELINES BY DEGREE OF URBANIZATION AND REGION

	Region				
Degree of Urbanization	**Northeast**	**Midwest**	**South**	**West**	**Total**
Large City	24% (7% - 46%)	4% (0% - 26%)	6% (0% - 22%)	17% (2% - 41%)	15% (6% - 27%)
Suburb of Large City	47% (19% - 76%)	20% (2% - 52%)	0% (0% - 15%)	16% (2% - 41%)	23% (10% - 39%)
Small City	23% (1% - 46%)	60% (23% - 91%)	16% (4% - 33%)	4% (0% - 41%)	24% (17% - 47%)
Rural Area	-- --	27% (10% - 49%)	7% (0% - 25%)	-- --	20% (8% - 36%)
Total	33% (17% - 51%)	43% (27% - 66%)	11% (4% - 21%)	15% (4% - 31%)	20% (12% - 30%)

Note: -- "Large city" includes housing units located in central cities where the PMSA/MSA population is over 1 million. "Suburb of large city" includes housing units located outside central cities but in PMSAs/MSAs with population over 1 million. "Small city" includes housing units located in PMSAs/MSAs with population under 1 million. "Rural Area" includes those housing units that are not located in a PMSA/MSA.

-- Numbers in parentheses are 95% confidence intervals for the estimated percents. For example, the 95% confidence interval for the percent of housing units with lead-based located in a large city and constructed between 1960-1979 is 0% - 21%.

-- A "--" represents a cell with less than 10 housing units in the sample.

of studies cite soil in large cities with the highest levels of lead in soil.[11] This finding may be an artifact the survey methodology, in which soil samples were taken by core samples. No surface scrapings were taken from paved surfaces. Many of the sampled homes in large cities had no unpaved surfaces suitable for core sampling.

Table 2-28, which displays the estimated percentage of housing units with lead in soil by urbanization and region shows that the south has a significantly lower percentage of housing units with lead in soil when compared to the Northeast and Midwest. Again, this may be related to the relative ages of the housing stock in the different regions of the country.

2.2 Public Housing

Below are the results of the public housing data analysis. The presentation parallels the private housing presentation.

2.2.1 Prevalence of Lead-Contaminated Paint, Dust and Soil in Public Housing

Table 2-29: Estimated Number and Percent of Housing Units with Lead-Based Paint by Selected Characteristics - An estimated 86 percent (i.e., about 782,000 units) of all public housing units in the United States built before 1980 have lead-based paint somewhere in the building.

"Somewhere" refers to lead-based paint on one or more of the following locations: the interior of the unit; the exterior walls; or the common areas of multi-family structures. As with private housing, a surface with lead-contamination is defined here, and by HUD, as having a measured paint lead loading of 1.0 mg/cm^2 or greater.

Although the data collected during the National Survey suggests that older public units are more likely to have lead-based paint than newer units, the differences are not as great as those predicted for the private dwelling units (see Table 2-1). Because the sample sizes are small, (i.e., only 97 were units were sampled) and stratified by construction year, conclusions may not truly represent all public housing units in the United States and readers are cautioned in their interpretation.

Table 2-30: Number and Percentage of Housing Units with Lead-Based Paint by Concentration and Sample Location - Table 2-30 shows the impact of four different paint lead loading thresholds on the prevalence of lead-based paint in public housing. Included thresholds are the Maryland, Federal, and Massachusetts standards. Also included are dwelling units with concentrations of lead at 2.0 mg/cm^2 or higher. Similar to Table 2-2 for private housing, modifying the threshold concentration substantially changes the number of dwelling units characterized as having lead-based paint on interior painted surfaces. However, the different thresholds have less of an effect on the prevalence of exterior lead-based paint than on interior paint.

Prevalence of Nonintact Lead-Based Paint - Less than 10 public housing units in the sample had more than five square feet of nonintact lead-based paint. Because the sample size was small, projecting meaningful national estimates and analyzing relationships between nonintact lead-based paint

---

[11]One such study reporting higher soil lead levels in inner cities was reported at the Trace Substances in Environmental Health-XXV conference held in Columbia, Missouri, May 20-23, 1991 entitled *Dust Control as a Means of Reducing Inner-City Childhood Pb Exposure* by H.W. Mielke et al.

TABLE 2-29

ESTIMATED NUMBER AND PERCENT OF PUBLIC HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT, BY SELECTED CHARACTERISTICS

(Paint Lead Concentration >= 1.0 mg/sq cm)

Characteristic		Total Public Housing Units (000)	Housing Units With Lead-Based Paint Somewhere in Building: Percent	Housing Units With Lead-Based Paint Somewhere in Building: Number (000)	Number of Housing Units in Sample
Total Public Housing Units Built Before 1980		910	86%	782	97
		100%	(78% - 94%)	(705 - 858)	
Construction Year:	1960-1979	455	79%	359	43
		50%	(66% - 92%)	(299 - 419)	
	1950-1959	273	90%	246	24
		30%	(77% - 100%)	(209 - 273)	
	Before 1950	182	97%	177	30
		20%	(88% - 100%)	(160 - 182)	

Notes: (1) Numbers in parentheses are 95% confidence intervals for the estimated percents and numbers.
(2) Categories with small sample sizes should be interpreted with caution.

TABLE 2-30

NUMBER AND PERCENTAGE OF PUBLIC HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT BY LEAD CONCENTRATION AND LOCATION OF LEAD-BASED PAINT

	Percentage of Homes			
	Paint Lead Concentration (mg/sq cm)			
Location	**>=0.7**	**>=1.0**	**>=1.2**	**>=2.0**
Unit Interior	80%	75%	70%	46%
Interior Common Areas	44%	38%	36%	31%
Building Exterior	71%	68%	67%	59%
Playgrounds	12%	12%	11%	--
Somewhere in Building	90%	86%	85%	77%

	Number of Homes (000)			
	Paint Lead Concentration (mg/sq cm)			
Location	**>=0.7**	**>=1.0**	**>=1.2**	**>=2.0**
Unit Interior	730	685	633	417
Interior Common Areas	401	347	331	279
Building Exterior	647	623	612	540
Playgrounds	112	112	99	--
Somewhere in Building	821	782	774	697

Note:
A "--" indicates that there were less than 10 housing units in the sample with the lead characteristic.

with dust and soil lead was not possible. Furthermore, the dust lead data for the public housing units was suspect because a large number of vacant apartments were sampled. Since dust lead loadings are a function of total dust present, and because the unoccupied units were thoroughly cleaned prior to the sampling visits, the representativeness of the data is unknown. Thus, scaling the results to project national estimates is not advisable.

2.2.2 Amounts of Lead Paint in Public Housing

The previous section reported the prevalence of public housing units in the United States that have lead-based paint somewhere on their surfaces. This section presents national estimates of how much surface area is covered with lead-based paint.

Table 2-31: Amounts of Lead-Based Paint on Interior Surfaces by Component/Substrate - Table 2-31 presents data on the prevalence of interior lead-based paint by architectural component and material substrate categories. An estimated 12 percent or 252 million square feet of all painted interior surfaces are covered with lead-based paint. Twelve percent is also estimated for private dwelling units (Table 2-9). On average, each public housing unit with lead-based paint has approximately 447 square feet of interior lead-based paint. Although painted walls, ceilings, and floors account for more area, painted metal components are much more likely to be lead-based. The component breakdown shows that the "Walls/ceiling/floors" component has 193 million square feet of lead-based paint accounting for approximately 76 percent of all interior lead-based paint. However, only 10 percent of the paint on walls, ceilings, and floors is lead-based. Paint on "metal components" (e.g., radiators, doors, air heat vents) is much more likely to be lead-based, even though the total surface areas covered with lead-based paint are far less. The metal component only has one tenth the area of lead-based paint, but this represents 33 percent of all painted metal components. The separate breakdown by material substrate shows the "Drywall" category with the largest amount of lead-based paint with 136 million square feet, or 54 percent of all interior lead-based paint.

Table 2-32: Amounts of Lead-Based Paint on Exterior Surfaces by Component/Substrate - Table 2-32 presents data on the prevalence of exterior lead-based paint by architectural component and material substrate categories. The data indicates there is less exterior surface area painted with lead-based paint than interior surface area. This is the opposite of private housing findings (Table 2-10). An estimated 83 million square feet of lead-based paint covers exterior surfaces (7 percent of all exterior painted surfaces on public housing), with an average of 214 square feet per public housing unit. The component breakdown shows that the non-metal components have 44 million square feet of lead-based paint accounting for 53 percent of all exterior lead-based paint. Non-metal components include such items as trim, window sills, doors, soffit, and fascia. The breakdown by material substrate shows that wood has the largest amount of lead-based paint with 40 million square feet, or 48 percent of all exterior lead-based paint on public housing.

2.2.3 Levels of Lead in Paint, Dust and Soil in Public Housing

Table 2-33: Arithmetic Mean Paint Lead Loadings by Characteristics - For exterior painted surfaces, a clear trend is apparent in paint lead loadings (mg/cm^2) from newer to older public housing units. Old lead-paint has more lead in it than newer lead-based paint. This is consistent with the paint manufacturing trends, where the amount of lead added to paint has dropped since the 1940's. For more information, tables with geometric means (which approximate the median) are given in the Appendix B of this document (Table B-4). Although geometric means are the same for interior and exterior surfaces, the arithmetic means indicate exterior surfaces have higher paint lead loadings. This reflects very high values measured on a few exterior surfaces which distort the arithmetic means (see Table 2-35).

TABLE 2-31

AMOUNTS OF LEAD-BASED PAINT (LBP) ON INTERIOR SURFACES BY ARCHITECTURAL COMPONENT AND SUBSTRATE FOR PUBLIC HOUSING UNITS

(LBP Concentration >= 1.0 mg/sq cm)

Component/Substrate	National Total Amount of LBP (millions of sq ft)	National Total Amount of LBP (percent of all paint on component/substrate)	Amount LBP (1) Per Housing Unit With Lead-Based Paint (square feet)
Components:			
Walls/ceiling/floor	193	10%	282
Metal component (2)	22	31%	32
Non-metal component (3)	32	19%	47
Shelves/other (4)	4	8%	6
Totals	252	12%	367
Substrates:			
Wood	35	18%	51
Metal (5)	23	29%	34
Drywall or plaster	136	9%	198
Concrete	51	21%	74
Undetermined	7	8%	10
Totals	252	12%	367

(1) Base equals the estimated 685,000 units with lead-based paint on interior surfaces.
(2) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(3) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(4) Includes shelves, cabinets, fireplace, and closets, on any substrate.
(5) Metal substrate refers to any architectural component on metal substrate.

Note: Because of rounding, totals may not be exactly the same as the sums of the numbers.

TABLE 2-32

AMOUNTS OF LEAD-BASED PAINT (LBP) ON EXTERIOR SURFACES BY ARCHITECTURAL COMPONENT AND SUBSTRATE FOR PUBLIC HOUSING UNITS

(LBP Concentration >= 1.0 mg/sq cm)

Component/Substrate	National Total Amount of LBP		Amount LBP (1) Per Housing Unit With Lead-Based Paint (square feet)
	(millions of sq ft)	(percent of all paint on component/substrate)	
Components:			
Walls	8	1%	13
Metal component (2)	28	16%	45
Non-metal component (3)	44	15%	71
Other (4)	3	3%	4
Totals	83	7%	133
Substrates:			
Wood	40	12%	64
Metal (5)	28	14%	45
Drywall or plaster	0	0%	0
Concrete	8	3%	13
Undetermined	7	4%	11
Totals	83	7%	133

(1) Base equals the estimated 622,860 units with lead-based paint on exterior surfaces.
(2) Includes only metal windows, doors, soffit and facia, columns, and railings.
(3) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(4) Includes porches, balconies, stairs, etc., on any substrate.
(5) Metal substrate refers to any architectural component on a metal substrate including aluminum siding on exterior walls.

Note: Because of rounding, totals may not be exactly the same as the sums of the numbers.

Table 2-34: Arithmetic Mean Paint Lead Loadings by Component/Substrate and Construction Year - This table further breaks down Table 2-33's construction year category by component/characteristic. The data shows the same trend -- an increase in paint lead loadings from newest to oldest. For additional information on the geometric means of this data, refer to Appendix B, Table B-5.

Table 2-35: Percentiles and Mean for MAP/XRF Measurement Statistics by Sample Location - Arithmetic means, standard deviations, and selected percentiles are provided for the actual MAP/XRF measurements taken at public housing units. The descriptive statistics are grouped by sample location (interior of unit, exterior of unit, and all common areas). Two important findings outlined by the table are that a substantial number of samples had no detectable lead and the highest measurements were recorded from exterior painted surfaces.

Table 2-36: Percentiles and Mean for Soil Lead Measurement Statistics by Sample Location - Descriptive statistics, including arithmetic means, standard deviations and selected percentiles, are presented for lead concentrations in soil samples collected during the National Survey. Soil samples came from three locations on the property of each dwelling: a drip-line sample near an exterior wall of the dwelling, potentially contaminated with deteriorated lead-based paint; an entrance sample collected near the most commonly used entrance, to measure the potential for track-in lead; and a remote sample, intended to measure background lead from sources other than lead-based paint. These locations are analyzed separately in this table, and overall statistics are presented.

Arithmetic mean concentrations are generally lower in the public housing soil samples than in soil samples collected at private housing sites (see Table 2-16). One evident cause is that the public housing distribution is tighter, without extremely high values. In the private housing samples, concentrations ranged from 1 to 22,000 ppm, and the data is very skewed to the right. These large values increase the private housing arithmetic means. By examining the medians, however, the public and private housing appear more similar.

A major problem with the public housing soil lead data collected during the National Survey is that most units (about 70 percent) did not have exposed soil present to collect samples; most were surrounded by pavement. This is reflected in the small number of samples collected at each location (see Table 2-36). Therefore, the representativeness of public housing units with soil nearby to all public housing units is unknown. Since many public housing units are in inner cities, and soil in inner cities is usually cited as having the highest average lead concentrations (although this was not observed in the private housing data), it would be expected that the soil samples collected from public housing should be higher in lead than from private housing. This was not the case, however, and the small sample sizes and/or the unknown representativeness of the data could be the reason.

Another conclusion from the table is that the three sampling locations for public housing are much more similar to each other than the three locations for private housing.

Figure 2-3: Boxplots of Dust Lead Loadings by Location - See Figure 2-1 for examples of boxplots. Because the data is approximately log normal (skewed to the right), it is plotted on a log scale. In doing so, the data approximates a normal distribution, reflected in the symmetry of the boxplots. From this display of the data it is possible to visually compare lead loadings from all of the sample locations inside dwellings, simultaneously. As with the private housing data, the highest lead loadings are generally found in window wells.

TABLE 2-33

ARITHMETIC MEAN PAINT LEAD LOADINGS IN PUBLIC HOUSING UNITS BUILT BEFORE 1980, BY SELECTED CHARACTERISTICS

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Total Public Housing Units Built Before 1980	0.4 (0.3 , 0.5)	1.2 (0.4 , 1.5)
Construction Year:		
1960-1979	0.4 (0.2 , 0.5)	0.3 (0.1 , 0.5)
1950-1959	0.4 (0.2 , 0.5)	1.1 (0.1 , 2.1)
Before 1950	0.5 (0.3 , 0.8)	2.3 (0.5 , 4.1)

TABLE 2-34

ARITHMETIC MEAN PAINT LEAD LOADINGS BY PAINTED COMPONENT AND CONSTRUCTION YEAR FOR PUBLIC HOUSING UNITS

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Walls/ceilings/floor		
1960-1979	0.4 (0.2 , 0.5)	0.3 (0.1 , 0.5)
1950-1959	0.3 (0.1 , 0.5)	0.9 (-0.1 , 1.3)
Before 1950	0.5 (0.3 , 0.8)	1.0 (-1.1 , 6.9)
Metal		
1960-1979	0.5 (0.1 , 0.8)	0.3 (0.0 , 0.6)
1950-1959	1.3 (0.5 , 2.0)	0.6 (-0.1 , 1.3)
Before 1950	1.0 (0.6 , 1.4)	2.9 (-1.1 , 6.9)
Non-metal		
1960-1979	0.4 (0.2 , 0.7)	0.3 (0.0 , 0.6)
1950-1959	0.4 (0.2 , 0.7)	2.8 (-0.3 , 5.8)
Before 1950	0.7 (0.2 , 1.3)	5.4 (1.6 , 9.3)
Other		
1960-1979	0.2 (0.1 , 0.3)	1.1 (-0.2 , 2.4)
1950-1959	0.4 (0.0 , 0.9)	0.3 (-0.1 , 0.6)
Before 1950	0.3 (0.0 , 0.7)	1.4 (-0.7 , 3.5)

TABLE 2-35

PERCENTILES AND MEAN FOR XRF MEASUREMENTS FOR PUBLIC HOUSING UNITS BY SAMPLE LOCATION (UNWEIGHTED)

(Paint Lead Concentrations in mg/sq cm)

	Location		
	Interior	**Exterior**	**Common Areas**
Minimum	0.00	0.00	0.00
1%	0.00	0.00	0.00
5%	0.00	0.00	0.00
10%	0.00	0.00	0.00
25%	0.05	0.00	0.06
Median	0.21	0.14	0.31
75%	0.68	0.72	1.08
90%	1.74	3.44	2.42
95%	2.64	7.18	4.58
99%	4.78	22.52	23.28
Maximum	12.76	34.50	23.96
Mean	0.58	1.22	1.17
Std. Dev.	2.81	9.64	7.898
No. of Samples	1,731	267	553

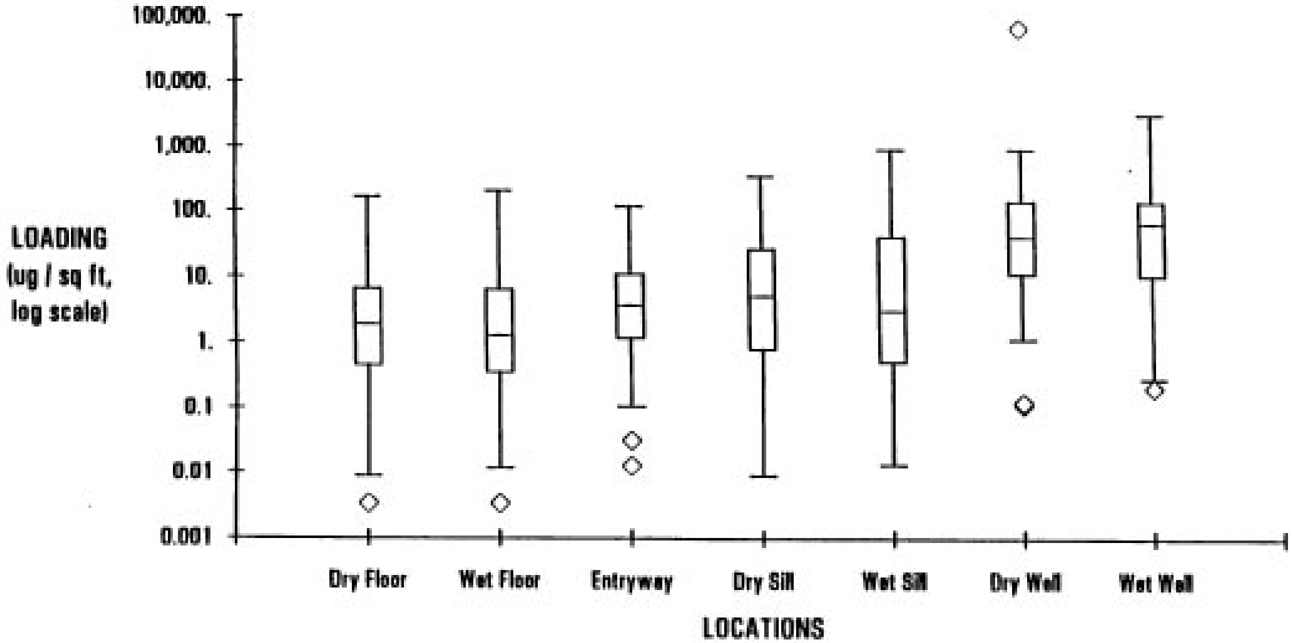
TABLE 2-36

PERCENTILES AND MEAN FOR LEAD IN SOIL SAMPLES FROM PUBLIC HOUSING UNITS BY SAMPLE LOCATION
(Soil Lead Concentrations in ppm)

	All Locations	Drip Line	Entrance	Remote
Mean	92	104	112	79
Minimum	5	8	11	6
1%	5	8	11	6
5%	9	9	14	9
10%	12	15	15	11
25%	20	25	23	25
Median	39	47	49	49
75%	126	186	167	101
90%	206	438	265	219
95%	424	483	499	243
99%	753	527	872	615
Maximum	753	527	872	615
No. of Samples	89	28	26	29

FIGURE 2-3

BOXPLOTS OF DUST LEAD LOADINGS FOR PUBLIC HOUSING, INTERIOR LOCATIONS



3. SOURCES OF ERROR IN THE NATIONAL SURVEY DATA

An evaluation of data quality is necessary in order to assess the utility of the survey data for lead research and policy development. It is the errors in the data that generally determine data quality. In this context, "error" refers to deviations of obtained survey results from those that are true reflections of the population. These errors are a function of the processes that generated the data at the various stages of the survey: sample design, sample selection, field sampling and data collection, laboratory measurement, data processing, and data analysis.

This chapter describes the quality of the data collected from the national survey of lead-based paint and the statistical techniques used to identify and measure error. Much of the chapter focuses on errors due to the MAP/XRF equipment used to detect and quantify the amount of lead in painted surfaces. As will be seen, it is possible to adjust the data to "correct" for these errors. Section 3.2 reports a detailed analysis of nonresponse and other potential sampling biases of estimates for both private and public housing. Section 3.3 examines measurement errors for MAP/XRF devices. Section 3.4 presents a bias and variable analysis of the MAP/XRF measurement errors on classifying homes with lead-based paint. Section 3.5 explores classification error due to incomplete sampling of painted surfaces within dwelling units. Finally, Section 3.6 looks at the quality of the laboratory measurements and the effects of small dust sample weights on the findings.

3.1 Statistical Concepts and Terminology

An *error* is simply the difference between the sample estimate and the population parameter that we wish to estimate. We can talk about errors for a single measurement, or for an average based on many observations. For example, suppose the national average of lead in paint on wet room window sills is 0.8 mg/sq cm. If a single measurement for a wet room window sill is 1.5 mg/sq cm, then the error is 0.7 mg/sq cm. Similarly, if the (weighted) average of all measurements taken from window sills in wet rooms is 0.6 mg/sq cm, then the error is -0.2 mg/sq cm.

There are two types of errors: bias and variable error.

- **Bias -** is a constant error because all possible surveys using the same design would overestimate (or underestimate) the population parameter, on the average. Biases arise from a number of sources, including differences between the sample frame and the target population, differential response rates from different census blocks (i.e, segment) of the sampled population, uncalibrated or mis-calibrated field or laboratory measurement equipment, and some types of data-reduction procedures.

- **Variable Error -** is an error caused by the random variation inherent in any sampling or measurement process. A variable error is specific to each measurement and generally cannot be estimated or statistically corrected and remain in the data even if all systematic errors have been eliminated. Variable errors result from both sampling and measurement processes, and are typically reported as a variance or standard deviation.

- **Precision -** refers to the size of the variable error. If the variable error is small, we say that the estimate is precise.

Throughout this chapter we will make use of the following term:

- **Sample Weight** - is the number of housing units in the target population that a sampled unit represents. The sample weight can be calculated by taking the inverse of the probability of selection for that unit. Thus, if the probability of selection is .01, the sample weight is 100. With multi-stage samples, the overall probability of selection is the product of the conditional probabilities of selection at each stage.

3.2 Response Rates and Potential for Non-Response Bias

The *Comprehensive and Workable Plan for the Abatement of Lead-Based Paint in Privately Owned Housing: Report to Congress* includes a brief discussion of response rates in its Appendix A. That Appendix reports national response rates by construction year stratum for single and multi-family housing units.

The objective of the analysis described in this section is to estimate the potential impact of national survey nonresponse on the estimated prevalence of lead-based paint in housing. To accomplish this, a detailed analysis of response rates at each stage of the survey was conducted for each of the 150 census blocks that were surveyed (see Appendix I for a description of the survey methodology, including definitions of terms). The analysis looked at the relationship between response rates and factors such as ethnicity, geographic location and economic measures of wealth (rent, home value, and income) that might be related to response rates. In addition, the analysis studied the association between housing units in the same or nearby census blocks with respect to the presence or absence of lead in paint, dust, and soil.

3.2.1 Private Housing

For private housing, no statistically significant relationship was observed between response rate and ethnicity, income or age of the housing units. In addition, there was a strong positive association between inspected housing units in the same census block with respect to the presence or absence of lead in paint, dust, and soil. On the other hand, the lowest rent category homes (< $200/month) and the highest market value homes (<$150,000) appear to be slightly under represented, and the South is somewhat overrespresented in the sample. On balance, these findings suggest that the potential bias due to nonresponse is likely to be small. Therefore, the estimates of lead-based paint prevalence were not adjusted for non-response.

The first step in the non-response analysis was the calculation of the rates of being successfully contacted, being eligible for inclusion in the survey, and being a respondent in the surveys for each census block. Then these rates were analyzed by region and economic variables using frequency distributions, Pearson correlation coefficients, and tabulations of rates versus the different variables. The results of the analysis are presented in Tables 3-1 through 3-5.

Table 3-1 shows national contact rates, response rates, and eligibility rates, where applicable. These rates can be viewed as conditional success rates -- each were calculated using only those potential respondents who reached that particular stage.

The meaning of the term "complete" varies with each stage of data collection. A case was considered completed at the screener stage if an interview was completed, regardless of eligibility. Thus, the screener response rate measures the rate at which the screening interview reached its logical conclusion (eligible or ineligible), and did not terminate due to other causes (refusals, break offs, etc.). At the telephone interview stage, only eligible respondents who completed interviews and scheduled appointments were considered "completes." Finally, a case was considered "complete" at the inspection stage if an inspection was completed.

A "refusal" occurred during the telephone interview when an eligible respondent refused to schedule an appointment for an inspection. A "refusal" occurred during the inspection when an eligible respondent refused to schedule an appointment for an inspection, or scheduled an appointment and then refused the field team entry into the housing unit. A "break off" occurred when the field team was refused an inspection after the start of the inspection interview.

TABLE 3-1

NATIONAL RESPONSE, CONTACT, AND ELIGIBILITY RATES AT EACH DATA COLLECTION STAGE

Data Collection Stage	Contact Rate	Response Rate	Eligibility Rate
Screener	77%	63%	89%
Telephone Interview	83%	55%	
Inspection		90%	

National rates in the table were calculated as follows:

- Screener Contact Rate = (Total Attempts - Not at Homes - Vacant)/(Total Attempts - Vacant)
  Screener Response Rate = (Completes)/(Total Attempts - Vacant Homes)
  Screener Eligibility Rate = (Eligibles)/(Completes)

- Telephone Interview Contact Rate = (Completes + Language Problems + Refusals)/(Total Attempts)
  Telephone Interview Response Rate = (Completes)/(Total Attempts)

- Inspection Response Rate = (Completes)/(Completes + Refusals + Break Offs)

Table 3-2 shows the number of census blocks, by Census Region and stage of data collection, that were lost because they had a response rate of zero. Percentages of census blocks lost ranged by region from 17 percent (6 of 35) in the Midwest to 32 percent (8 of 25) in the West. Overall, 23 percent (35 of 150) of all census blocks surveyed were lost.

A number of reasons for Census block loss were documented. Ten Census blocks were lost during the listing and screening stages of data collection. The main reasons for these losses were: being located in drug or other dangerous areas with potential for endangering the field staff; all housing units being ineligible because they were public, post-1980 construction, or commercial; and all housing units being inaccessible (all buildings in the segment were secured). Twenty-five census blocks were lost during the listing and screening stages of data collection. The main reasons for these losses were refusals and inability to schedule inspections with respondents during the 1-2 week period when the field team was in town. Most (51 percent) of the lost census blocks occurred during the telephone interview stage.

TABLE 3-2

CENSUS BLOCKS LOST BY CENSUS REGION AND SURVEY DATA COLLECTION STAGE

Region	Segments in Regions	Survey Data Collection Stage				Total
		Listing	Screening	Telephone Interview	Inspection	
Northeast	40	2	2	6	1	11
Midwest	35	1	0	3	2	6
West	25	1	2	4	1	8
South	50	1	1	5	3	10
Total	150	5	5	18	7	35

One important measure of the respresentiveness of the National Survey is to examine how the distributions of the housing characteristics and socioeconomic and ethnic factors in the National Survey compare to national distributions. National distributions were obtained from the American Housing Survey (AHS) for 1987 performed by the Bureau of the Census and the Department of Housing and Urban Development. The distributions of ethnicity, income, building age, region of the country, monthly rent and market value from the National Survey were compared to their respective national distributions and presented in Table 3-3. Chi-square tests were used to determine how the distributions in the National Survey compared to those from the American Housing Survey. In most cases, the distribution of households in the National Survey were not significantly different from those in the American Housing Survey. No significant differences were observed in the distributions of ethnicity, building age, or household income. The few cases where there were significant differences involve the following. There were fewer tenant-occupied housing units with lower rents (rent less than $200 a month) than expected given national estimates; more expensive owner-occupied housing (market value greater then $150,000) than expected given national estimates; and more dwelling units located in the South than expected given national estimates.

TABLE 3-3

CHI-SQUARE RESULTS FOR DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES

a. Ethnicity

Ethnicity	Black	Hispanic	Other[1]
Observed frequency in National Survey	29	24	230
National distribution, from AHS (000)	9,261	4,977	64,943
Expected frequency, from AHS*	33	18	232
Individual Chi-square values*	0.508	2.169	0.019

*The chi-square statistic was calculated assuming a fixed total of 283 homes with data on race (3 cells and 2 degrees of freedom).

Total Chi-square statistic 2.812
P-value with 2 degrees of freedom 0.245

[1]Other race includes non-Hispanic whites, Asians and Pacific Islanders, Eskimos and American Indians and other non-blacks

b. Building Age

Building Age	pre-1940	1940 to 1959	1960 to 1979
Observed frequency in National Survey	77	87	120
National distribution, from AHS (000)	21,215	21,001	36,965
Expected frequency, from AHS**	76	75	133
Individual Chi-square values**	0.011	1.810	1.194

**The chi-square statistic was calculated assuming a fixed total of 284 homes with data on building age (3 cells and 2 degrees of freedom).

Total Chi-square statistic 3.015
P-value with 2 degrees of freedom 0.221

c. Region

Region of the Country	Northeast	Midwest	South	West
Observed frequency in National Survey	52	69	116	46
National distribution, from AHS (000)	17,618	20,344	25,589	15,628
Expected frequency, from AHS**	63	73	91	56
Individual Chi-square values**	1.911	0.190	6.585	1.740

*The chi-square statistic was calculated assuming a fixed total of 283 homes with data on region (4 cells and 3 degrees of freedom).

Total Chi-square statistic	10.423
P-value with 3 degrees of freedom	0.015

d. Household Income

Household Income	< $10,000	$10,000 to $19,999	$20,000 to $29,999	> $30,000
Observed frequency in National Survey	51	49	56	107
National distribution, from AHS (000)	15,482	17,090	15,102	31,147
Expected frequency, from AHS**	53	57	50	103
Individual Chi-square values**	0.050	1.062	0.680	0.121

**The chi-square statistic was calculated assuming a fixed grand total of 263 homes (4 cells and 3 degrees of freedom).

Total Chi-square statistic	1.913
P-value with 3 degrees of freedom	0.590

e. Monthly Rent

Monthly Rent	< $200	$200-$399	> $400
Observed frequency in National Survey	12	47	40
National distribution, from AHS (000)	5,886	12,230	8,560
Expected frequency, from AHS**	22	45	32
Individual Chi-square values**	4.436	0.057	2.132

*The chi-square statistic was calculated assuming a fixed grand total of 105 homes (3 cells and 2 degrees of freedom).

Total Chi-square statistic 6.625
P-value with 2 degrees of freedom 0.036

f. Current Market Value

Current Market Value	<$40,000	$40,000-$59,999	$60,000-$79,999	$80,000-$99,999	$100,000-$150,000	>$150,000
Observed frequency in National Survey	39	21	25	16	29	42
National distribution, from AHS (000)	11,885	10,228	9,173	5,582	6,281	7,405
Expected frequency, from AHS**	41	35	31	19	21	25
Individual Chi-square values**	0.051	5.472	1.235	0.471	2.724	11.211

**The chi-square statistic was calculated assuming a fixed grand total of 172 homes (6 cells and 5 degrees of freedom).

Total Chi-square statistic 14.406
P-value with 5 degrees of freedom 0.013

Table 3-4 shows the distribution of the overall Census block response rate by region. Census blocks in the Northeast had the lowest overall response rate where only 35% of the census blocks had response rates over 25 percent. In contrast, census blocks in the Midwest and South had the highest overall response rates where 66% and 68% of the respective census blocks had response rates over 25 percent.

TABLE 3-4

NUMBER AND PERCENT OF CENSUS BLOCKS BY OVERALL RESPONSE RATE[1] AND CENSUS REGION

	Segments with Indicated Overall Response Rate				
	Overall Response Rate				
Region	**Equal to 0%**	**0% and < 25%**	**> = 25% and = < 75%**	**More than 75%**	**Total**
Northeast	11 (28%)	15 (37%)	13 (33%)	1 (2%)	40 (100%)
Midwest	6 (17%)	6 (17%)	21 (60%)	2 (6%)	35 (100%)
West	8 (32%)	4 (16%)	12 (48%)	1 (4%)	25 (100%)
South	10 (20%)	6 (12%)	29 (58%)	5 (10%)	50 (100%)
Total	35 (23%)	31 (21%)	75 (50%)	9 (6%)	150 (100%)

1 The Overall Response Rate = (Screener Response Rate)(Telephone Response Rate)(Inspection Response Rate)

Table 3-5 addresses the question: "If one housing unit in a segment has (or does not have) a particular lead characteristic, then do all the other housing units in the same segment of similar age (year of construction) have the same lead characteristic?" For purposes of this table, census blocks with less than two housing units inspected were excluded. Fifty percent of the census blocks surveyed had two or more inspected housing units.

The first column in Table 3-5 displays the number of census blocks where all housing units in a segment had the same lead characteristic. The last two columns break down census blocks where there were differences into two categories: census blocks where there were differences within houses in the same construction year category; and census blocks where differences could always be explained by construction year category. Table 3-5 shows that lead characteristics were the same for most census blocks for each of the four types of lead characteristics analyzed, irrespective of the year of construction. For example, if we look at the characteristic "lead-based paint" we find that all houses within a segment had the same lead characteristic for 64 percent of the census blocks. Of the remaining 36 percent where there were differences, 8 percent is explained by construction year category. Differences in lead-based paint characteristic between house built in different year categories were found in only 28 percent of the census blocks.

TABLE 3-5

ASSOCIATION BETWEEN INSPECTED HOUSING UNITS IN THE SAME CENSUS BLOCKS WITH RESPECT TO THE PRESENCE/ABSENCE OF LEAD IN PAINT, DUST, AND SOIL

Presence or Absence of Lead	Segments[1] Where All Housing Units Have the Lead Characteristics or All Housing Units Do Not Have the Lead Characteristics	Segments[1] Where Some Housing Units Have and Some Do Not Have the Lead Characteristics	
		All Housing Units in the Same Construction Year Category[2] Have the Same Lead Characteristics	Some Housing Units in the Same Construction Year Category[2] Have Different Lead Characteristics
Lead-based paint	48 (64%)	6 (8%)	21 (28%)
Lead in dust above guidance	61 (81%)	3 (4%)	11(15%)
Lead in soil above guidance	59 (79%)	10 (13%)	6 (8%)
Damaged lead-based paint	51 (68%)	9 (12%)	15 (20%)

Note: • The results are based on unadjusted XRF readings of 1.0 mg/sq cm or greater.
• Of the 150 segments that were surveyed, 75 segments had two or more housing units in the segment.

1 The percentage of segments in the three columns were calculated using the total number of segments with two or more dwelling units as the base.
2 Housing units were grouped into 3 categories based on construction year: Pre 1940, 1940 - 1959, and 1960 - 1979.
3 At least 5 square feet of damaged lead-based paint.

3.2.2 Public Housing

For the public housing component of the national survey of lead-based paint, a sample of 110 projects from a national frame of public housing projects[12] was drawn according to the design described in Chapters 2 and 3 of Appendix I of this report. The survey design specified a visit to one randomly selected housing unit from each of these 110 projects. The survey was completed in 97 of the 110 sampled projects. Of the 13 nonrespondent projects, eight were excluded because they were found to be out of scope -- they either had no family units, did not exist, or were built since 1980. The remaining five projects were not completed because of problems encountered in scheduling interviews during the short period when the field team was in town.

The nation's public housing projects were well represented in the sample. The response rate was 95 percent (97 completes among 102 eligible projects). Despite the fact that the sample design was unbiased and there was no evidence of bias in the selection of the projects, vacant housing units appeared to be over-represented in the sample. Of 97 eligible units, 44 were vacant. Such a large number were undoubtedly

[12] A public housing project is a complex of housing units developed and built at the same time, or units acquired from different locations but grouped together for administrative process.

vacant because field technicians were sometimes steered to vacant units selected by the public housing project manager.

Although a comparison of the prevalence of lead in paint between vacant and occupied housing units showed no significant differences, the impact on other target parameters is unknown. For example, it is difficult to estimate the percentage of public housing units with both lead-based paint and children under the age of seven because so few occupied apartments were sampled. Furthermore, dust samples collected in vacant apartments probably do not represent samples collected in occupied apartments. Because the public housing sample sizes were small and the occupied housing unit sizes were even smaller, definitive conclusions resulting in estimates of dust lead levels based on the national survey data are not possible and are not recommended.

The sample was carefully designed to be representative. A sample is said to be representative if the distributions of its characteristics are about the same for the sample as they are for the target population. Because of sampling error, perfect agreement is not expected. The design for the national survey is described in Appendix I, Chapter 3, Sample Design and Selection. The design has the property of being statistically unbiased. That is, every public housing project in the nation had a known, positive probability of being selected into the sample. The sample weights were calculated to properly reflect these varying probabilities.

We can assess representativeness by comparing the characteristics of the weighted sample to the characteristics of the target population. Tables 3-6, 3-7, and 3-8 present tabulations of the public housing frames and sample by construction year, census region, and size of the public housing authority (PHA), respectively. The first column of all three tables is the same and shows the steps that were taken to develop the sample from the target population:

- **Full National Frame --** is the inventory of HUD's public housing units data file that constitutes the target population.

- **County Extract, Unedited --** refers to the 30 primary sampling units (or PSU's) that were selected from the full national frame.

- **County Extract, Updated -** refers to a revision to the number of housing units in the sampled counties that was made after contacting the PHAs. Details concerning these sampling procedures are contained in Appendix I, Chapter 3 of this report.

- **Final Sample --** refers to the sample drawn from the 30 country updated extract.

The projected figures given for the county extracts and the final sample were obtained by applying the sample weights to the data.

We can assess the representativeness of the sample in Tables 3-6, 3-7, and 3-8 by comparing the actual percentages for the full national frame to the projected percentages shown for the final sample. When examined by construction year category (Table 3-6), we see that the distribution of the sample projections is similar to the distribution for the full national frame. When examined by geography (Table 3-7), the sample projections are relatively high in the Northeast (44% versus 31%) and low in the Midwest (11% versus 17%) and the South (34% versus 41%), compared to the figure for the full

TABLE 3-6

DISTRIBUTION OF PUBLIC HOUSING FAMILY UNITS BY CONSTRUCTION YEAR

File	Construction Year Stratum			Total
	pre-1950	1950-1959	1960-1979	
Full National Frame (1)				
Number of units	161,501	246,680	388,475	796,656
Percent	20%	31%	49%	100%
30 County Extract, Unedited (1)				
Number of units	37,060	56,580	86,355	179,995
Percent	21%	31%	48%	100%
Projected to Nation (2)				
Number of units	158,534	244,610	394,496	797,640
Percent	20%	31%	49%	100%
30 County Extract, updated (3)				
Number of units (4)	44,700	77,189	80,073	201,962
Percent	22%	38%	40%	100%
Projected to Nation (2)				
Number of units (5)	165,233	281,529	353,427	800,189
Percent	21%	35%	44%	100%
Final Sample				
Number of units	30	24	43	97
Percent	31%	25%	44%	100%
Projected to Nation (6)				
Number of units	182,000	273,000	455,000	910,000
Percent	20%	30%	50%	100%

NOTES:

(1) Source: HUD's inventory of public housing units data file.
(2) National Projection obtained by use of the PSU weights.
(3) Source: Revisions to HUD's inventory made after contacting the respective PHA's.
(4) Excludes 2,519 units in 8 projects found to be out of scope during field work
(5) Excludes 62,555 units projected from the 8 projects in note 4.
(6) National Projection obtained by use of the final sampling weights, which were adjusted to conform to HUD's counts of the number of family units in each construction year stratum.

TABLE 3-7

DISTRIBUTION OF PUBLIC HOUSING FAMILY UNITS BY CENSUS REGION

File	Census Region				Total
	Northeast	Midwest	South	West	
Full National Frame (1)					
Number of units	271,924	151,661	361,280	90,154	875,019
Percent	31%	17%	41%	10%	100%
30 County Extract, Unedited (1)					
Number of units	99,926	36,747	35,339	14,198	186,210
Percent	54%	20%	19%	8%	100%
Projected to Nation (2)					
Number of units	328,091	107,932	283,800	103,463	823,286
Percent	40%	13%	34%	13%	100%
30 County Extract, updated (3)					
Number of units (4)	391,240	102,960	305,660	110,140	910,000
Percent	43%	11%	34%	12%	100%
Projected to Nation (1)					
Number of units (5)	374,844	84,550	260,566	83,710	803,670
Percent	47%	11%	32%	10%	100%
Final Sample					
Number of units	43	11	32	11	97
Percent	44%	11%	33%	11%	100%
Projected to Nation (6)					
Number of units	391,236	102,962	305,661	110,140	910,000
Percent	43%	11%	34%	12%	100%

NOTES:

(1) Source: HUD's inventory of public housing units data file, including post-1980 buildings.
(2) National Projection obtained by use of the PSU weights.
(3) Source: Revisions to HUD's inventory made after contacting the respective PHA's.
(4) Excludes 2,519 units in 8 projects found to be out of scope during field work
(5) Excludes 62,555 units projected from the 8 projects in note 4.
(6) National Projection obtained by use of the final sampling weights, which were adjusted to conform to HUD's counts of the number of family units in each construction year stratum.

national frame. Finally, when examined by PHA size (Table 3-8), the sample projections are relatively low for the smallest size category (29% versus 40%) and high for the largest size category (52% versus 32%).

Thus, although the sample appears fairly representative with regard to age, some differences appear for both geography and PHA size. A second look at Tables 3-7 and 3-8 helps explain where these differences occurred. For both tables, the distribution appear fairly consistent among the final sample and 30 county extracts. It is at the first stage of sampling, when the 30 counties were selected from the full national frame, that the major differences appear to have occurred. The sample of 30 counties was designed to be representative of privately-owned housing and not public housing. The survey design is described in Appendix I, Chapter 3, of this report entitled: Sample Design and Selection.

Techniques such as post-stratification can be used to adjust the sample weights so the distribution by characteristics is more similar to the distribution of units on the full national frame. Representativeness in itself is desirable because it eliminates a possible source of bias. But post-stratification for this sample would also increase the variance of the estimates. Consider, for example, representation with regard to size (Table 3-8). If we post-stratify by PHA size (Table 3-8), then the weights for housing units in large PHAs would become smaller and the weights in small PHAs would become larger. This would increase the heterogeneity in the weights, since the large PHAs would now tend to have the small weights and the small PHAs the large weights. Increased heterogeneity in the weights results in increased variance, or reduced precision.

There was a second reason why post-stratification was undesirable. Post-stratification would have required the use of the HUD public housing data for the entire nation. We encountered numerous errors when updating the 30-county extract, which suggested that the error rate is high in the entire national file. Of course it was not possible to remove the errors in the entire national file. Thus, there is no way to assess the accuracy of the distributions of characteristics derived from the full national file.

Thus, while post-stratification would eliminate one source of bias, it would tend to increase the variance of estimates due to the increased heterogeneity of the sample weights, and it would introduce a second source of bias and variance: the errors in the full national frame. On balance, the disadvantages of post-stratification were felt to outweigh the advantages, especially since there was no need to present the survey results by PHA size or census region.

The public housing survey design called for randomly selecting one housing unit from each sampled project. This design was difficult to implement in the field. As mentioned above, 44 of the 97 housing units surveyed were apparently vacant. The number of vacant units was so large because some public housing project managers apparently steered the field team to vacant housing units. This is a potential source of bias since the public housing project managers may have steered the field teams to housing units that were not representative of the units in the housing projects with respect to the survey variables. A systematic difference between vacant and occupied housing units would be an indication of such a bias. Since only one unit was selected from each project, no estimates of bias could be constructed.

Table 3-9 gives the number of housing units sampled broken down by occupancy status and prevalence of selected lead-related characteristics. To examine the possibility of "steering bias," the percentages of occupied and vacant housing units that had each lead characteristic were recompared.

TABLE 3-8

DISTRIBUTION OF PUBLIC HOUSING FAMILY UNITS BY PHA SIZE

	PHA Size (No. of Family Units)				
File	**<1,250**	**1,250-2,350**	**2,351-10,000**	**> 10,000**	**Total**
Full National Frame (1)					
Number of units	368,777	93,396	169,807	301,593	933,573
Percent	40%	10%	18%	32%	100%
30 County Extract, Unedited (1)					
Number of units	18,737	6,192	15,927	145,354	186,210
Percent	10%	3%	9%	78%	100%
Projected to Nation (2)					
Number of units	252,774	60,437	113,355	396,720	823,286
Percent	31%	7%	14%	48%	100%
30 County Extract, updated (3)					
Number of units (4)	13,885	7,882	13,634	167,645	203,046
Percent	7%	4%	7%	83%	100%
Projected to Nation (1)					
Number of units (5)	224,182	49,688	89,694	442,970	806,534
Percent	28%	6%	11%	55%	100%
Final Sample					
Number of units	22	7	14	54	97
Percent	23%	7%	14%	56%	100%
Projected to Nation (6)					
Number of units	263,165	64,548	109,793	472,506	910,000
Percent	29%	7%	12%	52%	100%

NOTES:

(1) Source: HUD's inventory of public housing units data file, including post-1980 buildings and buildings outside the 48 contiguous states.
(2) National Projection obtained by use of the PSU weights.
(3) Source: Revisions to HUD's inventory made after contacting the respective PHA's.
(4) Excludes 2,519 units in 8 projects found to be out of scope during field work
(5) Excludes 62,555 units projected from the 8 projects in note 4.
(6) National Projection obtained by use of the final sampling weights, which were adjusted to conform to HUD's counts of the number of family units in each construction year stratum.
(7) Sums may not equal total due to rounding error.

TABLE 3-9

PAINT LEAD, PAINT DAMAGE AND DUST LEAD IN PUBLIC HOUSING, BY OCCUPANCY UNWEIGHTED SAMPLE COUNTS

	Public Housing			
	Occupied		Vacant	
	53 Units in Sample		44 Units in Sample	
Characteristic	Number	Percent	Number	Percent
Interior:				
Lead-based paint (LBP)	42	79%	35	80%
Damaged LBP (1)	1	2%	3	7%
Damaged Paint (1)	10	19%	15	34%
High lead dust (2)	1	2%	3	7%
Exterior:				
Lead-based paint	35	66%	32	73%
Damaged LBP (1)	4	8%	2	5%
Damaged paint (1)	10	19%	8	18%
Priority hazard present (3)	6	11%	6	14%

(1) Paint is considered to be "damaged" if more than five square feet of it is peeling, chipped, or otherwise damaged.

(2) Dust is considered to be "high lead dust" if the dust lead level exceeds the clearance levels in the HUD Guidelines, 200ug/sq ft on floors , 500 ug/sq ft on window sills, and 800 ug/sq ft on window wells.

(3) Priority hazard is present if the dwelling unit has LBP, either inside or outside and either high lead dust or more than five square feet of damaged LBP (total, inside and outside).

Using interior damaged paint as an example, the test statistic, z, was calculated to be 0.71 using the standard formula for comparing two independent proportions.[13] The critical value for the significance test is 1.96. Since 1.2 is less than 1.96, we cannot conclude that there is any difference between occupied and vacant housing units with regard to the prevalence of interior damaged paint. Similar calculations for the other lead characteristics and no significant differences were found between vacant and occupied housing units. Thus, there is no statistical evidence of any differences between vacant and occupied housing units with respect to the selected lead-related characteristics.

3.3 Correcting for Measurement Bias

This section describes the methods used to correct measurements for calibration bias and censoring bias (to be defined below).

Definitions of Four Types of Measurements

Each measurement passed through a series of transformations from the time it was first calculated internally by the MAP/XRF instrument, to when it was fully corrected for bias. To avoid confusion, different terminology will be used to refer to the measurement as it passed through different stages of processing. Four types of measurements are now defined:

- An **internal measurement** is a value the MAP/XRF instrument calculated internally before it displayed a number. The technician cannot observe this number. Although the true lead concentration cannot be negative, the internal measurement on surfaces with little or no lead paint can be negative either because of measurement bias or variable error.

- A **field measurement** is the number displayed by the MAP/XRF instrument and recorded in the field. The field measurement is different from the internal measurement when the latter is negative, in which case the field measurement is zero. This is called censoring.

- A **recalibrated measurement** is a field measurement after being corrected for calibration bias.

- A **corrected measurement** is a field measurement after being fully corrected for both calibration and censoring bias. The national estimates for the prevalence of lead-based paint are based on the corrected measurements.

Measurement bias is a common phenomenon in a field study that uses equipment such as MAP/XRF instruments. According to our field procedures, a single measurement was taken on each surface measured within a dwelling unit. Compared to the true lead concentration, the measurement tended to be larger (upward bias) in some situations, and smaller in others (downward bias). There were two types of measurement bias:

- **Calibration** bias occurred when the internal measurement tended to be systematically different from the true lead concentration being measured; and

[13] $Z=(p_1 - p_2)/(\text{var } p_1 + \text{var } p_2)$ 1/2, where var p=p(1-p) D/n and D is the design effect, D=1.45, obtained in the *Comprehensive and Workable Plan* report, page A-18.

- **Censoring bias** occurred when a negative internal measurement was displayed as zero. All zero field measurements were said to be censored because the internal measurement could not be observed. The only thing known about the internal measurement corresponding to a censored field measurement is that the internal measurement was less than or equal to zero.

The remainder of this section describes in detail the methods used to correct the field measurement first for calibration and then for censoring bias.

3.3.1 Adjusting Field Measurements for Calibration Bias

The possibility of calibration bias was anticipated in advance of the field period. In order to estimate and subsequently correct for bias, a provision was added to the survey procedures to collect "validation" measurements.

Validation Measurements

Validation measurements consisted of field measurements taken on surfaces (called *shims*) with known lead concentrations. The shims were prepared by painting 3 by 4 inch heavyweight paper sheets with lead-based paint. According to NIST, the shims had lead concentrations of 0.6 and 2.99 mg/cm^2. Technicians placed shims over substrates made from four different types of material: wood, steel, drywall, and concrete. One set of shims and one set of substrate samples were kept with each of eight MAP/XRF instruments. The survey field technicians collected validation measurements on the shims during three different time periods:

- **Baseline** validation measurements were taken before the field period. After the MAP/XRF instruments were received from the manufacturer but before they were used in the field for the survey, technicians collected eight replicate measurements on each of the four substrates using each of the two shims, for a total of 64 measurements. In addition, the technicians took measurements using the substrate material alone with no shim (roughly equivalent to a shim with 0.0 mg lead/sq. cm).

- **Daily** validation measurements were taken during the actual field period. Field technicians took one measurement of each shim on each substrate at the beginning and end of each day, for a total of 16 measurements per day. The Daily validation measurements were used to develop the calibration equations.

- **Closeout** validation measurements were taken after the field period. These measurements were taken in a fashion similar to the Baseline validation measurements, except they were taken after the field period but before the instruments were returned to the manufacturers.

The validation measurements were studied and found to be consistent with the following assumptions:

- The MAP/XRF instrument calculates an internal value, referred to here as the internal measurement.

- The MAP/XRF instrument displays the field measurement which is equal to the maximum of zero and the internal measurement.

- The relationship between the internal measurement and the true lead concentration in the shims is approximately linear.

- The distribution of the variable error for a measurement, or measurement variance, for the internal measurements is approximately normal.

- The standard deviation of the measurement variance does not depend on the true lead concentration in any substantial way.

It should be noted that corrected measurements analyzed by NIST[14] are also consistent with these distributional assumptions.

Outliers in the Validation Data

During the preliminary exploratory data analysis, and later while processing the data, a few unusual measurements and patterns were identified. After examining the original data sheets and any notes written by the field technicians, 24 measurements (less than 1% of the data) were classified as outliers because they were either very unusual, or taken at the same time as several other measurements that were also unusual. With these outliers excluded, the data are consistent with the assumption that the measurement variance had a normal distribution. It was then valid to perform the statistical tests available for linear regression procedures which assume normality.

Estimating the Mean of the Internal Measurements

For purposes of constructing the calibration equations, it was necessary to estimate the theoretical mean of the internal measurements, say μ_{ij}, for each combination of the i^{th} substrate and j^{th} shim. With uncensored and normally distributed data, regression is the preferred method of statistical analysis. But censoring complicates the estimation of model equations because it results in upward bias for the regression estimate of μ_{ij} and downward bias for the regression estimate of the sample variance. An alternative estimator, the sample median, is an unbiased estimator of μ_{ij} (assuming the distribution of measurements is symmetric) when the proportion of censored values is less then 50 percent.

Which is the better estimator of μ_{ij}: the sample mean or the sample median? A standard accepted measure of the goodness of an estimator is the mean squared error (MSE) or its square root, the root mean squared error (RMSE). The MSE is equal to the magnitude of the expected squared difference (or error) between the estimator and the parameter being estimated. The MSE for an estimator is a function of both its variance and bias as follows:

$$\text{MSE} = \text{Variance} + \text{Bias}^{2}.$$

One way to choose among estimators is to compare their RMSE's. The RMSE's for both the sample mean and the sample median are plotted in Figure 3-1 as a function of the percentage of zeros in the data.[15]

---

[14] McKnight, M. Byrd, W. Roberts, W., *Measuring Lead Concentration in Paint Using a Portable Spectrum Analyzer X-Ray Fluorescence Device,* Building Materials Division, National Institute of Standards and Technology, U.S. Department of Commerce, May 1990, NISTIR W90-650, and personal communication with Mary McKnight.

The RMSE for the sample mean varies with the percentage of zeros in the data. The bias for the sample mean is also graphed and can be seen increasing with the percentage of zeros in the data. The RMSE for the sample median is a constant, and the bias is zero, as long as the percentage of zeros is 50% or less.

With less than 7% censoring, both the sample mean and the sample median had little bias. But the sample mean had lower RMSE and was therefore the preferred estimate of μ_{ij}. With 37% zeros, the sample mean and sample median had the same RMSE, but the sample mean was quite biased. Thus, in this case, the sample median was the preferred estimator. When the percentage of zeros is between 7% and 37%, the choice between estimators was a matter of judgment.

Factors that Affected the Means of the Internal Measurements

The daily validation measurements were used in a regression analysis to determine the factors that affect the means of the internal measurements, to develop the model, and to calibrate equations. It was desirable to use as much of the data as possible and, at the same time, to minimize the effect of censoring. Thus all the data were used from instrument-shim-substrate combinations where the percentage of zeros was less than 20%. If there were fewer than 20% zeros in the data set, the mean was used; otherwise the median was used. A cutoff of 20% was chosen because it is midway between 7% and 37%. Since there were few data sets where the percentage of zeros was near 20%, the results of the analyses are not sensitive to the choice of the cutoff.

As will be described below, the magnitude of the measurement variance differed across instruments, substrates, and, to some extent, shims. Accordingly, weighted regression was used to identify factors that had a significant effect on the calibration equations. Each weight was calculated as the inverse of the estimated measurement variance within each instrument-substrate-shim combination. A preliminary weighted regression analysis showed that the model depends on a number of factors:

- **Period**. The same set of shims was kept with each MAP/XRF instrument throughout the field period. However, no attempt was made to use these same shims during the Baseline and Closeout validation periods. It is possible that different shims (with slightly different lead concentrations) may have been used for the different validation periods for some or all of the MAP/XRF instrument. Thus it was decided to use only the Daily validation measurements to derive the model and calibration equations.

- **Instrument, Substrate, and Shim**. It is not possible to estimate how much of the apparent instrument-to-instrument differences are due to shim differences, substrate differences, or other differences associated with the instruments. This is because only one set of shims and substrates were used for each instrument within a given period. The shims had lead concentrations of 0.6±0.02 and 2.99±0.30 mg/cm2, where the error ranges are standard deviations.[16]

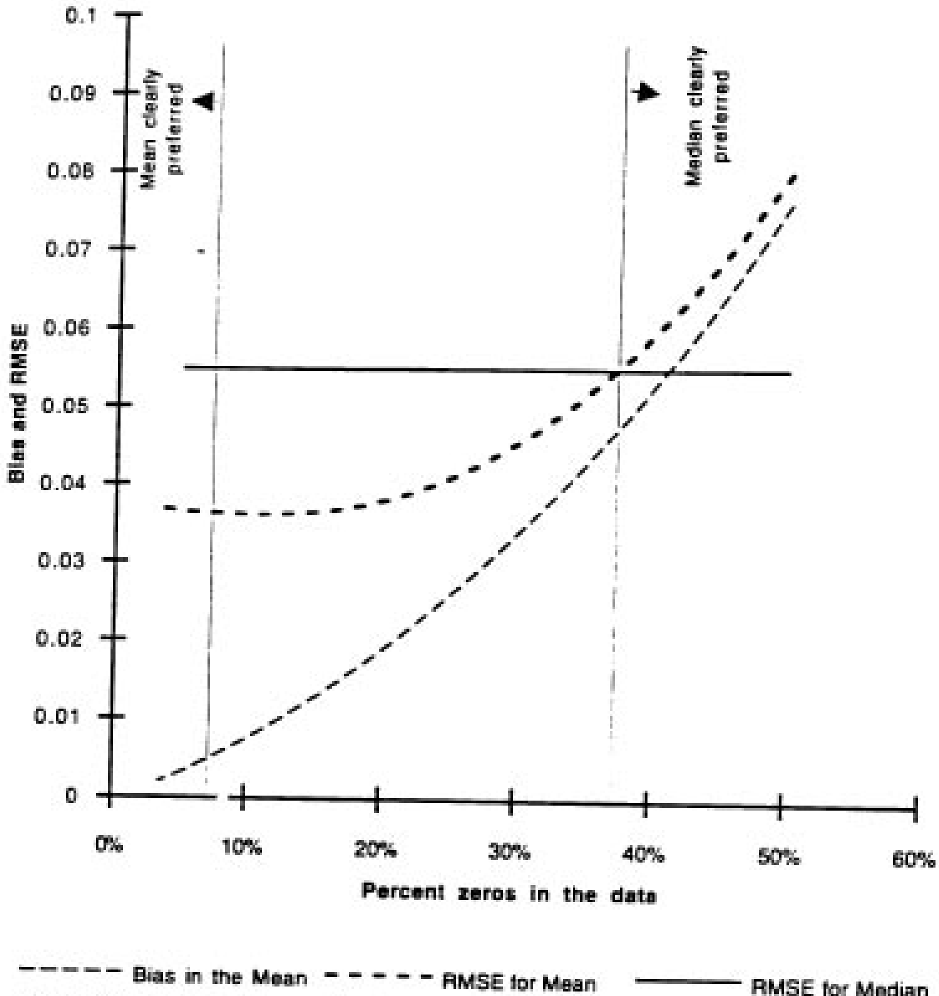
---

[15] The calculations are based on a sample of 45 normally distributed measurements (the average number within a substrate -shim group) with standard deviation of 0.25 mg/sq cm, rounded to the tenths of a unit. On the average, 45 measurements were taken on each substrate by each XRF instrument. The standard deviation used is similar to that for the wood and drywall measurements.

[16] We interpreted the error ranges on these shim concentrations as standard deviations, following consultations with Mary McKnight at NIST.

FIGURE 3-1

RMSE AND BIAS OF THE SAMPLE MEAN AND MEDIAN WHEN ESTIMATING THE MEAN OF INTERNAL MEASUREMENTS



Calculations assume 45 normally distributed values with standard deviation of 0.25 and rounding to tenths of a unit

- **Age of Photon Emission Source**. The rate of photon emissions is known to decay over time. The rates of decay were not significantly different across instruments or substrates. Therefore, a single formula was used for all measurements.

The effect of operator was investigated, but the results were not statistically significant. Using the Daily validation measurements, a separate model equation, and subsequently a separate calibration equation, was developed for each combination of MAP/XRF instrument and substrate.

Factors that Affected the Variances of the Internal Measurements

The spectrum analyzer MAP/XRF instrument measures the lead content in the wall by exposing a small portion of the wall material to radiation. In response to this exposure, the lead on both the surface and the substrate emit photons at specific frequencies. The MAP/XRF instrument measures the energy and number of photons emitted by the lead and substrate. The internal measurement is based on the relative number of the photons at different energies.

For the Daily validation measurements, there were significant differences in the measurement variance associated with each combination of shim, substrate, and MAP/XRF instrument ($p<.0001$). Consistent differences in variance were also found among substrates and shims across all instruments. Factors that affect variance include:

- **Instrument characteristics**. Characteristics such as the shim samples, substrate samples, calibration, temperature, battery power, and how the technician holds and operates the instrument, can all induce variance. Differences in the strength of the instrument's radioactive source or the sensitivity of the electronics can result in differing measurement variance among MAP/XRF instruments.

- **Photon Emission**. The number of photons emitted and detected during the sampling period has inherent variability. Because the field measurement depends on the number of photons detected per unit time, there is a natural measurement variance even with all other factors held constant. Using longer exposures reduces this variance. All the measurements collected in the national survey were based on a 60 second exposure. Based on physical principles, the number of photons detected should follow a Poisson distribution.

- **Surface Characteristics**. Variance in the homogeneity of the substrate material can result in differences in the measurement variance among substrates.

- **Lead concentration**. Considering the various factors that can affect the field measurement, the variability of the measurements is expected to either remain constant or increase as lead concentration increases. This relationship was investigated using the Daily validation measurements. The hypotheses could not be rejected that the variability of the measurements remains constant across different lead concentrations ($p=.62$).

Table 3-10 and Figure 3-2 show the pooled measurement standard deviation across all instruments for each combination of shim and substrate. Also shown are the measurement standard deviations determined by NIST for comparable 60-second sampling periods. The measurement variance for the data collected in the national survey is similar to that found by NIST. Any differences in the measurement variance between the NIST data and the Daily validation measurements may reflect instruments, instrument programming, or the conditions under which the measurements were taken.

TABLE 3-10

STANDARD DEVIATION OF REPLICATE FIELD MEASUREMENTS BY SUBSTRATE, POOLED ACROSS MAP/XRF INSTRUMENTS

Data Source	Substrate	Std. Deviation of Replicate Measurements		
		0.6 shim	2.99 shim	Pooled
Daily	Wood	0.17	0.31	0.25
Validation	Steel	0.16	0.25	0.21
Data	Drywall	0.24	0.26	0.25
	Concrete		0.44	0.49
NIST	Wood & Plaster			0.3[a]

[a] Applies to lead concentrations below 2.0 mg/cm^2.

Testing the Linearity of Measurements with Lead Concentration

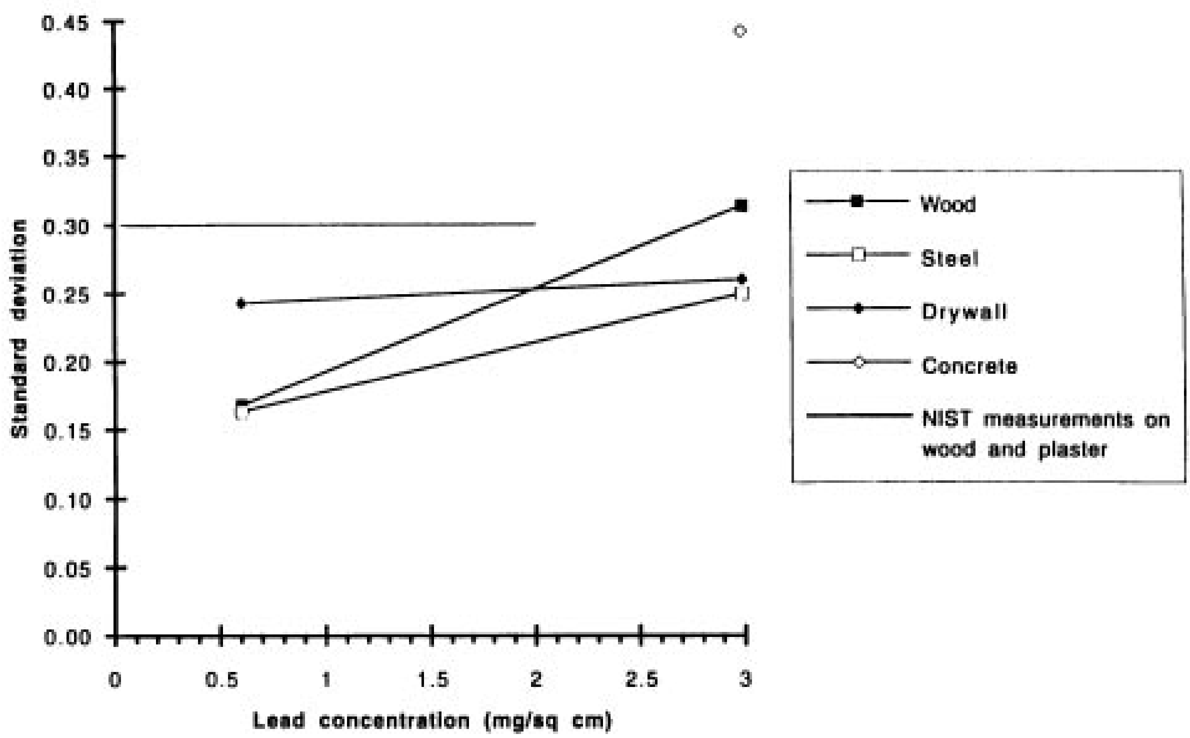
The existence of a linear relationship between the validation measurements and shim concentration was tested using available data for three lead concentrations: 0.0, 0.6, and 2.99 mg/cm^2. Because there were significant differences in the model between the Baseline, Daily, and Closeout periods, the results of this test were only approximate. Furthermore, sets of measurements without shims and with less than 20% zeros were available only for steel substrates with three instruments in the baseline period and four instruments in the closeout period. A quadratic curve was fit to these data (see Figures D-1 through D-7 in Appendix D). Although some quadratic terms were statistically significant, the direction of the curvature was not consistent across instruments, or between the baseline and closeout period for one instrument. Because of the limited data and inconsistent results, the assumption was made that the relationship between the field measurement and the shim concentration was linear.

A Formal Test of Factors

The effects of factors on the means of internal measurements were tested for significance using weighted regression on daily validation measurements for substrate-instrument-shim combinations where less than 20% of the measurement were censored. Factorial interaction terms were included for instrument, substrate, and shim concentrations. These terms were highly significant ($p<.0001$). These differences might have been due to differences among the substrate samples. Because some show quite different patterns on some substrates, it was assumed that there were significant differences between instruments beyond any variability due to the shims and substrates. Accordingly, a different model equation was fit for each combination of instrument and substrate.

FIGURE 3-2

POOLED STANDARD DEVIATION OF 60 SECOND VALIDATION MEASUREMENTS BY SUBSTRATE



The model also included a term for time. The rate of photon emissions is known to decay over time. With all instrument and substrates combined, the decay was statistically significant (p=.017), as expected. Furthermore, the rates of decay were not found to be significantly different across instruments or substrates. Therefore, a single formula was used to adjust all measurements for this decay. This was accomplished by including a single trend term for all instruments that represented time. The estimated change in the expected field measurement for each additional day in the field was -0.00099 mg/cm^2.

The Calibration Model

Based on the preliminary analyses described above, the calibration equation was assumed to have the following form:

$$Y = \alpha_{ij} + \beta_{ij} * X - 0.00099 * t,$$

where Y = the field measurement, X = the lead concentration, t = time (i.e. the number of days since 2/1/90), α = the intercept, β = the slope, and the subscripts i and j represent instruments and substrates, respectively.

Because the Daily validation measurements had measurements at only two shim concentrations, the model equations were simple to calculate. Since the regression line passes through the estimated means for the two shims (after adjusting the measurement for the time trend), it was a simple matter to calculate the estimates, a and b, of the model parameters α and β. The calibration equation was then determined by inverting the estimated model equation as follows:

$$x = \frac{(Y - a_{ij} + 0.00099 * t)}{b_{ij}} ,$$

where x is an estimate of the lead concentration X. Since the lead concentration cannot be less than zero, the Recalibrated measurement were calculated as

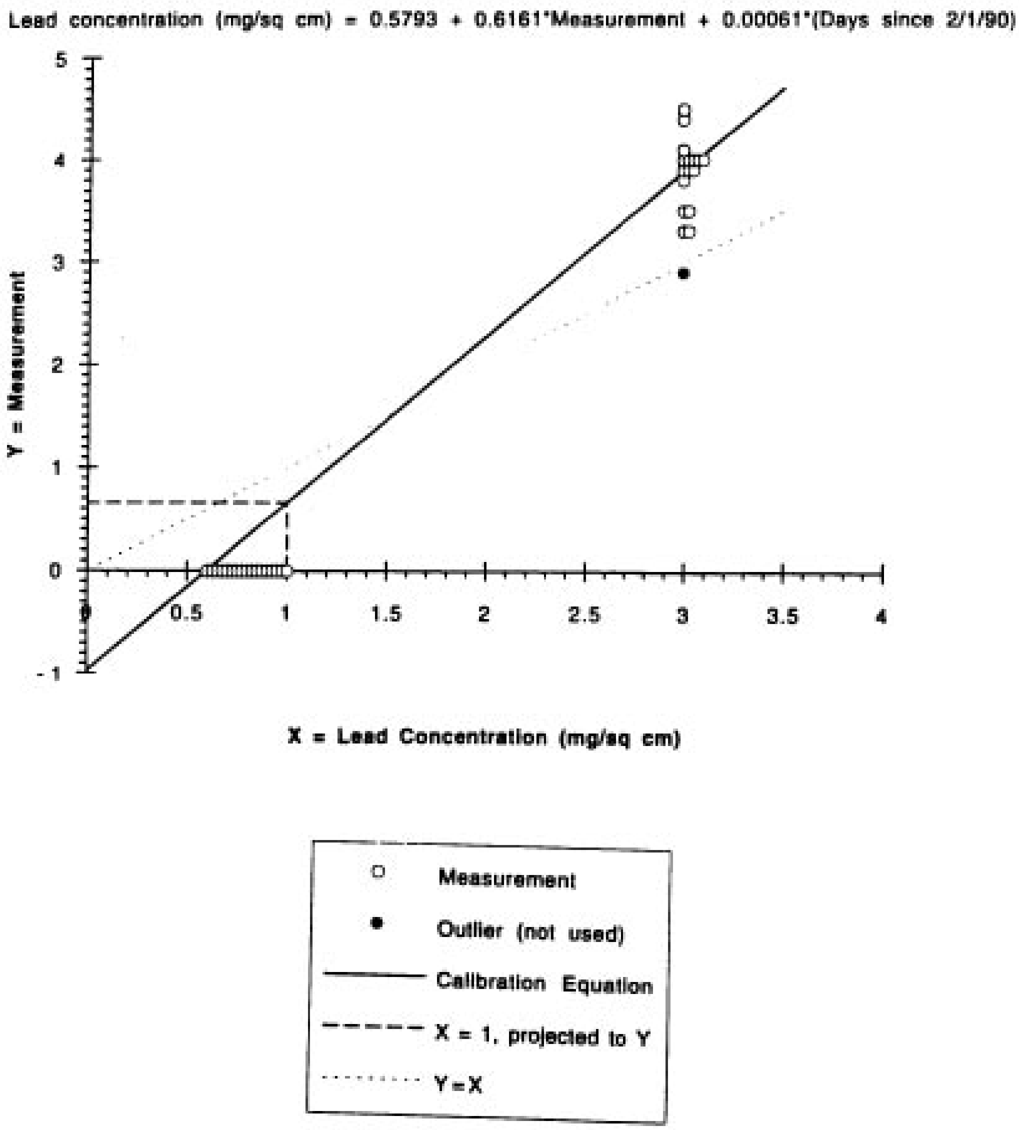
Recalibrated measurement = Maximum (x, 0).

Figure 3-3 illustrates the calibration equation for instrument #32 on steel. (Figures D-8 through D-38 in the Appendix D show the calibration equations for other combinations of instrument and substrate.)

- The **solid line** is the calibration equation.
- The **dotted line** is shown for reference purposes. This line has an intercept of 0 and a slope of 1 and shows the hypothetical case where no adjustment is necessary. The more the slopes of the solid and dashed lines differ, the more important it is to adjust for calibration.
- The **open circles** represent the actual measurements. They are plotted as a sideways histogram, rather than on top of each other, in order to reveal more detail in the plot.
- The **filled circles** represent outliers that were removed from the analysis.

FIGURE 3-3

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #2 ON STEEL



When X = 1, Y = approximately 0.65

Median used for 0.6 shim

- The **dashed line** is also given for reference purposes. This line illustrates the calibration adjustment for the special case where the true lead concentration is 1.0 mg/cm^2. This corresponds to a field measurement of about 0.68.

With censored validation data, there is a tendency for the Recalibrated measurement to have upward bias when the actual lead concentration is greater then 2.99 mg/cm^2. This problem is particularly severe for the calibration equation for instrument #34 on concrete. Thus, for this particular instrument and substrate combination, a different calibration procedure was used: the slope of the model equation was estimated as the average of the estimated slopes for the other seven MAP/XRF instruments when measuring on concrete.

Precision of the Recalibrated Measurements

Direct calculation of the precision of the recalibrated measurements is quite difficult because of the relatively complex procedure for defining the calibration equation. Therefore, simulations were used to estimate the precision of the recalibrated measurements. The simulations incorporate the variance in the nominal lead concentration of the shims, the internal measurements used to derive the calibration equation, and the field measurements which are to be recalibrated. Simulations were performed to answer the question: what are the mean, standard deviation, and percentiles of 1000 independent recalibrated measurements, each made on a surface with the same known true lead concentration.

One thousand simulations were performed for each of the following lead concentrations: 0.0, 0.4, 0.75, 1.0, 1.25, 2.0, 3.0, and 4.0 mg/cm^2. Five steps were used for each simulation:

1) simulate the true lead concentration in the two shims;

2) simulate 45 Daily validation measurements for each shim;

3) calculate the calibration equation;

4) simulate the field measurement corresponding to the true lead concentration in the measured surface; and

5) calculate the Recalibrated measurement for the field measurement from step 4 using the calibration equation from step 3.

The simulations required making assumptions about the true relationship between the lead concentration and the field measurements, which is the true model equation. A different true model equation was assumed for each substrate. As with the model equations, the true model equations were assumed to be linear. The true model equations were determined by pooling the measurements for each substrate across all MAP/XRF instruments and using the procedure defined above to derive the model equations. The measurement variance for measurements on each substrate are based on the variances shown in Table 3-10.

Figures D-39 through D-43 (see Appendix D) are sets of boxplots that show percentiles of the simulated recalibrated measurements. Each box covers the central 50% of the values. For some substrates and low concentrations, the lower percentiles of the adjusted lead measurements all fall at 0.6, causing the box to collapse halfway or entirely. The bar through the center of the box is plotted at the sample median. The simulated recalibrated measurements are clustered close to the dashed line (where the actual measurement = simulated measurement). This is convincing evidence of the effectiveness and validity of the recalibration procedures.

Table 3-11 presents the mean and standard deviations of the Recalibrated measurements calculated using the simulations. For concrete, unlike other substrates, the assumed true relationship between the field measurements and the lead concentrations is likely to be incorrect. This is because almost all the data for the 0.6 shim is censored (see the column under assumption 1 for concrete in Table 3-11). Therefore, the alternate assumption (assumption 2) also was made that the true model equation has a slope of 1.0 and passes through the median value for the 2.99 shim. The difference in the statistics in Table 3-11 for these two sets of assumptions demonstrates the uncertainty in the Recalibrated measurements for concrete, due to the censoring of the data.

TABLE 3-11

MEAN AND STANDARD DEVIATION OF SIMULATED RECALIBRATED MEASUREMENTS

True Lead Conc. (mg/cm^2)	Substrate									
	Wood		Steel		Drywall		Concrete (Assumption 1)		Concrete (Assumption 2)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0	0.60	0.00	0.12	0.16	0.60	0.00	0.70	0.26	0.60	0.00
.4	0.61	0.04	0.41	0.27	0.62	0.05	0.81	0.40	0.60	0.00
.75	0.78	0.16	0.74	0.29	0.77	0.17	1.02	0.54	0.61	0.09
1	1.01	0.20	1.02	0.30	1.00	0.20	1.16	0.61	0.64	0.17
1.25	1.28	0.22	1.26	0.31	1.27	0.24	1.33	0.68	0.69	0.26
2	2.04	0.27	2.02	0.34	2.02	0.28	2.06	0.82	1.46	0.77
3	3.07	0.39	3.01	0.41	3.07	0.39	3.09	0.87	3.13	1.01
4	4.11	0.51	4.02	0.53	4.08	0.51	4.16	0.95	4.92	1.32

Note: SD = Standard Deviation

3.3.2 Adjusting Recalibrated Measurements for Censoring Bias

Field measurements of zero are said to be censored because the corresponding original internal measurement could not be observed. In order to simplify the discussion that follows, distinctions are made between two types of Recalibrated measurements:

- **Censored Recalibrated** measurements are recalibrated measurements that correspond to field measurements that were zero. Most censored recalibrated measurements had values near 0.6 mg/cm^2.

- **Uncensored Recalibrated** measurements are recalibrated measurements that corresponded to field measurements that were not zero.

Figure 3-4 shows a histogram of the recalibrated measurements plotted on a log scale. The labels on the x-axis show the midpoint of every fourth interval. The gray bars show frequencies of the uncensored recalibrated measurements, and the stacked black bars show frequencies for the censored recalibrated measurements. The distribution of the uncensored recalibrated measurements above .63 is fairly regular in appearance. The irregular appearance of the lower portion of the histogram is due to the stacking of censored recalibrated measurements at specific values. Recall that a different calibration equation was developed for each combination of instrument and substrate. For each such combination, all of the zero (censored) field measurements were mapped into approximately the same value when recalibrated. Thus censored recalibrated measurements appear stacked at a limited number of values, each value corresponding to a combination of instrument and substrate.

Most censored recalibrated measurements corresponded to negative internal measurements, although certainly there were some that corresponded to zero internal measurements. Since the censoring process itself tended to increase the measurement, and never decreased it, it is easy to see how censoring induced upward bias in the field measurements. This upward bias carried over to the recalibrated measurements.

In the previous section, it was described how field measurements were corrected for calibration bias, applying the same calibration formula to zeros and non-zeros alike. In this section, the censored recalibrated measurements were singled out and statistically corrected for censoring bias. This was accomplished by replacing the censored recalibrated measurements with expected values of corresponding uncensored recalibrated measurements. The union of these expected values with the uncensored recalibrated measurements will be referred to as the corrected measurements.

Figure 3-5 shows a histogram of the corrected measurements. The black bars have been spread to the left and the histogram has a fairly regular appearance throughout its entire range. The gray bars are unchanged because the original uncensored recalibrated measurements needed no correction.

Determining Approximate Expected Values

As noted above, the censored recalibrated measurements were statistically corrected for censoring bias by replacing them with approximate expected values of corresponding uncensored recalibrated measurements. The distribution of the uncensored recalibrated measurements have an approximate lognormal distribution. If the parameters of this distribution can be estimated, then approximate expected values of all the ordered observations can be approximated as percentiles of the estimated distribution. The available data were censored. There are some well-established approaches to estimating parameters with censored data that include maximum likelihood estimation, probability plotting, and regression analysis.

A plot of the logarithms of the ordered uncensored recalibrated measurements versus expected standard normal order statistics will tend to be a straight line. Weighted regression was used to estimate the intercept and slope of the regression line, the weights being the inverses of the variances of the corresponding standard normal order statistics. The regressions were performed separately for each combination of instrument and substrate. Because our interest was in the distribution of the uncensored corrected measurements at low lead concentrations, and the lognormal distribution may not fit these measurements well at high lead concentrations, the regression was run using only uncensored recalibrated measurements below 4.0 mg/cm^2. This cutoff value was used because the relationship appeared reasonably linear in that range.

With the distribution of the uncensored measurements fully estimated, it was then possible to calculate expected values for all uncensored ordered measurements in the full sample. The censored recalibrated measurements were then replaced with these expected values which will be referred to as corrected recalibrated measurements.

FIGURE 3-4

HISTOGRAM OF RECALIBRATED MEASUREMENTS IN PUBLIC AND PRIVATE DWELLING UNITS

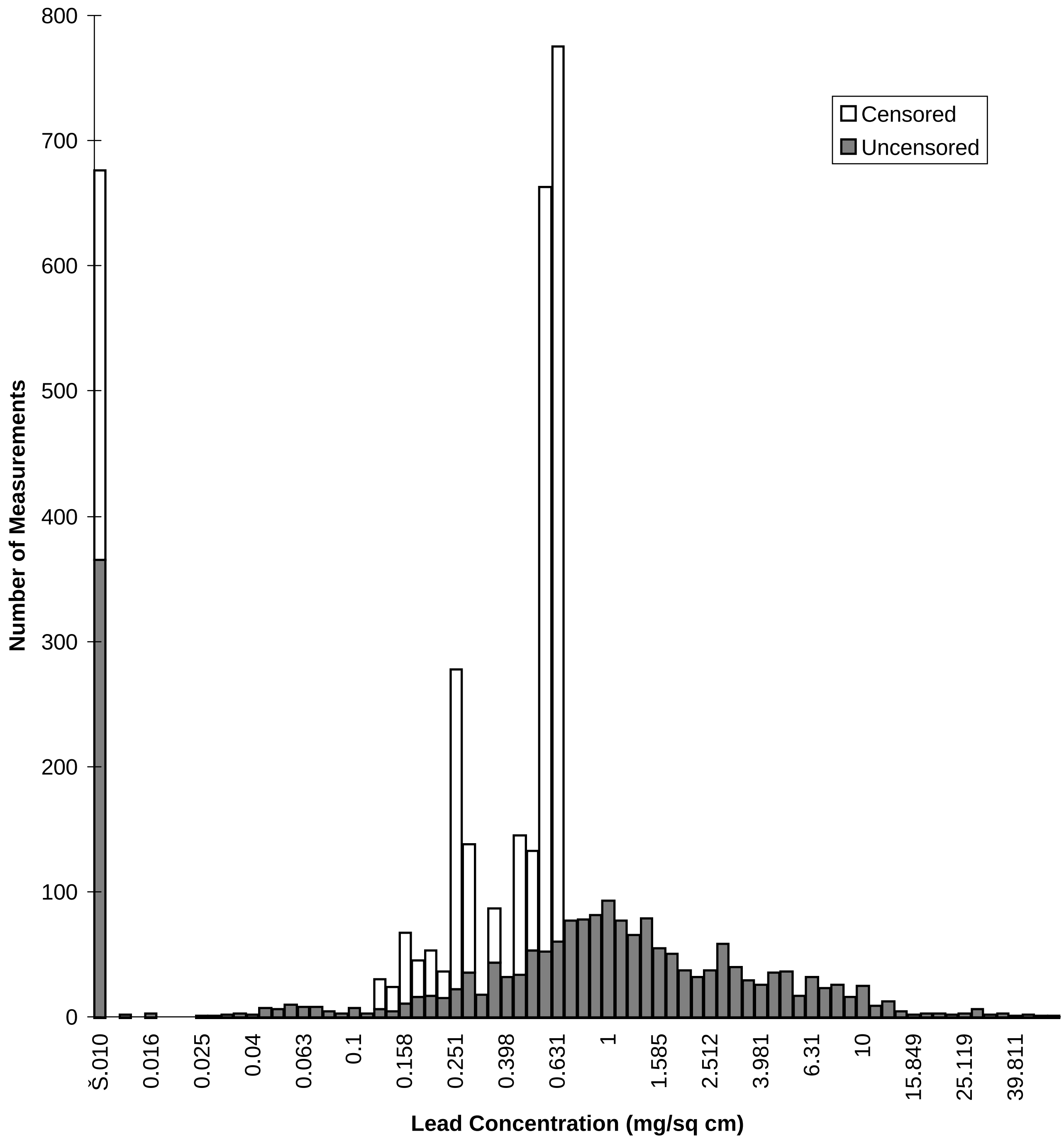
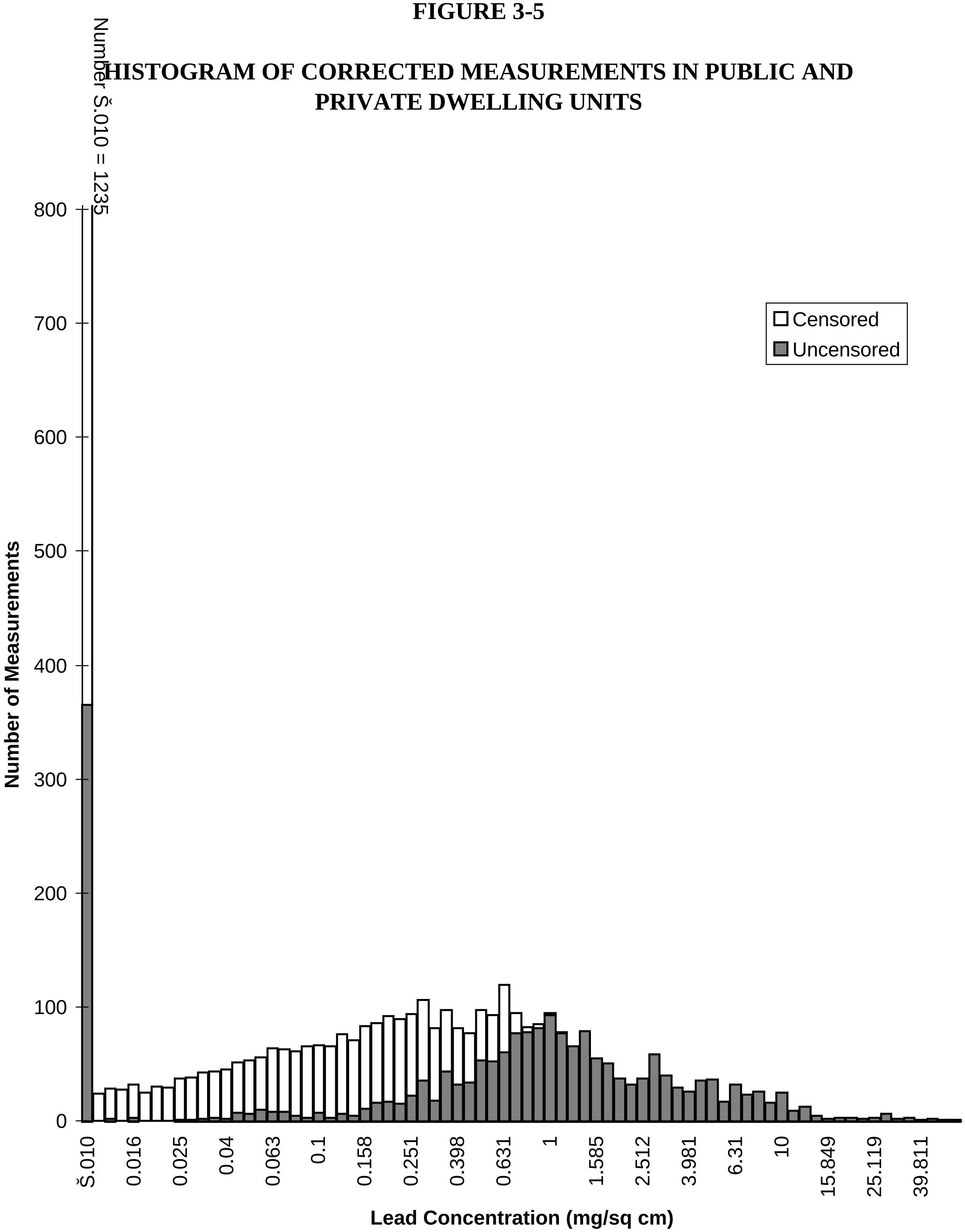


FIGURE 3-5

HISTOGRAM OF CORRECTED MEASUREMENTS IN PUBLIC AND PRIVATE DWELLING UNITS



Assigning Expected Values to Dwelling Units

A remaining task was to assign corrected measurements for the censored recalibrated measurements to specific dwelling units. While the corrected measurements have a natural order, the censored recalibrated measurements do not since, for a particular combination of instrument and substrate, all of the censored recalibrated measurements had about the same value.

One result of the simulation study, to be discussed below, was that corrected measurements on similar surfaces were more similar within dwelling units than between dwelling units. Therefore, the censored recalibrated measurements were ordered using the average estimated concentration of lead-based paint within a dwelling unit on surfaces in the same room with similar architectural type. In a few cases, the corrected measurement was undefined because the substrate was unknown, the MAP/XRF instrument serial number was missing, or there was not enough data to perform the regression analysis. In these cases, the corrected measurements were set equal to the recalibrated measurement.

3.4 Correcting for Bias in the National Estimate of Lead-Based Paint Prevalence

One of the primary objectives of the survey was to estimate the national prevalence of dwelling units with lead-based paint. This was done by computing the weighted proportion of sampled dwelling units that were classified as having lead-based paint. Each dwelling unit was classified as having lead-based paint on the basis of the maximum corrected measurement for that dwelling unit. But this sample maximum is a biased estimator of the maximum true lead concentration in the dwelling unit. This section describes the causes of this bias and the procedures used to correct for it in order to construct an unbiased estimate of lead-based paint prevalence.

3.4.1 Three Types of Estimates at the Dwelling Unit Level

National estimates based on the survey data are all weighted (inversely to the probability of selection) means of estimates made at the dwelling unit level. It is important to distinguish among three different types of estimates made at the dwelling unit level:

- **Averages of measurements**. Examples are the average (i.e., weighted mean) lead concentration in paint and the average surface area of lead-based paint in a dwelling unit.

- **Maximums of measurements**. One crucial example is the national estimate of the prevalence of lead-based paint. The maximum of multiple measurements is used to define a lead-based paint indicator for a dwelling unit. The reason for using the maximum stems from the following definition: a dwelling unit is said to have lead-based paint if any one surface in the dwelling unit has an average lead concentration greater than or equal to 1.0 mg/cm^2.

- **Classification Estimators.** The lead-based paint indicator defined above for a dwelling unit is a classification estimator since it classifies each sampled dwelling unit as having (or not having) lead-based paint.

The bias of an average of measurements is equal to the average of the biases of the measurements. Section 3.3 describes how calibration bias and censoring bias were removed from the individual measurements. Since the corrected measurements are unbiased (or have a very small bias), **estimates of average lead concentration and surface area in a dwelling unit are unbiased**. However, both maximums of measurements and classification estimators can be biased, even if the underlying measurements are unbiased..

Bias of the Sample Maximum

The bias of a sample maximum (for estimating the true maximum) depends on two additional factors, even if the corrected measurements are unbiased:

1) **Upward bias due to multiple miss-classification opportunities**. This kind of bias results from having measurement variance in conjunction with multiple measurement opportunities. To see how this bias arises, assume that a painted surface has a true lead concentration of 1.0 mg/cm^2. A single unbiased measurement with a symmetric measurement error distribution is equally likely to be above or below the true lead concentration, 1.0 mg/cm^2. However, the maximum of two independent measurements from this distribution has a 75 percent change of exceeding the true lead concentration ($1 - .5^2 = .75$). The maximum is therefore a biased estimator. With 12 independent measurements, the probability that the maximum exceeds the true lead concentration rises to $1 - .5^{12}$ or 99.98 percent.

2) **Downward bias due to incomplete sampling of rooms and surfaces**. By design, neither every room in a dwelling unit nor every surface in a room was sampled. This results in downward bias. To fix ideas, consider the extreme case when there is no measurement variance. Suppose a dwelling unit has 50 surfaces and the highest concentration of lead is on 5 of these surfaces (which could happen if they were all painted at the same time). If only one surface is randomly sampled, the probability of selecting one of these five surfaces is only 10%. If 12 surfaces are sampled, the probability rises to 69%. The probability reaches 100% only if 46 surfaces are sampled (since at least one must have the highest).[17]

Bias Due to Classification

Classification bias can occur when measurements are used to classify a unit. For example, a painted surface is classified as having lead-based paint if the measured lead content of the paint exceeds 1.0 mg/cm^2. Estimates of percentages derived from these classifications can be biased, even if the underlying measurements are unbiased. Classification bias depends on the density function of true lead concentrations in the vicinity of 1.0 mg/cm^2. If the density is increasing, then there tend to be more lead concentrations just above the threshold than just below. Because of measurement error, there is a tendency to make more Type II errors than Type I errors. This results in downward bias in our classification estimator. The reverse is true if the density is decreasing.

The Corrected Maximum Measurement

Although the three sources of bias discussed above will offset each other somewhat, it is virtually impossible that they will cancel each other out at the dwelling unit level. Furthermore, there was the possibility that the net bias in the national estimates could be substantial. Therefore it was important to develop a method to adjust the maximum corrected measurement for bias. Accordingly, the following definition is made:

- Corrected Maximum measurement is the maximum corrected measurement at the dwelling unit level after it has been further corrected for upward bias due to multiple miss-classification opportunities, downward bias due to incomplete sampling of rooms and surfaces, and classification bias when estimating lead-based paint prevalence.

---

[17] The test models the data with the hyper geometric distribution and uses Fisher's Exact Test. See *Introduction to Statistical Analysis,* by Wilfrid Dixon and Frank Massey, Jr. (1969), 3rd edition, page 243.

Purposive Samples

In an attempt to compensate for the downward bias due to incomplete sampling of rooms and surfaces, the sample design specified that purposive samples be taken. The first purposive sample was taken on the surface inside the dwelling unit which the technician believed to have the highest lead content. If that field measurement was zero, a second purposive measurement was taken inside the dwelling unit. One or two purposive samples also were taken on the exterior of the building, following the same sequential procedure. If the technicians were skilled at finding paint with the highest lead concentration, purposive measurements could have offset the downward bias due to incomplete sampling of rooms and surfaces. On the other hand, if technicians were unable to distinguish surfaces with higher lead content, the purposive measurements would be not much different than adding one or two more randomly selected surfaces.

In order to assess the effectiveness of the purposive measurements, purposive measurements were compared to non-purposive measurements in the same dwelling units, in rooms with the same wet-dry status, and from components with the same architectural type. The average difference was small and not statistically significant. This indicated that, within the precision which can be obtained with the survey data, once an architectural component and room type was selected by the technician, his ability to find a surface with the highest lead concentration was not much better than chance. (This analysis has not been performed.)

Regardless of the skill of the technician in finding lead-based paint, including the additional samples at the dwelling unit level reduces downward bias due to incomplete sampling of rooms and surfaces. But additional samples also increases upward bias due to multiple miss-classification opportunities. The net effect of purposive samples on the bias of prevalence estimates cannot be determined without further analysis. (This analysis has not been performed.)

3.4.2 A Simulation Approach

In order to correct for the bias in the national estimates of lead-based paint prevalence, it is necessary to estimate the biases caused by incomplete sampling of rooms and surfaces and by classifications based on measurements. This, in turn, requires knowledge of the distribution of lead loadings on all painted surfaces in a dwelling unit, including the means, variances, and correlations between different rooms and surfaces. Unfortunately, no applicable and adequate data set exists. While there are data sets with the needed data, e.g., the data from HUD's abatement demonstration project, none of these data are from general population samples of housing units that used the same methodology as the national survey to measure lead in paint. The only available alternative was to develop a simulation model and use it to correct for the bias in the national estimates of lead-based paint prevalence. The simulation also allowed assessment the effectiveness of the purposive samples. The model was developed using the following three steps:

(1) **Characteristics of non-sampled rooms and surfaces were simulated**. Room characteristics were known for sampled rooms, but not for non-sampled rooms. A hot deck imputation procedure was used to simulate characteristics in these non-sampled rooms. This completed the information on the population of surfaces and rooms from which would be sampled in the simulation process.

(2) **A simulation model was developed** that contained terms for all the components that were believed to determine measurement location and variance.

(3) **The model was calibrated** to be consistent with aggregate characteristics of the national survey data.

Hot Deck Simulation of Surfaces and Rooms

Imputation is a procedure by which missing data is replaced with data believed to be similar. The analytical process is usually simplified when there are no missing values. In this study, imputation made the simulations possible. There are many approaches to imputation. The simplest is to replace each missing value with the mean of comparable known values. One disadvantage of this approach is that it creates too much regularity in the data and variance estimates are often biased downward.

The imputation procedure chosen was the hot deck method. Hot deck imputation was chosen because it preserves both means and variances. With this method, missing data were replaced with actual data from the survey that came from a similar room in a similar dwelling unit. When choosing a similar room, the degree of similarity depended on: 1) the total number of rooms in the dwelling unit, 2) public or private status, 3) age category, 4) wet or dry status, 5) number of surfaces in the room, and 6) the most common substrates within architectural components. When multiple rooms were equally similar, one was chosen at random. A similar imputation was performed for common rooms where type of common room (laundry room, office, etc.) was also a consideration. For purposes of the simulations, data for purposive samples were retained and treated like other non-purposive samples.

Correlation Between Rooms

The question arose as to how to model the correlation of measurements taken in different rooms of the same dwelling unit. Such a correlation can result from using the same paint in multiple rooms over the painting history of the dwelling unit. Analysis of differences between the purposive and non-purposive samples for similar architectural components suggested that the magnitude of the variance component for differences between rooms was near zero. The same conclusion was reached when the differences between wet and dry rooms was assumed to be random rather than fixed. Data from the HUD Demonstration Project did not help answer the question because the measurement error was too large to permit the estimation between-room correlation.

While the paint on some surfaces in different rooms was the same, the choice of paint for the other surfaces was often made independently. The correlation of measurements taken in different rooms depends on the proportion of surfaces that were painted independently. Therefore, the model included an extra parameter[18] that represented this proportion. Because no data were available to support an estimate of this parameter, the parameter was set equal to 50% for the final simulations. However, the parameter was varied from 0% (paint chosen independently) to 100% (same paint used) in order to determine the sensitivity of the estimate of lead-based paint prevalence to the changes in the parameter. Results of these simulations indicated that the estimate was not very sensitive to the value of this parameter, except when it was close to zero. These simulations also provide a rough estimate of the component of variance in the estimate of lead-based paint due to the uncertainty in the value of the parameter.

[18] The parameter was incorporated into the simulations by performing a Bernoulli trial for each surface. Due to the nature of how it was incorporated, it does not appear explicity in the simulation model below.

The Simulation Model

Based on an accumulation of empirical evidence, the distribution of corrected measurements for lead-based paint is skewed to the right and can be fit reasonably well with a lognormal distribution. The simulation model had the following exponential form:

$$\text{Corrected Measurement} = e^{(\mu_{ikl} + \chi_{im} + \delta_{ijklmh} + I_p + \varepsilon)} + \nu$$

where:

μ_{ijkl} = mean of the log transformed lead concentration on surfaces of the same type.

χ_{im} = random dwelling unit component, normally distributed with variance σ^2_i which depends on the age category i.

δ_{ijklmh} = random surface component, normally distributed with variance σ^2_{ij}.

I_p = adjustment to account for the purposive samples which have a slightly different mean than the random samples.

ε = random within surface component, normally distributed[19] with variance σ^2.

ν = measurement variance associated with the MAP/XRF instrument, normally distributed with standard deviation[20] equal to (a_g+b_g*lead concentration).

h = index for an individual surface within a dwelling unit.

i = index for the vintage of the dwelling unit (Before 1920, 1920-1939, 1940-1949, 1950-1959, 1960-1969, 1970-1979).

j = index for the area type (interior, exterior, common rooms, playground).

k = index for the room type (wet, dry).

l = index for the architectural component (walls-ceiling-floor, metal substrates, nonmetal substrates, other).

m = index for dwelling unit within age category.

---

[19] The estimate of this parameter was obtained from an analysis of the HUD Demonstration Project data. The results of this analysis were sensitive to the handling of numerous outliers and are thus not precise. In the model, the standard deviation of the within surface variation was set to).30. We judged that this estimate was within a factor of two of the actual estimate. Therefore, simulations in which the within surface variance was changed by a factor of two were used to estimate the precision of the results due to the imprecision of the estimate of this parameter.

[20] The standard deviations show in Table 3-10 were used to develop a linear equation to estimate the unknown parameter, which depend on substrate g.

Most of the terms in the simulation model were selected after exploratory analysis of the survey data. Through this analysis, factors were identified that had a significant affect on the mean and variance of the corrected measurements, particularly in the vicinity of 1.0 mg/cm^2. The SAS procedure VARCOMP was used to estimate the proportion of the between dwelling unit variance which was associated with differences between surfaces. The estimates of variance components were obtained by subtraction.

As a check on the accuracy of the model, the actual survey data were compared to the simulated data and significant differences were found for certain characteristics. Calibration procedures were then performed that involved perturbing parameters until there were no longer any significant measurable differences. The main effect of this effort was to increase the between-surface variance relative to the between-dwelling-unit variance.

The simulations were repeated ten times for each dwelling unit. The simulation process itself contributed random error to the estimates. Simulation error can be reduced by increasing the number of simulations. But this requires additional computer time and resources. After 10 simulations, the simulation variance was approximated and it was determined that the benefits of additional simulations were negligible.

3.4.3 Adjusting the Maximum Corrected Measurement to Remove Bias in the Estimate of Lead-Based Paint Prevalence

The last step in the analysis was to determine the corrected maximum measurement. A linear adjustment was made so that the lead-based paint indicator based on the corrected maximum measurement was essentially unbiased for estimating the prevalence of lead-based paint.

A Linear Adjustment to the Maximum Corrected Measurement

As defined above, the corrected maximum measurement is the maximum corrected measurement at the dwelling unit level after it has been further corrected for upward bias due to multiple miss-classification opportunities, downward bias due to incomplete sampling of rooms and surfaces, and classification bias. The corrected maximum measurement allowed the prediction of which particular dwelling units had lead-based paint, and to cross-tabulate these predictions with other characteristics of dwelling units that were not incorporated into the simulations, such as the number of children in the dwelling unit.

The calibration of the model assured that there were no significant measurable differences between the actual survey and the simulated data at the aggregate level. Thus survey data characteristics drove the simulation model only through the formulation of the model and the determination of parameter values. When simulating measurements for a dwelling unit, no additional use was made of the actual survey measurements for that dwelling unit. In this sense, the simulated data were not paired to the actual survey data.

The maximum corrected measurement was computed at the aggregate level by transforming its percentiles to equal the percentiles of the maximum true lead concentrations that were developed from the simulations. A linear adjustment was deemed appropriate because the relationship between the percentiles of the distributions of the simulated and actual maximums was approximately linear in the range of 1.0 to 2.0 mg/cm^2. A different linear formula was determined for each vintage on dwelling unit.

After making this final correction, it was noted that the corrected maximum measurement was greater than the maximum corrected measurement. Recall that corrected maximum measurement is the maximum corrected measurement at the dwelling unit level after it has been further corrected for upward bias due to multiple miss-classification opportunities; and downward bias due to incomplete sampling of rooms and surfaces. Upward bias is corrected by decreasing the maximum corrected measurement. Downward bias is corrected by increasing the maximum corrected measurement. Since the net effect of the adjustment was an increase, this implies that the downward bias due to incomplete sampling of rooms and surfaces was greater than the upward bias due to multiple miss-classification opportunities.

For a few dwelling units, no corrected measurements were recorded. In the past these dwelling units have been treated as if they have no lead-based paint. The procedures used to calculate the predicted maximum average lead concentration make the same assumption. As a result, the adjusted estimates may slightly underestimate the proportion of dwelling units nationally which have lead-based paint.

Variance of the Unbiased Estimate of Lead-Based Paint Prevalence

The prevalence of lead-based paint is the proportion of dwelling units with lead-based paint. The variance of a sample proportion depends on the proportion itself, p. The arcsine transformation was used to stabilize the variance and make it independent of p. The variance of the transformed proportion is approximately 1/4n, where n is the sample size. This approximation was then multiplied by the design effect of 1.45. Finally, a constant was added to take into account the uncertainty of the model, simulation variance, and variance due to imprecision in our parameter estimates, particularly the within surface variance component and the effective proportion of surfaces with independent lead concentrations. This constant was estimated to be 0.00224 after computing the variance of results from multiple sets of simulations. This procedure yielded the following estimate of the variance:

$$\sigma_n^2 = \frac{1.45}{4n} + 0.00224$$

In terms of our lead-based paint prevalence estimate, 95% confidence limits are given by:

$$sin^2 \left(p \pm 1.96 * \frac{1.45}{4n} + 0.00224 \right).$$

Table 3-12 presents confidence intervals for selected values of p and n.

TABLE 3-12

NINETY-FIVE PERCENT CONFIDENCE INTERVALS FOR LEAD-BASED PAINT PREVALENCE BY PREVALENCE ESTIMATE AND SAMPLE SIZE

Prevalence Estimate	Sample Size		
	284	100	50
2%	0% to 7%	0% to 8%	0% to 11%
5%	1% to 11%	1% to 13%	0% to 16%
10%	4% to 18%	3% to 21%	2% to 24%
20%	12% to 30%	10% to 33%	7% to 37%
30%	20% to 41%	17% to 44%	14% to 49%
40%	29% to 52%	26% to 55%	22% to 59%
50%	38% to 62%	35% to 65%	31% to 69%
60%	48% to 71%	45% to 74%	41% to 78%
70%	59% to 80%	56% to 83%	51% to 86%
80%	70% to 88%	67% to 90%	63% to 93%
90%	82% to 96%	79% to 97%	76% to 98%
95%	89% to 99%	87% to 99%	84% to 100%
98%	93% to 100%	92% to 100%	89% to 100%

3.5 Laboratory Measurement Error and the Effects of Small Dust Sample Weights on the Findings

The objective of this section is to estimate the extent of laboratory measurement error associated with a number of sources, including dust sample weight, location of the sample, and analytical protocol. The data files constructed for this analysis consist of entries for samples analyzed by laboratories. The entries contain identification codes, site codes, laboratory codes, weight of the sample, kind of sample (dust or soil), location within the dwelling unit, micrograms of lead detected, weight of dust sampled, area vacuumed, and lead loading (micrograms of lead per square foot for the dust samples only). The analysis began with a review of the data file to determine if any data values should be trimmed from the file before proceeding.

3.5.1 Trimming the Dust Analysis File

Examination of the data showed that there were some unusually large values for both weight of the dust samples collected and for micrograms of lead measured in the dust. Some of the large weights are samples in which the dust was collected by a wet wipe (see discussion below). Some of the large readings may represent errors in data entry or recording, and some may represent actual dust samples with atypically heavy concentrations of lead. Also, when the laboratory results (micrograms of lead detected and weight of the sample analyzed) were converted to parts per million (ppm or micrograms of lead/grams of sample), a number of very large readings appeared. Even though some of the large readings could have been correct, it was judged that the review of dust samples having extremely large ppm of lead would contribute little to a judgment concerning the accuracy of the chemical analyses. Therefore, it was decided to leave all dust samples with lead values greater than 100,000 ppm out of the analyses and to consider other possible dust lead cutoffs based on lead loadings and sample weights as described below.

A few dust samples had very high lead loading readings (micrograms of lead per square foot of surface). This is an important statistic since it is probably the best measure of lead exposure to children under seven years old.[21] High values however, have little relevance in terms of the ability of the laboratories to detect levels of lead near threshold exposure levels (very low loading levels). For this reason, samples with lead loadings exceeding 2,000 micrograms per square foot vacuumed were dropped from the QC analysis of laboratory data.

Other dust samples eliminated from analysis were unauthorized wipe samples not collected according to the dust sampling protocol developed during the design phase of the survey. The protocol specified dust samples collected in homes with a vacuuming technique. A few however, were collected in homes with wet wipes when vacuuming was impossible (e.g., because no electricity was accessible) in order to avoid not collecting any dust samples at all. It was necessary however, to eliminate the wipe samples from analysis because their treatment during laboratory analysis resulted in substantial errors in the "total dust weight collected" determination. The cause of the error was that the weight of the wet wipes and the weight of the dust were reported as one number. Thus, dust lead concentration (ppm) determinations were impossible to calculate since the denominator (grams of dust + grams of baby wipe) was artificially large. This resulted in very small (usually fractional) ppm values. For this reason, and because of possible background contamination from the untested baby wipe commercial brands, the wet wipe samples were eliminated whenever they could be identified. Wet wipe samples also were eliminated from the dust analysis reported elsewhere in this report. The laboratory report showed that there were 81 such samples (out of 2,178 dust samples--see page 17 of the MRI Report). Of the 81 samples, only 36 are identified in the detailed laboratory data and another 17 were identified in internal analyses conducted for this report. One of the latter could not be located in the file.

After elimination of the identified wet wipe samples, samples greater than 100,000 ppm and lead loadings greater than 2,000 micrograms per square foot (see above), the remaining samples were classified by location within the homes, floors, window sills, and window wells. The average weights and 95th and 98th percentiles of weights were examined, as well as other possible cutoffs near the top of the scale. Based on the analysis, and because many of the wet wipe samples could not be positively identified, a decision was made to exclude all dust samples that weighed greater than 20 grams collected from floors, 5 grams collected from window sills, and 8 grams from window wells.

Table 3-10 shows the number of cases dropped from the analysis file for the various reasons specified above. Many of the 111 samples with missing data appear to have had sample weights too small to measure or at least too small to analyze. The actual samples are listed in Appendix C. A total of 1,974 dust samples remained in the file after the eliminations.

---

[21] Davies, D.J.A.; Thornton, I., Watt, J.M.; et. al: "Relationship between blood lead and intake in two year old urban children in the UK, "*Sci. of the Total Env.* 90:13 (1990).

TABLE 3-13

NUMBER OF DUST SAMPLES EXCLUDED FROM THE FILE BY REASON FOR EXCLUSION

Reason	Number of Samples
Single reason	
Missing data (for ppm or for lead loading)	111
Dust collected by wet wipe	21
Excessively large ppm	5
Excessively large lead loading	22
Dust sample weight too large (may include unknown wipes)	16
Two reasons	
Dust sample weight too large and dust collected by wet wipe	31
Dust sample weight too large and excessively large lead loading	1
Dust sample weight too large and missing data (for ppm or for lead loading)	1
Total	**208**

3.5.2 Trimming the Soil Analysis File

There was no need to truncate the weights of the soil samples since the analysis required approximately one gram of soil and generally, much more than this was collected in the field. There was a problem however, with large soil readings; some were as large as 43,000 ppm. Although there is no reason to believe that the large readings are not factual, such readings contribute little to an analysis of the precision of the laboratory work, as noted in the discussion above. Therefore, a decision was made to truncate the soil readings at 2,600 ppm because a natural break occurred in the data between samples with moderate and relatively high values. This cutoff only dropped 13 soil samples out of 869 in the file, leaving 856 for analytical purposes. The dropped samples are listed in Appendix Table C.

3.5.3 Laboratory Comparisons by Site for Dust Analysis

The laboratory analyses of the dust samples were performed by three laboratories: Midwest Research Institute in Kansas City, Missouri; Core Laboratories in Casper, Wyoming; and Core Laboratories in Aurora, Colorado. This section investigates whether each of the three laboratories, on the average, determined lead concentrations equally. If the samples were randomized in the field and sent in batches to the three laboratories equally, it could be assumed lead concentrations in the dust were equally distributed across the laboratories and a simple analysis comparing the results from all three laboratories would be possible. Therefore, an analysis across the 30 counties was conducted to determine how the dust samples were distributed among the three laboratories.

For the purpose of this analysis, the dust sample data was trimmed further to include only those samples with less than 5,000 ppm of lead. The reason for the exclusion of the larger values was that they distorted the comparison between the laboratories; a single large value could cause a laboratory average for a given county to vary by a factor of two. Even with this constraint, the variation between laboratories within a site was large because the distribution of samples from counties to the laboratories was uneven and the proportion analyzed by each laboratory in a single county was not equal. Table 3-14 shows that there was wide variation in the allocation of the dust samples sent to the laboratories (the samples were not randomized between laboratories). Even if the proportions had been equal, however, the large variation in dust lead content within each county would mask any possible distinction among the laboratories. The conclusion made from the analysis is that a determination of significant differences among the laboratories can not be made.

3.5.4 Laboratory Comparisons by Site for Soil Analysis

Soil samples were only analyzed by two of the laboratories mentioned above--Core Laboratories in Casper and Midwest Laboratories in Kansas City. Table 3-15 shows the total number of soil samples included in the analysis, the average amount of lead (in ppm) and the standard deviations for the two laboratories conducting the analysis. Table 3-12 shows that samples were not randomly distributed among the laboratories, therefore, no conclusions about the accuracy of the laboratory work can be drawn from this analysis for the same reasons as mentioned above.

3.5.5 Evaluation of the Laboratory Dust Analysis Method

Graphite furnace atomic absorption (GFAA) spectrometry was chosen for the analysis of dust samples over other methods, specifically over the less expensive inductively coupled argon plasma spectroscopy (ICP), because GFAA is more sensitive -- it can detect and quantify lower levels of lead. This decision, made in the design phase of the survey, was based on pretest information that suggested small dust sample sizes could be expected and that possible low dust lead concentrations in the samples should necessitate the most sensitive analytical method available (GFAA). Because GFAA is considerably more expensive than ICP analysis, a question arises as to what data would have been lost if the less expensive method had been used. By looking back at the data, and by using pretest information on minimum sample weights required for ICP analysis, an evaluation of the quality of data is possible. The number of samples from the trimmed file of 1,974 samples that would not have met the minimum dust weight requirement for analysis by the ICP is 210, or about 11 percent. How different are these samples in terms of ppm and lead loading?

The samples too small for analysis by the ICP method had smaller average concentrations (678 ppm) than the total dust sample file (1571 ppm), although this particular measure is subject to high variation for small samples as discussed in the next subsection. The samples that could not have been analyzed by the ICP method represent cases in which the lead loading (the measure probably most closely related to exposure) is quite small. These samples had a mean loading of 0.2 µg/sq ft versus an overall mean of 41.6 µg/sq ft. The maximum loading of these samples is about 1.3 micrograms of lead per square foot vacuumed. Compared to HUD dust lead guidelines, these small loading values are trivial and most likely do not represent a hazard. Since loading is reflective of the amount of dust collected, the tautology (i.e., a clean house has little dust and hence less dust lead exposure), is repeated here. In any case, it appears that little would have been lost in the identification of hazardous levels of lead in housing if the ICP method had been used in place of the GFAA method, but national estimates of the prevalence of dust lead levels would not have been as accurate.

Table 3-14

NUMBER OF DUST SAMPLES, AVERAGE PPM OF LEAD AND STANDARD DEVIATION BY LABORATORY AND SITE, FOR SAMPLES WITH LESS THAN 5,000 PPM OF LEAD

	Number of Samples				Averages				Standard Deviations		
PSU	C	A	M		C	A	M		C	A	M
1	43	0	0		232				411		
2	35	15	0		515	706			455	1,202	
3	38	60	0		725	444			720	664	
4	17	70	0		401	344			449	702	
5	0	55	0			393				446	
6	41	33	0		669	379			927	337	
7	34	24	0		937	935			922	1,144	
8	33	0	0		501				734		
9	35	9	0		887	1,081			1,182	1,186	
10	43	60	0		1,080	696			1,150	929	
11	0	75	0			542				775	
12	0	32	4			1,020	1,264			1,148	657
13	47	5	0		629	1,937			1,118	1,134	
14	5	33	0		935	471			1,096	907	
15	51	0	41		579		649		734		1,022
16	44	13	0		442	450			240	305	
17	57	31	0		519	262			668	274	
18	41	118	0		1,161	566			1,170	798	
19	51	33	0		754	626			708	761	
20	21	0	0		820				976		
21	0	50	0			591				679	
22	17	41	0		749	516			1,124	913	
23	4	53	0		1,598	747			1,370	782	
24	20	39	0		1,191	929			1,227	1,,127	
25	45	9	39		612	430	594		890	484	
26	15	25	0		505	511			832	722	
27	35	0	0		659				970		
28	24	42	0		415	693			330	920	
29	23	17	0		869	565			1,099	856	
30	24	0	0		1,139				865		
All Sites	843	942	84		700	585	653		863	806	857

NOTE:

C=CORE LAB, CASPER, WY
A=CORE LAB, AURORA, CO
M=MRI, KANSAS CITY, MO

Table 3-15

NUMBER OF SOIL SAMPLES, AVERAGE PPM OF LEAD, AND STANDARD DEVIATION BY LABORATORY AND SITE, FOR SAMPLES WITH LESS THAN 2,600 PPM OF LEAD

	Number of Samples		Average ppm		Standard Deviation of ppm	
Site	C	M	C	M	C	M
1	28	0	28.97		32.03	
2	30	0	371.98		530.95	
3	38	27	226.03	154.21	348.73	191.85
4	0	48		196.63		403.79
5	34	0	105.52		166.58	
6	23	0	67.15		52.84	
7	14	0	391.61		592.56	
8	17	0	97.37		225.10	
9	35	0	261.39		637.93	
10	0	53		341.16		551.01
11	29	0	119.17		148.30	
12	18	0	479.30		440.15	
13	36	0	108.88		210.71	
14	25	0	104.35		296.89	
15	21	39	30.57	242.19	43.82	222.76
16	2	0	452.88		74.50	
17	14	14	112.07	160.17	195.71	89.86
18	3	1	666.24	974.44	418.44	
19	0	2		852.08		115.13
20	12	0	68.83		99.54	
21	30	0	155.87		320.64	
22	36	0	38.36		91.34	
23	12	0	139.62		92.05	
24	23	0	285.09		453.38	
25	24	30	99.33	85.24	151.21	72.49
26	32	0	36.02		23.65	
27	21	0	137.55		370.50	
28	27	14	38.44	147.64	25.37	160.97
29	24	0	193.44		332.93	
30	20	0	384.09		583.36	
All Sites	628	228	160.2	97.12	320.09	112.29

NOTE:

C=CORE LAB, CASPER, WY
M=MRI

3.5.6 The Problem of Small Dust Weights

The Midwest Research Institute laboratory report states on page B-15, "A minimum of 10 mg of dust will be needed to achieve method detection limits suitable for the data quality objectives of this survey." Many of the dust samples however, weighed less than this minimum and were analyzed. Of the 1,974 dust samples in the analysis file, 669 or 35 percent weighed less than 10 mg and 448 or 23 percent weighed less than 5 mg. These fractions were consistent among the three laboratories. This subsection examines the ppm and lead loading for the smallest samples, i.e., those under 5 mg. The weights of the dust samples were "tap weights" obtained by tapping out the dust from the collection container. This method does not yield reliable weight data, especially for small dust samples.

Table 3-16 provides a summary by weight of the 448 samples under 5 mg. Note that the 51 smallest sample weights were presumably, rounded to 0.1 mg. Only one of these samples with 0.1 mg of weight was deleted from the analysis file, though (reading number 116 in Appendix Table C-1), and that sample was dropped because of an excessively high ppm reading value.

The average ppm in the total analysis file equaled 1,571. Table 3-16 shows that most of the ppm averages for samples under 5 mg exceed that average. This is not surprising in view of the smallness of the denominator in the ppm computation and the uncertainty with which that denominator is measured.

The lead loadings per square foot are more critical to the objectives of the study, however, and they tend to be quite low. Only five measurements among the samples weighing less than 5 mg exceeded 50 micrograms per square foot of surface. The five readings are 57, 58, 71, 86, and 152. Between the sample weights of 5 mg and 10 mg the samples representing more than 50 micrograms per square foot are 57, 75, 87, 88, 89, 287, and 529. Adherence to the 10 mg cutoff would have eliminated several sample cases that appear to be significant. The real and unknown issue however, is how accurate the readings are for such small sample weights. This issue could be explored further by comparing the post digestion duplicate results for small dust-weight samples. Yet, this comparison could not be made because the post-digestion readings for the duplicate procedures were not identified by weight of the dust sample in the analysis file.

TABLE 3-16

DISTRIBUTION OF NUMBER OF DUST SAMPLES, AND AVERAGE AND STANDARD DEVIATIONS OF PPM AND LEAD LOADING PER SQUARE FOOT (µg/ft^2) FOR SAMPLES WEIGHING LESS THAN 5 MILLIGRAMS

		ppm		Lead Loading per sq. ft.	
Dust weight	No. of samples	Average	Std. Dev.	Average	Std. Dev
0.1	51	7,185	10,192	0.59	0.97
0.2	13	2,397	2,795	0.29	0.39
0.3	18	6,231	11,843	1.16	2.10
0.4	14	1,378	1,130	0.66	0.87
0.5	13	5,295	11,218	1.82	5.17
0.6-1.0	51	4,116	11,957	1.86	4.19
1.1-1.5	50	1,687	2,753	1.88	3.65
1.6-2.0	42	3,674	13,813	6.37	23.93
2.1-2.5	42	1,319	3,046	3.36	11.09
2.6-3.0	38	986	1,903	1.27	2.68
3.1-3.5	35	3,874	16,543	4.71	15.23
3.6-4.0	35	886	1,226	4.07	10.80
4.1-4.5	19	1,228	2,956	5.38	14.32
4.6-4.9	27	986	1,452	3.61	6.29
All wts.	448	3,026	9,033	2.77	10.37

3.5.7 Quality of Laboratory Data--Conclusion

The laboratory work was carefully done, both in terms of the design of the accountability, the quality control procedures, and the execution. The resulting high data quality allows for meaningful statistical analysis to predict national estimates of residential lead in dust and soil.

4. CONCLUSIONS

4.1 Major Conclusions

This section presents a brief discussion of the major conclusions to be derived from the experience of the national survey.

4.1.1 Study Findings

Lead-based paint is widespread in housing. An estimated 64 million homes, 83 percent of the privately owned housing units built before 1980, have lead-based paint somewhere in the building. Twelve million of these homes are occupied by families with children under the age of seven years old. An estimated 49 million privately owned homes have lead-based paint in their interiors. There are no statistically significant differences in the prevalence of lead-based paint by type of housing, market value of the home, amount of rent payment, household income, geographic region or degree of urbanization.

Thirteen million homes - 17 percent of the pre-1980 stock - have dust lead levels in excess of the federal guidelines, regardless of whether or not they have lead-based paint. However, excessive dust lead levels are associated with the presence of damaged lead-based paint. Fourteen million homes, 19 percent of the pre-1980 housing stock, have more than five square feet of damaged lead-based paint. Nearly half of them (47 percent) have excessive dust lead levels.

While a large majority of pre-1980 homes have lead-based paint, most of them have relatively small amounts of it. The average privately-owned housing unit with lead-based paint has an estimated 601 square feet of it on interior surfaces and 869 square feet on exterior surfaces. Over half of the leaded paint is on walls, ceilings, and floors. The amounts of lead-based paint per housing unit vary with the age of the dwelling unit. Pre-1940 units have, on average, about three times as much lead-based paint as units built between 1960 and 1979.

Lead paint is even more widespread in public housing; 86 percent of all pre-1980 public housing family units have lead-based paint somewhere in the building. While most public housing units have some lead-based paint, most of them have small amounts of it. The average public housing unit with lead-based paint has an estimated 367 square feet on interior surfaces and 133 square feet on exterior surfaces. Most of the interior lead-based paint is on walls, while very little of the exterior walls are painted.

4.1.2 Impact of Measurement Error and Lead Concentration Variations on the Data Analysis

The data analyses and reports of findings should incorporate instrument and laboratory measurement error.

The spectrum analyzer MAP/XRF instrument produces readings with measurement errors. They are systematically different from the actual lead concentrations in the painted surfaces and have random variation. Similarly, the laboratory protocols used to measure the lead in dust and soil have measurement errors. These measurement errors can induce systematic errors in the estimated extent of the lead hazard from paint, dust and soil, and can also result in underestimates of the uncertainty in the estimated hazard.

Therefore, the field procedures must provide for the collection of the QA data necessary to estimate the measurement errors and their impact on the study findings. Further, the data analyses must explicitly estimate and correct for the impact of the measurement errors.

The data analyses and reports of findings should take account of the inherent variation in the painted surfaces.

In the national survey, it was not possible to test for the lead content of every painted surface in every sampled housing unit. Consequently, surfaces were sampled, as described in Appendix I. This sampling protocol was designed to control the project costs and respondent burden by controlling the amount of time required for an inspection. However, it did not adequately provide for the estimation of variation in lead concentrations within surfaces, between surfaces in the same room, and between rooms. These sources of variation need to be addressed to produce accurate estimates of the uncertainty in the national estimates.

It is therefore recommended that multiple MAP/XRF readings be taken at randomly selected locations on a subset of the selected surfaces; at two or more components of the same type in the same room; and in two wet rooms and two dry rooms. Although it may not be feasible to do this in all homes, a subset should be selected for these additional readings.

4.1.3 Use of the Spectrum Analyzer MAP/XRF

In contrast to the HUD Interim Guidelines, substrata correction is a necessary step in the accurate determination of the presence and amount of lead-based paint on surfaces, when using the MAP/XRF.

In the national survey, the MAP/XRF generally produced readings that were systematically different from the amount of lead in the paint being tested. The direction and magnitude of the systematic differences were related to the substrate material, the lead loading in the paint, and, to a lesser extent, the age of the Co^{57} source. The exact nature of the relationships varied significantly from one individual MAP/XRF to another. Furthermore, the precision of the readings depended on the substrate. Therefore, substrate corrections are needed to obtain accurate measurements of lead loadings. There are two possible ways to do this:

1. Take frequent validation readings, as described in Appendix I, and analytically correct the readings using methods as described in Chapter 3. Readings need to be done on three or more different shims however, not just two shims as in the national survey. With only two shims, only a linear model can be used to correct the MAP/XRF readings. It is possible however, that a non-linear model better describes the relationship between the readings and the actual lead concentrations in the painted surfaces.

2. Perform substrate corrections in the field. The HUD Interim Guidelines describe substrate correction procedures appropriate for direct reading MAP/XRFs, not the spectrum analyzer MAP/XRF. At present, HUD is developing field substrate correction procedures for the spectrum analyzer MAP/XRF.

4.2 Additional Conclusions and Recommendations

The main purpose of this section is to identify lessons learned during the conduct of the National Survey and to develop recommendations for future field operations.

Objective of Recommendations

The objective of many of the following recommendations is to improve the representativeness and statistical validity of the data, e.g., develop methods to enhance respondent participation. Other recommendations concern the logistics of moving inspection teams around the country in an efficient and cost-effective fashion. Recommendations for improved in-home protocols are aimed at ensuring that inspections are conducted correctly and completely. In an effort to improve lead-based paint testing, a full section is devoted to recommendations concerning the use of the MAP/XRF.

Applicability of Recommendations

The National Survey operated under constraints that are not typical of other lead-based paint studies. Therefore, some of the following recommendations would not be appropriate under other research designs. For example, most lead studies are conducted in a single location, alleviating the problems associated with moving teams and equipment around the nation under a tight schedule. The units in the National Survey were randomly selected from the national population of dwelling units; most other studies limit themselves to a smaller, targeted population of units which have some underlying rationale for being included, e.g., they were FHA-mortgaged properties or they were in a targeted area of an inner city. A rationale for each recommendation is provided and the reader must use his or her judgment in determining if the recommendation is appropriate for another design.

The recommendations presented below stem from lessons learned in the pretests and in implementing the final design. They are grouped and discussed under the following headings:

1. Background of Field Operations Procedures
2. Pretest
3. Sample Frame Development for Private, Multi-family, and Public Housing
4. Field Activities
5. Dwelling Unit Visit and Inspection Protocol
6. In-Field Environmental Sampling
7. Use of the Spectrum Analyzer MAP/XRF

Background of Field Operations Procedures

Plans for field operations were designed and pretested as part of a separate contract effort in advance of the National Survey. Based on the results of the pretest, a final survey design and field recommendations were issued. The contractor for the full National Survey evaluated the design and recast certain portions to accommodate changes in research objectives. In addition, the survey contractor developed the schedule, budget, and detailed field procedures that were not present in the original design. The revised plan was pretested and further modifications were made based on the pretest results. The most significant and far-reaching change was the decision not to sample dwelling unit rooms and architectural features (components) in the field. Sampling was done based on the information gathered in a telephone interview. This allowed the inspection team to complete work in one visit per dwelling unit as opposed to the two visits (i.e., one for statistical sampling, one for inspecting and environmental sampling) called for in the original design.

Pretest

- **Pretests should be performed that test all study design features and technologies employed in a study and pretest subjects should represent the diversity of respondents and situations expected in the full study.**

For field studies, pretests done in diverse settings are essential to evaluating the study design and technologies employed. Although the initial pretest of the original design rendered many valuable lessons, a number of aspects were not thoroughly pretested prior to making the recommendations for the original design. This led to a number of unanticipated problems, specifically:

- The initial pretest included a large percentage of vacant units. This tended to present the field problems involved with resident contact.
- The initial pretest sample lacked diversity of circumstances, e.g., working in bad weather or at night.
- The initial pretest created lists of dwelling unit owners/managers of rental properties by reviewing telephone books and working with owner/managers to reach rentals. That was an impractical solution given the time and cost constraints of the full study. The sampling approach used in the National Survey (contacting dwelling unit resident first) was never pretested.

Sample Frame Development for Private (Single and Multi-family) and Public Housing

- **All listing and screening should be completed before beginning unit inspections**

Completion of listing and screening before beginning unit inspections means that the full sample frame can be developed before sampling is begun. Because of the tight field schedule of the National Survey, it was necessary to begin inspections before listing and screening were completed in some counties. Although it was possible to take samples from completed counties based on estimates of the final national distribution of eligible housing, it was necessary after screening was totally completed to adjust the final sample to the actual distribution. There were two ways to accomplish this: retroactively adjust the sample in the counties which had been previously sampled; or sample the remaining counties to balance the sample to the national distribution. Although the former approach yields a better sample statistically, it would have created costs that were deemed to outweigh the gain in statistical power. The National Survey adopted the second course, adjusting the sample in the counties that had not yet been sampled.

If schedule and budget permit, it is desirable to complete all listing and screening before beginning the sampling process, to produce a superior statistical sample.

- **When screening vacant and high security buildings, additional methods (other than knocking on the door) of reaching dwelling unit residents should be in place.**

A number of buildings sampled for screening were inaccessible because of high security measures, uncooperative doormen, and unknown/inaccessible management companies. The field staff did not have time to develop means of accessing these buildings (e.g., find owners and convince them to participate). Inaccessibility resulted in having to substitute dwelling units. The negative consequences could have been lessened if clerical staff had been in place at the listing/screening stage to lend assistance to field interviewers in contacting owners and management companies. "Crisscross" directories linking addresses to phone numbers could be referenced, for example. If a phone number is available for a dwelling unit, the resident can be contacted and screened over the phone.

- **During the initial listing and screening period, increased time should be allowed for listing and screening of dwelling units in rural areas.**

The sampling plan required the listing of *all* homes in the selected segments. The number of dwelling units listed in a county varied from a low of 220 in Cascade, Montana (a small rural county), to a high of 3,239 in Fairfax, Virginia (a large urban county). The time required to list all homes varied immensely, not in relation to the number of dwelling units, but in relation to the area of the county in square miles. Small rural counties took about 50 percent longer to list than the more densely populated urban counties. Similar increases in "time to complete" were found for the screening phase, as well.

- **Increased time should be allowed for listing and screening of dwelling units if weather is inclement.**

Listing and screening of dwelling units requires that the interviewers spend a large amount of time outdoors. Their progress will be substantially delayed if there is serious inclement weather. During bad weather, the Survey had a higher rate of illness and attrition rate among interviewers. A significant number of new interview staff had to be recruited, and in January, six weeks after the start of the field period, a second 3-day training session was held.

Attempts to screen during holidays were also unproductive. On the days near Thanksgiving and Christmas, there was a notably higher rate of incomplete screener interviews due to people not at home and refusals to take the time to answer screener questions. Interviewers were less willing to work over the holidays, and the request that they do so contributed to the loss of some of the original field interviewers.

The net result of working during holidays and inclement weather was a higher non-participation rate among dwelling units and waste of field staff effort.

- **Homes with two construction dates should carry the oldest date in the records.**

Some homes had a section built many years after the original construction date. In some cases a newly added room may have been added after the cut-off date for inclusion in the study (1979). The home was entered into the data base using the earliest date.

- **For public housing dwelling unit sampling, several additional weeks should be allowed for dwelling unit sample selection in large PHAs.**

The larger the PHA, the longer time it took to select the dwelling units. The reason varied, e.g., the contact was on vacation, there was difficulty finding the contact, the contact needed authorization, etc. Without exception, selection of PHA dwelling units in the five largest PHAs in the sample took several weeks longer than it did for the 25 other PHAs.

Field Activities

- **When dealing with PHAs, interviewers should closely coordinate with the PHA representative.**

Coordination with the PHA management staff is imperative to effective sampling and inspection of public housing. To ensure that all responsible and affected parties are involved in the process, it is advisable to establish a liaison with both a PHA headquarters representative and the manager of the specific housing project.

Often teams met the PHA representative at the PHA office and then went to the site. The PHA staff "smoothed" the way many times. In neighborhoods where the inspection teams felt unsafe, PHA staff provided a vital service as escorts .

Because, the PHA staff are busy, every effort must be made to accommodate their involvement. For example, PHA offices often are not located close to the dwelling unit. Inspection teams should schedule their time to allow them to meet the PHA staff at the PHA office and go from there to the dwelling unit. The PHA representatives typically tried to inspect as many units on each trip as possible. This consideration, combined with frequent unexpected delays, suggests that the inspection team should not schedule other inspections too soon after the PHA work, to avoid schedule conflicts.

Upon meeting a PHA representative in the field, the inspection team should not assume that he or she has been fully briefed on the study or the team's mission. The interviewer needs to go over the objectives and procedures of the survey and the importance of the PHA dwelling unit inspections.

- **Field staff should be trained to respond to anyone's questions concerning what they are doing and what the study is about. Do not attempt to use local police or similar agencies, or the regional HUD office, to substantiate or verify presence or activities.**

In the National Survey each field person carried an ID badge, copies of the license that allowed him to handle the MAP/XRF and its radioactive source, endorsement/introductory letters from survey and agency principals, the 800 number for the survey coordinator, and a number for the Washington HUD representative. Field staff were questioned on several occasions by police officers, building mangers, or nearby residents for walking around a house after dark to take the MAP/XRF readings and soil samples (the MAP/XRF looks like a gun, especially at night). The documentation materials listed above explained and substantiated the activities of the inspection team and proved invaluable.

Although it seems like a good idea, contacting local police in advance and then referring questioners to the police can lead to a bigger problem; i.e., the police contacts inadvertently denying knowledge of the survey. If the person who gets the call at the police station has not been apprised of the study, he or she will deny knowing anything about it, even though the police department has been informed of the activities. There is also a problem with overlapping jurisdictions. The study might notify the county sheriff, although the inquiry comes into the state police.

- **The inspection team should have special training and support in working in potentially unsafe neighborhoods. Two sub-recommendations follow. One concerns the safety of the inspection team, the second the safety of the equipment.**

It was reported in both the initial pretest of the original design and in the National Survey that inspection teams were reluctant to enter certain public housing. The issue was not public housing, per se, but one of being a stranger in a potentially unsafe neighborhood, carrying expensive equipment and driving a rental car. Having a PHA official escort the inspecting team significantly helped to alleviate these concerns. Official escorts also meant that the visit took place during working hours, Monday through Friday. Hence, lessened anxiety when escorted may have been a function of the time of day much as the escort itself. Visits in the evenings and on the weekends were regarded quite differently by the field teams.

- **Safety of Inspection Team: Potentially unsafe neighborhoods should be identified at the screening stage. Trips should be scheduled to occur during daylight hours on weekdays; operational and budgetary provision should be made for a paid escort (e.g., off-duty police officer) to accompany the inspection team.**

Operating in potentially unsafe neighborhoods proved very problematic for the two person inspection teams. Listing and screening were more problematic because the interviewer, usually a woman, was working alone. Screening and inspection completion rates were low in unsafe neighborhoods. Accommodations requested by the field staff included: extended time in the county so appointments could be scheduled exclusively during weekday, daylight hours; and paid escort, e.g., an off-duty police officer. Addressing the perceived and real dangers to the field staff is essential if they are to produce complete results in certain neighborhoods.

- **Safety of Equipment: The within-unit procedures need to be sequenced so all tests with one set of equipment are completed before the other set is needed.**

In regard to the safety of the equipment, there are two issues: the safety of the equipment (MAP/XRF, vacuum, supplies, etc.) from being stolen or damaged, and the safety of the general public if unauthorized persons tamper with unattended equipment.

The equipment in use was considered adequately supervised. Equipment in a car trunk also was generally secure. Therefore, the difficulty lay in assuring that only the set of equipment currently being used (MAP/XRF or vacuum) was out of the car.

The easy part of the solution was to remove only one set from the car trunk, perform all sampling requiring that equipment, return it permanently to the car, and retrieve the next set. Unnecessary trips to the car wasted time, disconcerted the residents, and brought undue attention to the activity. The described approach minimized trips back and forth to the car. However, difficulties with this approach arose when:

- There were several common areas to be inspected, e.g., playgrounds, laundry room.
- The inspector had to park the car some distance away, forcing him to spend the time walking and putting himself at risk carrying equipment around unsafe areas.
- The equipment was in jeopardy of being abused or stolen while in the car.
- The equipment was at risk of being damaged while in the trunk (e.g., excessive heat or water leakage).
- There was a danger the car would be stolen with equipment in it.

Dwelling Unit Visit and Inspection Protocol

- **A member of the inspection team, preferably the team member with responsibility for operating testing and sampling equipment, should have general training in engineering and architectural terms and project-specific training in terms used for the data collection.**

The initial pretest and the survey itself both encountered difficulties in uniformly and consistently categorizing the conditions of walls and paint, identifying substrate materials, and categorizing architectural components into study-specific categories. Specific architectural, construction, and engineering expertise should be brought together in advance with survey staff to work out the exact type of background needed for the inspector and the study-specific categories for recording architectural features and conditions. The effort should go as far as specifying training in pertinent architectural/engineering areas and in the use of the study-specific categories.

The underlying issue is the uniformity and accuracy with which architectural components were named (categorized) and evaluated.

- **Sampling of rooms and components should be conducted in advance of the inspection visit to the home, based on home inventories conducted over the telephone (or a previous visit) and data collection forms should be customized in advance for the sampled rooms and components.**

The final pre-field test of the original design found that in-field sampling of dwelling unit rooms and components was likely to create a source of error. Additionally, he pre-field test had the potential to waste the residents' time and convey a sense of disorganization on the part of the field team and the entire survey. In-field selection lengthened the inspection by 30 to 50 minutes, taken up with what appeared to the residents as confused paper shuffling. Frequently, just as the team was finally prepared to begin the actual inspection, the resident was ready for them to leave.

In an effort to minimize in-field selection, the final in-field sampling protocol divided the room and component sampling task into three parts: randomly numbering the four exterior walls of the house or apartment building; selecting a wet and a dry room; and selecting architectural components.

When the inspection team arrived at a dwelling unit, the interviewer noted which wall of the house faced the street (as named in the address of the house). That became wall 1. Going clockwise, the remaining three walls were numbered 2, 3, and 4. That scheme was applied to interior and exterior sampling. This was the only sampling that the inspection team needed to generate in the field.

With regard to wet or dry room selection, residents of units initially sampled from the field screening provided an inventory of the home on the phone. This allowed the wet or dry room sampling to be done at the field headquarters, prior to the field effort, and printed on the Interior Observation Form customized for each unit. Using the Interior Observation Form, the field team could easily designate which components should be selected and tested, based on a random sampling priority specifically generated for that unit's components.

The use of pre-printed, customized forms helped alleviate the problems of wasted time, errors, and omissions by the field teams while performing the inspection and testing in the unit.

- **As long as the inspection can be completed in one trip, dwelling unit sketches are not necessary.**

Because the original design called for two trips to the dwelling unit, one to gather sampling information and one to inspect, dwelling unit sketches were needed under this plan. Under the single visit design, there was no need for sketches.

- **Allow additional time when inspecting public housing and private multi-family units for common areas and coordination with escorts.**

Inspections that include the additional common areas cannot be accomplished in the same amount of time as single-family units. Longer time periods should be automatically blocked out on the field team's scheduling calendar whenever the inspection is to include common areas.

- **XRF scores (measurements) should not be repeated out loud in the dwelling unit.**

The original design called for the inspector to read the XRF reading out loud and for the interviewer to copy it down. This procedure raised the resident's curiosity and led to a barrage of questions. Although it was important to be open with the residents concerning all aspects of the study, reciting technical data in the midst of the inspection needlessly alarmed them and often had the opposite effect.

In-Field Environmental Sampling

- **The purposive XRF paint lead reading should be collected either from a pre-established location in each dwelling unit or only in dwelling unit areas previously entered as part of other testing.**

The study design called for the field technician to perform an XRF reading at a spot he thought had a high chance of having leaded paint. It was not reasonable, though, to expect the dwelling unit resident to allow the inspector to walk through the entire house looking around. Further, such a foray took a lot of time. One of the following alternatives seems preferable:

- Specify on a study-wide basis that a reading will be taken in a standard location that has a high probability of having lead-based paint.

- Employ the final protocol established for the National Survey, viz., limit the search to dwelling unit areas entered as part of the rest of the inspection, including the two rooms sampled and areas walked through getting to those rooms.

- **The analytic model concerning dust analysis and pathways needs to be fully specified before an adequate dust sampling protocol can be designed.**

Dust sampling and test result analysis are the subject of continuing investigations. Insufficient dust was a recurring problem in the field. Outside of insufficient dust, there were few field problems encountered with dust sampling.

A traveling team operates under two major constraints -- it must carry its own equipment and it must limit its time in the dwelling unit to a reasonable length. Getting more dust by using a bigger, more powerful vacuum or vacuuming a larger area were limited options. The vacuum carried by the survey team weighed ten pounds, plus tubing, nozzles, and extension cords. The National Survey found that it took four minutes to completely collect a dust sample (vacuuming and re-vacuuming a 4-square-foot area as specified in the dust sampling protocol). A bigger vacuum or more time spent vacuuming were not reasonable in the context of the National Survey.

At one point in the procedure design stage it was suggested that in-field technicians evaluate the quantity of dust collected in a cassette. They would collect more dust in the cassette if they determined there was insufficient dust. This suggestion proved to be impractical. First, inspectors could not reliably determine by visual inspection if enough dust had been collected. Second, movement of the cassette while attached to the vacuum could cause dust to fall out of the cassette. Third, allowing the technician individual discretion to collect samples would lead to inconsistencies in the procedure and findings. Last, the amount of time spent vacuuming had to be limited to keep the visit to a reasonable length. Therefore, this suggestion was not implemented in the National Survey.

A final suggested approach to increasing dust sample yield was "wet wipe" testing, used in conjunction with or following vacuuming, to pick up leftover dust. This technique would not have improved effective yield, because wet wipe test results could not be compared or added in any meaningful way to vacuumed dust sampling results.

Static electricity posed a problem to the inspection team. Dust would cling to the vacuum nozzle, the edges of the template, etc. Efforts were made to build the template out of a material that did not accumulate static electricity but the phenomenon still occurred.

- **XRF's can be transported safely by air (as luggage) or by Federal Express or other carrier in plain wrapping. Interior wrapping must warn of the radioactive source. Authorizations for transport should be packed inside of the exterior wrapping but still accessible without opening the interior XRF case.**

There was much discussion about transporting the XRF, given its radioactive source. Because the National Survey equipment was never detained in transit, the above packaging approach appears to be an effective one. One member of the team must be authorized and licensed to transport and use the XRF.

Use of the Spectrum Analyzer MAP/XRF

- **Make sure licenses for MAP/XRF use are amended for the intensity of source used.**

A member of the inspection team must have a state-issued license that allows him to transport and use the MAP/XRF. Licenses are, typically, tied to the intensity of the radiation source and must be valid for the intensity being used. In most cases, the authorization for the higher millicurie level cost more, sometimes by several hundred dollars. The licenses should be applied for at least two months before the scheduled beginning of the field period.

All but one of the states in the National Survey required detailed information about where the MAP/XRF would be, when and how long it would be in the state, and who would be responsible for it. Several states required the addresses of the sampled homes and dates of inspections. The study staff needs to be prepared to supply this information to those responsible for the MAP/XRFs and their licensing.

- **Use a "full intensity" radiation source in the XRF.**

The XRF should have a full intensity radiation source to get the best readings in the minimum amount of time. Full intensity radiation presented no additional risk to residents, technicians, or the general population. The National Survey used a 40 millicurie source and 1-minute-long readings.

- **The XRF reading can be taken at any convenient place on a sampled architectural component; valid sampling does not require readings of a randomly selected spot on the component. This eliminates readings being taken in hard-to-reach locations, such as corners.**

The original design called for sampling a random location on the surface of a component for testing. Though this is subject to further verification, it appears that components with common paint history produced similar XRF readings. Another stage of sampling would significantly burden the inspection process. Correct assessment of "common paint history" is more pivotal here.

- **The MAP/XRF should not be used to scan or take an "average" reading on a component.**

The original MAP/XRF design called for "scanning" components with the MAP/XRF by running the "eye" across all parts of a component. The initial pretest results did not support this practice. The final survey protocol eliminated scanning. Experience with the MAP/XRF leads to the belief that the readings produced this way could be subject to unpredictable error because of the inability to evenly and smoothly scan a component's surface.

- **All MAP/XRF readings should be recorded on paper by the inspector.**

The MAP/XRF used in the National Survey was programmed to store the spectrum results in memory. Serious problems were encountered, though, when the memory was in use. The memory would fill up rapidly and the MAP/XRF would stop operating. The equipment had to be turned off, and then back on, in order to restore operation. All memory was lost in the process. The need to link readings to a location necessitated recording of a certain amount of data on paper in any event, so that recording MAP/XRF readings on paper involved little extra effort.

The equipment problem aside, using the MAP/XRF's on-board electronic storage necessitates some procedures for downloading memory contents during the field period. In practice it did not prove reasonable or practical to return the MAP/XRF to the survey operations office periodically for downloading. Teaching field staff how to download the memory and transmit it to the field operations office would have meant providing them with a properly-configured PC and modem in the field and providing PC training. Adding responsibilities and equipment to the technicians' load was deemed inadvisable.

- **MAP/XRF should be equipped with new batteries frequently so they are always running at full power. Batteries can freeze in transit (e.g., in the cargo hold of a plane) and lose their charge.**

Uneven power supply from batteries seemed to be associated with quirkiness in the operating of the MAP/XRF. The reliable response seemed to be simply installing new batteries (purchased locally) upon arrival in each county and again after several dwelling unit inspections. As noted above, if new batteries do not clear up the quirkiness in MAP/XRF performance, the console needs to be returned to the manufacturer for adjustment. There are no in-field adjustments possible.

APPENDIX A

Additional Data Tables for

Private Housing

TABLE A-1

PREVALENCE OF LEAD-BASED PAINT (LBP) BY LOCATION IN THE BUILDING - PRIVATELY OWNED OCCUPIED HOUSING UNITS

	Occupied Housing Units With Lead-Based Paint	
Location of LBP	**Number (000)**	**Percent (1)**
Unit Interior	48,986	63%
Interior Common Area	3,596	5%
Building Exterior	56,495	73%
Playground	525	1%
Somewhere in Building	64,443	83%

(1) Base equals all 77,177,000 housing units built before 1980.
(2) Numbers based on small sample sizes should be interpreted with caution.

TABLE A-2

ESTIMATED NUMBER OF PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT BY SELECTED CHARACTERISTICS (Paint Lead Concentration >= 0.7 mg/sq cm)

Characteristic	Total Housing Units (000) (1)	Housing Units With Lead-Based Paint Somewhere in Building		Number of Housing Units in Sample
		Percent	Number (000)	
Total Housing Units Built Before 1980	77,177	87% (5%)	66,831 (3,670)	284
One or More Children Under Age 7	13,912	92% (6%)	12,783 (863)	90
Construction Year				
1960-1979	35,681	82% (8%)	29,195 (2,708)	120
1940-1959	20,476	94% (5%)	19,210 (1,099)	87
Before 1940	21,018	88% (11%)	18,426 (2,254)	77
Housing Type				
Single Family	66,418	87% (5%)	57,926 (3,462)	227
Multifamily	10,759	83% (11%)	8,905 (1,160)	57

(1) Total units data are from the 1987 American Housing Survey.

Note: Numbers in parentheses are approximate half-widths of 95% confidence intervals for the estimated percents and numbers. For example, the approximate 95% confidence interval for the percent of housing units with some lead-based paint is 87% +/- 5% or 82% to 92%.

TABLE A-3

ESTIMATED NUMBER OF PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT BY SELECTED CHARACTERISTICS (Paint Lead Concentration >= 1.2 mg/sq cm)

Characteristic	Total Housing Units (000) (1)	Housing Units With Lead-Based Paint Somewhere in Building		Number of Housing Units in Sample
		Percent	Number (000)	
Total Housing Units Built Before 1980	77,177	80% (6%)	61,475 (4,336)	284
One or More Children Under Age 7	13,912	85% (8%)	11,873 (1,118)	90
Construction Year				
1960-1979	35,681	69% (9%)	24,769 (3,236)	120
1940-1959	20,476	89% (7%)	18,281 (1,411)	87
Before 1940	21,018	88% (11%)	18,426 (2,254)	77
Housing Type				
Single Family	66,418	80% (6%)	53,423 (4,113)	227
Multifamily	10,759	75% (12%)	8,052 (1,333)	57

(1) Total units data are from the 1987 American Housing Survey.

Note: Numbers in parentheses are approximate half-widths of 95% confidence intervals for the estimated percents and numbers. For example, the approximate 95% confidence interval for the percent of housing units with some lead-based paint is 80% +/- 6% or 74% to 86%.

TABLE A-4

ESTIMATED NUMBER OF PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT BY SELECTED CHARACTERISTICS (Paint Lead Concentration >= 2.0 mg/sq cm)

Characteristic	Total Housing Units (000) (1)	Housing Units With Lead-Based Paint Somewhere in Building		Number of Housing Units in Sample
		Percent	Number (000)	
Total Housing Units Built Before 1980	77,177	68% (6%)	52,690 (5,013)	284
One or More Children Under Age 7	13,912	73% (10%)	10,128 (1,407)	90
Construction Year				
1960-1979	35,681	48% (10%)	17,219 (3,509)	120
1940-1959	20,476	83% (8%)	17,045 (1,703)	87
Before 1940	21,018	88% (11%)	18,426 (2,254)	77
Housing Type				
Single Family	66,418	69% (7%)	45,602 (4,810)	227
Multifamily	10,759	66% (14%)	7,088 (1,457)	57

(1) Total units data are from the 1987 American Housing Survey.

Note: Numbers in parentheses are approximate half-widths of 95% confidence intervals for the estimated percents and numbers. For example, the approximate 95% confidence interval for the percent of housing units with some lead-based paint is 68% +/- 6% or 62% to 74%.

TABLE A-5

ASSOCIATION BETWEEN LEAD IN SOIL AND EXTERIOR LEAD-BASED PAINT CONDITION FOR PRIVATELY OWNED HOUSING UNITS (Numbers Represent Thousands of Housing Units)

Presence and Condition of Exterior Lead-Based Paint	Lead in Soil															
	Entrance				Drip line				Remote				Any Location			
	Within Guideline (1)		Exceeding Guideline (1)		Within Guideline (1)		Exceeding Guideline (1)		Within Guideline (1)		Exceeding Guideline (1)		Within Guideline (1)		Exceeding Guideline (1)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Exterior Walls																
No LBP	18,266	98%	441	2%	17,506	92%	1,447	8%	18,725	100%	- -	- -	17,719	91%	1,660	9%
LBP Present, Intact	38,782	88%	5,448	12%	37,066	86%	5,989	14%	40,748	92%	3,387	8%	35,914	80%	9,215	20%
LBP Present, Non-Intact	5,099	59%	3,537	41%	3,976	48%	4,310	52%	6,741	82%	1,450	18%	3,815	44%	4,821	56%
Playgrounds																
No LBP	61,625	87%	9,426	13%	58,026	83%	11,746	17%	65,692	93%	4,837	7%	56,926	78%	15,695	22%
LBP Present, Intact	522	100%	- - (2)	- -	522	100%	- -	- -	522	100%	- -	- -	522	100%	- -	- -
Total	62,147	87%	9,426	13%	58,547	83%	11,746	17%	66,214	93%	4,837	7%	57,448	79%	15,695	21%

(1) Although there is no federal standard for residential soil lead contamination, many experts agree that 500 ppm is a feasible threshold to designate "high" soil lead contamination in residential environments. EPA's interim guidance on soil lead cleanup levels at Superfund sites sets the cleanup levels at 500 to 1000 ppm [US EPA(Sept. 7, 1989), Interim Guidance on Establishing Soil Lead Cleanup Levels at Superfund Sites (OSWER Directive #9355.4-02)].

(2) All the paint was within the Guidelines.

(3) There were no playgrounds in the sample with non-intact lead-based paint present.

TABLE A-6

AMOUNTS OF LEAD-BASED PAINT (LBP) ON INTERIOR SURFACES BY PAINTED COMPONENTS AND YEAR CONSTRUCTED FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS (LBP Concentration >= 1.0 mg/sq cm)

	National Total Amount of LBP		Amount LBP Per Housing Unit With LBP
	(millions of sq ft)	(percent of all paint)	(square feet)
Components:			
Walls/ceiling/floor			
1960-1979	4,920	5%	281
1940-1959	7,121	15%	505
Built before 1940	6,106	11%	351
Metal component (1)			
1960-1979	24	2%	1
1940-1959	45	6%	3
Built before 1940	37	3%	2
Non-metal component (2)			
1960-1979	328	4%	19
1940-1959	873	9%	62
Built before 1940	5,971	47%	343
Shelves/other (3)			
1960-1979	6	0%	0
1940-1959	208	7%	15
Built before 1940	3,798	68%	218

(1) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(2) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(3) Includes shelves, cabinets, fireplace, and closets, on any substrate.

Note: Because of rounding, totals may not be exactly the same as the sum of the numbers.

TABLE A-7

AMOUNTS OF LEAD-BASED PAINT (LBP) ON EXTERIOR SURFACES BY PAINTED COMPONENT AND YEAR CONSTRUCTED FOR PRIVATELY OWNED OCCUPIED HOUSING UNITS
(LBP Concentration >= 1.0 mg/sq cm)

	National Total Amount of LBP		**Amount LBP Per Housing Unit With LBP (square feet)**
	(millions of sq ft)	**(percent of all paint)**	
Components:			
Walls/ceiling/floor			
1960-1979	8,825	28%	405
1940-1959	10,423	45%	625
Built before 1940	19,199	80%	1,066
Metal component (1)			
1960-1979	76	4%	3
1940-1959	146	8%	9
Built before 1940	180	13%	10
Non-metal component (2)			
1960-1979	1,575	15%	72
1940-1959	1,857	39%	111
Built before 1940	6,098	78%	338
Porches/other (3)			
1960-1979	26	2%	1
1940-1959	208	19%	12
Built before 1940	492	13%	27

(1) Includes only metal windows, doors, soffit and facia, columns, and railings.
(2) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(3) Includes porches, balconies, stairs, etc., on any substrate.

Note: Because of rounding, totals may not be exactly the same as the sum of the numbers.

TABLE A-8

GEOMETRIC MEAN PAINT LEAD LOADINGS IN PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980, BY SELECTED CHARACTERISTICS (Paint Lead Concentration >= 1.0 mg/sq cm)

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Total Occupied Housing Units Built Before 1980	0.1 (0.1 , 0.2)	0.3 (0.2 , 0.5)
Construction Year:		
1960-1979	0.1 (0.0 , 0.1)	0.1 (0.1 , 0.2)
1940-1959	0.2 (0.1 , 0.3)	0.4 (0.2 , 0.8)
Before 1940	0.3 (0.2 , 0.6)	1.6 (0.4 , 1.1)
Housing Type		
Single Family	0.2 (0.1 , 0.2)	0.4 (0.2 , 0.6)
Multifamily	0.1 (0.1 , 0.2)	0.2 (0.1 , 0.6)
One or More Children Under Age 7	0.1 (0.1 , 0.2)	0.3 (0.2 , 0.5)
Census Region		
Northeast	0.3 (0.2 , 0.6)	0.8 (0.3 , 2.1)
Midwest	0.1 (0.0 , 0.2)	0.6 (0.3 , 1.1)
South	0.2 (0.1 , 0.2)	0.2 (0.1 , 0.3)
West	0.1 (0.0 , 0.2)	0.2 (0.0 , 0.6)

(1) Numbers in parentheses are 95% confidence intervals for the respective geometric means.

TABLE A-9

GEOMETRIC MEAN LEAD LOADINGS IN PRIVATELY OWNED OCCUPIED HOUSING UNITS BUILT BEFORE 1980, BY ARCHITECTURAL COMPONENT AND CONSTRUCTION YEAR (Paint Lead Concentration >= 1.0 mg/sq cm)

Component/Construction year	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Walls/ceilings/floor		
1960-1979	0.05 (0.0 , 0.1)	0.1 (0.0 , 0.2)
1940-1959	0.1 (0.1 , 0.2)	0.3 (0.1 , 0.7)
Before 1940	0.2 (0.1 , 0.3)	2.5 (1.3 , 4.9)
Metal (1)		
1960-1979	0.05 (0.0 , 0.2)	0.003 (0.0 , 0.1)
1940-1959	0.05 (0.0 , 0.3)	0.1 (0.0 , 1.1)
Before 1940	0.05 (0.0 , 0.5)	0.1 (0.0 , 1.8)
Non-metal (2)		
1960-1979	0.1 (0.1 , 0.2)	0.2 (0.1 , 0.4)
1940-1959	0.2 (0.1 , 0.4)	0.6 (0.3 , 1.1)
Before 1940	0.9 (0.5 , 1.6)	1.9 (1.1 , 3.4)
Other (3)		
1960-1979	0.01 (0.0 , 0.0)	0.01 (0.0 , 0.0)
1940-1959	0.03 (0.0 , 0.1)	0.06 (0.0 , 0.3)
Before 1940	0.20 (0.1 , 0.4)	0.3 (0.1 , 1.0)

Note: Numbers in parentheses are 95% confidence intervals for the respective arithmetic means.

Interior:
(1) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(2) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(3) Includes shelves, cabinets, fireplace, and closets, on any substrate.

Exterior:
(1) Includes only metal windows, doors, soffit and facia, columns, and railings.
(2) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(3) Includes porches, balconies, stairs, etc., on any substrate.

TABLE A-10

DESCRIPTIVE STATISTICS FOR THE DUST LEAD CONCENTRATION MEASUREMENTS (WEIGHTED)

Set of Data	Dry room floor	Entry way floor	Wet room floor	Dry room window sill	Dry room window well	Wet room window sill	Wet room window well
Number of measurements summarized	270	269	265	207	78	131	71
Arithmetic mean (ppm)	631	813	440	5,264	10,186	5,083	13,132
Percentiles (ppm)							
maximum	11,287	18,563	8,376	96,492	457,178	104,368	83,633
upper quartile	378	922	483	2,466	6,326	2,583	7,450
median	188	380	198	735	1,962	826	2,432
lower quartile	102	201	83	259	536	289	575
minimum	3	21	6	1	5	1	22
Geometric mean (ppm)	224	423	204	925	1,792	1,011	3,236
Mean of the log transformed measurements	5.41	6.05	5.32	6.83	7.49	6.92	8.08

TABLE A-11

DESCRIPTIVE STATISTICS FOR THE DUST LEAD LOADING MEASUREMENTS (WEIGHTED)

Set of Data	Dry room floor	Entry way floor	Wet room floor	Dry room window sill	Dry room window well	Wet room window sill	Wet room window well
Number of measurements summarized	273	274	275	233	84	158	74
Arithmetic mean (ug\sq. ft.)	6.92	12.68	4.14	91.20	841.40	96.80	790.62
Percentiles (ug\sq. ft.)							
maximum	205	380	233	2,638	40,455	11,899	7,139
upper quartile	3.43	8.05	2.51	24.70	475.40	11.98	528.10
median	0.96	2.59	0.68	5.04	85.90	2.15	90.26
lower quartile	0.31	0.71	0.21	0.95	15.40	0.44	18.83
minimum	0.00	0.03	0.00	0.00	0.04	0.00	0.19
Geometric mean (ug\sq. ft.)	1.12	2.44	0.74	5.17	95.10	2.50	121.98
Mean of the log transformed measurements	0.11	0.89	-0.30	1.64	4.55	0.93	4.80

TABLE A-12

DESCRIPTIVE STATISTICS FOR THE DUST LOADING MEASUREMENTS(WEIGHTED)

Set of Data	Dry room floor	Entry way floor	Wet room floor	Dry room window sill	Dry room window well	Wet room window sill	Wet room window well
Number of measurements summarized	269	74	203	222	48	150	71
Arithmetic mean (ug\sq. ft.)	912,523	2,190,863	18,099,923	2,374,765	316,816	1,388,937	558,829
Percentiles (ug\sq. ft.)							
maximum	40,000,000	23,143,421	1,658,775,736	153,676,471	8,018,868	16,239,316	16,666,667
upper quartile	579,832	1,734,428	5,716,381	427,842	37,445	466,473	64,103
median	180,995	387,136	1,269,646	82,701	11,870	91,792	16,906
lower quartile	61,876	90,841	298,705	12,133	3,478	18,182	4,975
minimum	1,895	4,473	45	59	48	148	423
Geometric mean (ug\sq. ft.)	2,246	22	100	49	4	26	10
Mean of the log transformed measurements	7.71	3.08	4.60	3.88	1.49	3.28	2.32

APPENDIX B

ADDITIONAL DATA TABLES

FOR PUBLIC HOUSING

TABLE B-1

ESTIMATED NUMBER AND PERCENT OF PUBLIC HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT, BY SELECTED CHARACTERISTICS

(Paint Lead Concentration >= 0.7 mg/sq cm)

Characteristic		Total Public Housing Units (000)	Housing Units With Lead-Based Paint Somewhere in Building		Number of Housing Units in Sample
			Percent	Number (000)	
Total Public Housing Units Built Before 1980		910	90% (82% - 98%)	821 (748 - 895)	97
Construction Year:	1960-1979	455	88% (76% - 99%)	399 (347 - 450)	43
	1950-1959	273	90% (77% - 100%)	246 (209 - 273)	24
	Before 1950	182	97% (88% - 100%)	177 (160 - 182)	30

Note: Numbers in parentheses are 95% confidence intervals for the estimated percents and numbers.

TABLE B-2

ESTIMATED NUMBER AND PERCENT OF PUBLIC HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT, BY SELECTED CHARACTERISTICS

(Paint Lead Concentration >= 1.2 mg/sq cm)

Characteristic		Total Public Housing Units (000)	Housing Units With Lead-Based Paint Somewhere in Building: Percent	Housing Units With Lead-Based Paint Somewhere in Building: Number (000)	Number of Housing Units in Sample
Total Public Housing Units Built Before 1980		910	85% (77% - 93%)	774 (697 - 850)	97
Construction Year:	1960-1979	455	77% (64% - 91%)	351 (290 - 412)	43
	1950-1959	273	90% (77% - 100%)	246 (209 - 273)	24
	Before 1950	182	97% (88% - 100%)	177 (160 - 182)	30

Note: Numbers in parentheses are 95% confidence intervals for the estimated percents and numbers.

TABLE B-3

ESTIMATED NUMBER AND PERCENT OF PUBLIC HOUSING UNITS BUILT BEFORE 1980 WITH LEAD-BASED PAINT, BY SELECTED CHARACTERISTICS

(Paint Lead Concentration >= 2.0 mg/sq cm)

Characteristic		Total Public Housing Units (000)	Housing Units With Lead-Based Paint Somewhere in Building: Percent	Housing Units With Lead-Based Paint Somewhere in Building: Number (000)	Number of Housing Units in Sample
Total Public Housing Units Built Before 1980		910	77% (67% - 86%)	697 (614 - 780)	97
Construction Year:	1960-1979	455	65% (50% - 81%)	297 (228 - 367)	43
	1950-1959	273	82% (65% - 98%)	223 (177 - 268)	24
	Before 1950	182	97% (88% - 100%)	177 (160 - 182)	30

Note: Numbers in parentheses are 95% confidence intervals for the estimated percents and numbers.

TABLE B-4

GEOMETRIC MEAN PAINT LEAD LOADINGS IN PUBLIC HOUSING UNITS BUILT BEFORE 1980, BY SELECTED CHARACTERISTICS

Characteristic	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Total Occupied Housing Units Built Before 1980	0.2 (0.2 , 0.3)	0.2 (0.1 , 0.4)
Construction Year:		
1960-1979	0.2 (0.1 , 0.3)	0.1 (0.0 , 0.2)
1950-1959	0.2 (0.1 , 0.4)	0.3 (0.0 , 2.3)
Before 1950	0.3 (0.2 , 0.5)	1.2 (0.6 , 2.4)
Census Region		
Northeast	0.2 (0.1 , 0.4)	0.2 (0.0 , 1.3)
Midwest	0.2 (0.1 , 0.8)	0.0 (0.0 , 0.8)
South	0.2 (0.1 , 0.3)	0.2 (0.1 , 0.4)
West	0.2 (0.0 , 0.8)	0.3 (0.1 , 0.8)

Note: Numbers in parentheses are 95% confidence intervals for the respective geometric means.

TABLE B-5

GEOMETRIC MEAN PAINT LEAD LOADINGS IN PUBLIC HOUSING UNITS BUILT BEFORE 1980, BY ARCHITECTURAL COMPONENT AND CONSTRUCTION YEAR

Component/Construction year	Interior Surfaces (mg/sq. cm.)	Exterior Surfaces (mg/sq. cm.)
Walls/ceilings/floor		
1960-1979	0.1 (0.1 , 0.3)	0.1 (0.1 , 0.5)
1950-1959	0.1 (0.1 , 0.3)	0.9 (0.9 , 0.9)
Before 1950	0.2 (0.1 , 0.5)	0.6 (0.2 , 1.4)
Metal (1)		
1960-1979	0.03 (0.0 , 0.3)	0.001 (0.0 , 0.0)
1950-1959	0.8 (0.4 , 1.8)	0.2 (0.0 , 6.3)
Before 1950	0.4 (0.1 , 1.8)	0.9 (0.4 , 2.3)
Non-metal (2)		
1960-1979	0.2 (0.1 , 0.4)	0.1 (0.1 , 0.2)
1950-1959	0.2 (0.1 , 0.5)	0.9 (0.2 , 5.7)
Before 1950	0.3 (0.2 , 0.6)	3.3 (1.5 , 7.2)
Other (3)		
1960-1979	0.04 (0.0 , 0.1)	0.1 (0.0 , 0.8)
1950-1959	0.04 (0.0 , 0.2)	0.3 (0.1 , 1.0)
Before 1950	0.1 (0.0 , 0.2)	0.6 (0.1 , 3.2)

Note: Numbers in parentheses are 95% confidence intervals for the respective arithmetic means.

Interior:

(1) Includes metal trim, window sills, molding, doors, air/heat vents, and radiators.
(2) Includes non-metal trim, window sills, molding, doors, and air/heat vents.
(3) Includes shelves, cabinets, fireplace, and closets, on any substrate.

Exterior:

(1) Includes only metal windows, doors, soffit and facia, columns, and railings.
(2) Includes non-metal windows, doors, soffit and facia, columns, and railings.
(3) Includes porches, balconies, stairs, etc., on any substrate.

APPENDIX C

Dust and Soil Samples Excluded
From Data Quality Analysis

TABLE C-1

	LBP_ID	SD_ID	PB_ADJ	SAMP_WGT	SD_VACU	PPM	SD_LOAD	WET_WIPE flg1	MISNG_DATA_FLG flg2	PPM_FLG flg3	LOAD_FLG flg4	WGT_FLG flg5
1	131102	63	366.36	0.0012	4.00	384306.17	91.3180	0	0	1	0	0
2	211102	66	0.10	0.0000	1.33	0.00	0.0728	0	1	0	0	0
3	221101	66	0.46	0.0000	1.00	0.00	0.4600	0	1	0	0	0
4	310102	71	0.21	0.0000	1.33	0.00	0.1066	0	1	0	0	0
5	320101	64	0.06	0.0000	1.00	0.00	0.0630	0	1	0	0	0
6	320408	64	0.06	0.0000	0.67	0.00	0.0727	0	1	0	0	0
7	320408	66	0.06	0.0000	0.67	0.00	0.0934	0	1	0	0	0
8	320507	64	0.44	0.0000	0.67	0.00	0.0645	0	1	0	0	0
9	321307	64	0.01	0.0000	1.33	0.00	0.0060	0	1	0	0	0
10	321307	65	0.25	0.0000	1.00	0.00	0.2500	0	1	0	0	0
11	321307	66	0.01	0.0000	1.33	0.00	0.0080	0	1	0	0	0
12	[illegible]	64	0.06	0.0000	1.33	0.00	0.0094	0	1	0	0	0
13	350001	63	0.45	0.0000	4.00	0.00	0.1121	0	1	0	0	0
14	410100	66	2560.00	0.0570	1.00	44117.06	2560.0000	0	0	0	1	0
15	[illegible]	66	4.00	3.2700	2.00	0.76	2.0000	0	0	0	0	1
16	440002	66	16.00	5.3000	1.33	3.02	12.0000	0	0	0	0	1
17	520700	65	2020.00	0.0700	1.00	25472.00	2020.0000	0	0	0	1	0
18	520700	66	1120.00	0.0002	1.00	5660000.00	1120.0000	0	0	1	0	0
19	[illegible]	70	0.48	0.0000	4.00	0.00	0.1200	0	1	0	0	0
20	[illegible]	71	0.23	0.0000	2.00	0.00	0.1106	0	1	0	0	0
21	[illegible]	62	2.10	0.0000	4.00	0.00	0.5475	0	1	0	0	0
22	[illegible]	64	0.31	0.0000	0.67	0.00	0.4000	0	1	0	0	0
23	720300	68	10.50	0.0000	4.00	3.20	4.6746	1	0	0	0	0
24	720300	69	303.02	7.0010	4.00	42.70	75.7546	1	0	0	0	0
25	720300	70	13610.02	0.0000	4.00	19653.44	3404.0000	0	0	0	1	0
26	720300	71	8502.96	0.2300	2.00	36402.36	4251.4810	0	0	0	1	0
27	730006	64	11500.00	0.1167	1.00	98543.27	11500.0000	0	0	0	1	0
28	730006	65	6000.00	0.2573	1.00	20610.06	6000.0000	0	0	0	1	0
29	730006	67	30100.00	2.0019	1.00	15063.06	30100.0000	0	0	0	1	0
30	750405	66	1.54	0.0000	1.33	0.00	1.1532	0	1	0	0	0
31	760100	61	107.00	0.0005	4.00	214000.00	26.7500	0	0	1	0	0
32	760106	64	0.47	0.0000	1.33	0.00	0.3406	0	1	0	0	0
33	[illegible]	61	0.41	0.0000	4.00	0.00	0.1026	0	1	0	0	0
34	911404	66	0.83	0.0000	1.00	0.00	0.9262	0	1	0	0	0
35	911503	64	0.26	0.0000	0.67	0.00	0.3066	0	1	0	0	0
36	920001	66	0.11	0.0000	2.67	0.00	0.0400	0	1	0	0	0
37	[illegible]	66	0.42	0.0000	1.00	0.00	0.4241	0	1	0	0	0
38	930701	66	0.37	0.0000	1.00	0.00	0.3670	0	1	0	0	0
39	940700	61	0.32	0.0000	4.00	0.00	0.0811	0	1	0	0	0
40	941005	61	1.40	0.0000	4.00	0.00	0.3734	0	1	0	0	0
41	950402	67	0.00	0.3015	1.00	0.00	0.0000	0	1	0	0	0
42	1011501	66	1.25	0.0000	1.00	0.00	1.2510	0	1	0	0	0
43	1011600	61	2.00	0.0004	0.80	5000.00	0.0000	0	1	0	0	0
44	1011600	62	5.00	0.0000	0.80	1000.00	0.0000	0	1	0	0	0
45	1011600	63	3.00	0.0008	0.80	3750.00	0.0000	0	1	0	0	0
46	1011600	64	1.90	0.0000	0.80	0.00	0.0000	0	1	0	0	0
47	1011600	66	1.00	0.0000	0.80	0.00	0.0000	0	1	0	0	0
48	1011700	66	0.20	0.0000	1.00	0.00	0.1000	0	1	0	0	0
49	1020502	64	0.16	0.0000	1.00	0.00	0.1634	0	1	0	0	0
50	1020700	66	0.06	0.0000	1.00	0.00	0.0044	0	1	0	0	0
51	[illegible]	62	1.25	0.0000	4.00	0.00	0.3128	0	1	0	0	0
52	1041007	61	2.00	0.0000	4.00	0.00	0.0000	0	1	0	0	0
53	1041007	64	0.37	0.0000	1.00	0.00	0.3740	0	1	0	0	0
54	1041007	66	0.90	0.0000	2.00	0.00	0.4500	0	1	0	0	0
55	[illegible]	66	0.13	0.0000	2.00	0.00	0.0630	0	1	0	0	0
56	1051200	66	0.04	0.0000	1.00	0.00	0.0430	0	1	0	0	0
57	1051406	63	0.16	0.0000	4.00	0.00	0.0306	0	1	0	0	0
58	1100200	62	0.31	0.0000	4.00	0.00	0.0785	0	1	0	0	0
59	1100200	64	0.25	0.0000	0.80	0.00	0.5670	0	1	0	0	0
60	[illegible]	61	31.31	23.0000	4.00	1.31	7.8273	1	0	0	0	1
61	[illegible]	62	43.00	22.0000	4.00	1.04	10.8721	1	0	0	0	1
62	[illegible]	63	7.33	23.0022	4.00	0.32	1.8318	1	0	0	0	1
63	[illegible]	64	60.77	5.2001	0.67	9.06	70.1617	1	0	0	0	1
64	[illegible]	66	30.02	5.0014	1.00	7.02	30.0231	1	0	0	0	1
65	[illegible]	70	16.04	33.2721	4.00	0.67	4.2102	1	0	0	0	1
66	[illegible]	71	10.32	5.5140	1.33	1.87	7.7433	1	0	0	0	1
67	[illegible]	61	80.80	30.0000	4.00	1.88	15.2224	1	0	0	0	1
68	[illegible]	62	17.86	31.0004	4.00	0.57	4.3062	1	0	0	0	1
69	[illegible]	63	58.30	30.0043	4.00	1.97	14.6652	1	0	0	0	1
70	[illegible]	64	33.34	7.7165	1.00	4.32	23.3406	1	0	0	0	1
71	[illegible]	65	204.04	8.0041	1.00	32.04	204.6446	1	0	0	0	1
72	[illegible]	66	186.31	7.2072	2.00	22.02	82.6365	1	0	0	0	1
73	[illegible]	67	404.42	8.0727	1.00	46.07	404.4240	1	0	0	0	1
74	[illegible]	70	25.72	33.1063	4.00	0.77	6.4290	1	0	0	0	1
75	[illegible]	71	25.01	0.4472	1.00	3.86	25.6142	1	0	0	0	1
76	[illegible]	72	346.30	0.0011	1.00	30.83	345.1963	1	0	0	0	1
77	1100704	70	80.88	32.4001	4.00	2.15	17.4282	1	0	0	0	1
78	1101106	61	224.06	30.0000	4.00	8.36	56.1816	1	0	0	0	1
79	1101106	62	166.90	27.8710	4.00	3.83	26.4784	1	0	0	0	1
80	1101106	63	101.87	30.5867	4.00	3.46	25.4673	1	0	0	0	1
81	1101106	64	4.47	7.2202	1.00	0.82	4.4000	1	0	0	0	1
82	1101106	65	53.30	8.1120	1.00	6.57	53.3044	1	0	0	0	1

TABLE C-1

DUST SAMPLES EXCLUDED FROM ANALYSIS (Continued)

	LBP_ID	SD_ID	PB_ADJ	SAMP_WGT	SD_VACU	FRM	SD_LOAD	WET_WIPE flg1	MISNG_DATA_FLG flg2	FRM_FLG flg3	LOAD_FLG flg4	WGT_FLG flg5
83	1161106	66	10.82	0.9082	1.80	1.46	10.8182	1	0	0	0	1
84	1161106	67	135.12	7.7380	1.80	17.47	135.1106	1	0	0	0	0
85	1161106	70	11.51	32.8574	4.80	0.36	2.8771	1	0	0	0	1
86	1161207	61	43.47	29.8863	4.80	1.63	10.8175	1	0	0	0	1
87	1161207	62	4.41	28.8888	4.80	0.17	1.1017	1	0	0	0	1
88	1161207	63	4.19	27.2735	4.80	0.15	1.0486	1	0	0	0	1
89	1161207	64	154.24	8.4380	1.80	18.29	154.2437	1	0	0	0	1
90	1161207	65	71.18	0.0102	1.80	10.78	71.8861	1	0	0	0	0
91	1161207	66	148.21	0.0042	1.80	22.38	148.2112	1	0	0	0	1
92	1161207	67	172.36	0.2916	1.80	27.88	172.3680	1	0	0	0	0
93	1161207	70	9.72	28.6871	4.80	0.33	2.4290	1	0	0	0	1
94	1221602	63	0.90	0.0000	4.80	0.00	1.7280	0	1	0	0	0
95	1221602	66	5000.00	0.0004	2.33	64876.06	2486.7140	0	0	0	1	0
96	1221602	67	6000.00	0.0040	2.33	12806.42	2828.5710	0	0	0	1	0
97	1311606	63	0.86	0.0000	4.80	0.00	0.0121	0	1	0	0	0
98	1311606	64	0.86	0.0000	1.80	0.00	0.0486	0	1	0	0	0
99	1322001	64	0.30	0.0000	0.67	0.00	0.4483	0	1	0	0	0
100	1322000	63	0.86	0.0000	4.80	0.00	0.0121	0	1	0	0	0
101	1362000	67	0.86	0.0000	1.33	0.00	0.0004	0	1	0	0	0
102	1363706	64	2.78	0.0000	1.80	0.00	2.7783	0	1	0	0	0
103	1363706	66	0.94	0.0000	1.80	0.00	0.9375	0	1	0	0	0
104	1380106	67	0.28	0.0000	2.67	0.00	0.0004	0	1	0	0	0
105	1410406	66	0.19	0.0000	2.67	0.00	0.0720	0	1	0	0	0
106	1441302	64	0.01	0.0000	2.67	0.00	0.0046	0	1	0	0	0
107	1510806	62	4.14	0.0000	4.80	0.00	1.0344	0	1	0	0	0
108	1510806	65	0.28	0.0000	1.80	0.00	0.2842	0	1	0	0	0
109	1520204	66	1.30	0.0000	0.86	0.00	2.3676	0	1	0	0	0
110	1521500	66	3.41	0.0000	0.67	0.00	5.1078	0	1	0	0	0
111	1530500	64	0.46	0.0000	1.80	0.00	0.4846	0	1	0	0	0
112	1531201	67	1546.18	0.0000	0.75	1719.29	2061.5440	0	0	0	1	0
113	1531200	66	3.52	0.0000	0.80	0.00	7.0486	0	1	0	0	0
114	1540400	62	0.80	0.1216	4.80	0.00	0.0000	0	1	0	0	0
115	1561704	64	1.25	0.0000	1.80	0.00	1.2510	0	1	0	0	0
116	1561704	67	39.23	0.0001	0.67	392290.10	58.8300	0	0	1	0	0
117	1600406	62	95.01	11.1464	4.80	8.52	23.7532	1	0	0	0	0
118	1600406	64	191.47	1.0020	1.80	101.74	100.7722	1	0	0	0	0
119	1600406	66	304.38	2.4216	3.30	162.06	119.5113	1	0	0	0	0
120	1600406	69	440.50	10.0000	4.80	44.83	112.3000	1	0	0	0	0
121	1600406	70	282.87	11.7826	4.80	24.86	73.2434	1	0	0	0	0
122	1600506	69	1145.40	10.0020	4.80	106.26	286.3503	1	0	0	0	0
123	1600506	70	5921.07	10.2004	4.80	576.83	1480.2670	1	0	0	0	0
124	1600804	69	28.16	11.8834	4.80	2.41	7.0366	1	0	0	0	0
125	1600703	69	236.72	10.1461	4.80	23.53	59.6796	1	0	0	0	0
126	1600703	70	20.62	10.0000	4.80	1.93	5.1561	1	0	0	0	0
127	1731603	62	0.00	0.0000	4.80	0.00	0.0215	0	1	0	0	0
128	1730106	62	0.00	0.0000	4.80	0.00	0.0000	0	1	0	0	0
129	1750106	64	0.26	0.0000	1.30	0.00	0.2663	0	1	0	0	0
130	1760306	69	152.23	10.1806	4.80	15.00	38.0577	1	0	0	0	0
131	1760306	69	42.06	11.1380	4.80	3.78	10.5181	1	0	0	0	0
132	1760306	70	140.15	10.0037	4.80	13.18	35.0378	1	0	0	0	0
133	1760306	71	210.57	2.0000	4.80	103.83	52.6431	1	0	0	0	0
134	1760503	69	77.06	8.0002	4.80	8.02	19.4146	1	0	0	0	0
135	1760503	70	146.18	8.0002	4.80	16.40	36.2962	1	0	0	0	0
136	1820802	66	0.10	0.0000	1.80	0.00	0.0000	0	1	0	0	0
137	1820802	71	8180.00	1.7000	0.67	3586.24	9150.0000	0	0	0	1	0
138	1830801	70	432.00	27.0000	4.80	13.62	108.0000	0	0	0	0	1
139	1830801	71	5200.00	5.7800	1.80	904.36	5200.0000	0	0	0	1	1
140	1830801	72	71000.00	7.2000	1.80	9783.10	71000.0000	0	0	0	1	0
141	1830800	64	0.10	0.0000	0.33	0.00	0.3051	0	1	0	0	0
142	1840306	64	0.53	0.0000	0.33	0.00	1.5001	0	1	0	0	0
143	1840503	67	4.00	0.0000	1.80	0.00	4.0000	0	1	0	0	0
144	1880106	71	0.01	0.0000	1.80	0.00	0.0130	0	1	0	0	0
145	1880804	69	29.00	28.3100	4.80	1.10	7.2500	0	0	0	0	1
146	1880804	70	50.00	24.7000	4.80	2.82	12.5000	0	0	0	0	1
147	1880804	71	11.90	5.2000	1.80	2.27	11.9000	0	0	0	0	1
148	1880803	64	0.01	0.0000	1.80	0.00	0.0130	0	1	0	0	0
149	1880800	69	119.00	25.0000	4.80	4.63	29.7500	0	0	0	0	1
150	1880800	70	121.00	25.4000	4.80	4.78	30.2500	0	0	0	0	1
151	1880807	67	33000.00	3.7007	0.80	8001.70	66000.0000	0	0	0	1	0
152	1881004	69	86.00	27.2700	0.80	2.40	0.0000	0	1	0	0	1
153	1881004	70	16.30	27.0000	4.80	0.90	4.0750	0	0	0	0	1
154	1881103	71	0.10	0.0000	2.67	0.00	0.0004	0	1	0	0	0
155	1881202	69	90.00	21.0000	4.80	2.70	14.7500	0	0	0	0	1
156	1881202	70	129.00	34.0100	4.80	5.20	32.2500	0	0	0	0	1
157	1881400	69	40.10	28.0000	4.80	1.40	10.0250	0	0	0	0	1
158	1881400	70	60.00	30.0100	4.80	2.00	15.0000	0	0	0	0	1
159	1881400	71	2.45	6.2000	1.80	0.38	2.4500	0	0	0	0	1
160	1881500	70	22.20	28.0000	4.80	0.78	5.5500	0	0	0	0	1
161	1821700	67	5300.00	0.7130	1.20	7430.35	5300.0000	0	0	0	1	0
162	1831600	64	0.13	0.0000	0.67	0.00	0.2004	0	1	0	0	0
163	1832300	69	1200.00	28.0000	4.80	47.44	320.0000	0	0	0	0	1
134	[illegible]	67	2070.00	0.8676	0.67	3646.05	3186.0000	0	0	0	1	0

TABLE C-1

DUST SAMPLES EXCLUDED FROM ANALYSIS
(Continued)

	LBP_ID	SD_ID	PB_ADJ	SAMP_WGT	SD_VACU	PPM	SD_LOAD	WET_WIPE flg1	MSNG_DATA_FLG flg2	PPM_FLG flg3	LOAD_FLG flg4	WGT_FLG flg5
165	1962500	66	0.20	0.0000	2.80	0.00	0.1000	0	1	0	0	0
166	1963000	63	0.05	0.0000	4.80	0.00	0.0125	0	1	0	0	0
167	1963000	68	0.12	0.0000	4.80	0.00	0.0895	0	1	0	0	0
168	1963000	70	1.80	0.0000	4.80	0.00	0.2500	0	1	0	0	0
169	1980802	61	0.00	0.0024	4.80	0.00	0.0000	0	1	0	0	0
170	1980500	63	11.71	0.0000	4.80	0.00	2.0085	0	1	0	0	0
171	2030802	66	0.06	0.0000	0.67	0.00	1.4436	0	1	0	0	0
172	2110808	67	0.37	0.0000	1.33	0.00	0.2775	0	1	0	0	0
173	2141805	61	0.46	0.0000	4.80	0.00	0.1125	0	1	0	0	0
174	2240405	64	2382.44	0.3085	0.67	7780.40	3672.8570	0	0	0	1	0
175	2280805	65	2180.00	0.7418	0.67	2830.96	3150.0000	0	0	0	1	0
176	2230805	65	3668.98	0.2782	0.81	13332.06	4867.7410	0	0	0	1	0
177	2342802	64	0.23	0.0000	0.83	0.00	0.2722	0	1	0	0	0
178	2420400	64	1.81	0.0000	0.67	0.00	2.7100	0	1	0	0	0
179	2420400	65	2000.00	0.3110	0.67	6730.30	3125.0000	0	0	0	1	0
180	2441088	65	4830.00	0.1017	1.80	44642.77	4830.0000	0	0	0	1	0
181	2511088	64	0.16	0.0000	0.67	0.00	0.3402	0	1	0	0	0
182	2531004	61	0.30	0.0000	4.80	0.00	0.6747	0	1	0	0	0
183	2551000	64	0.32	0.0000	0.67	0.00	0.4000	0	1	0	0	0
184	2611101	61	0.10	0.0000	4.80	0.00	0.0054	0	1	0	0	0
185	2621704	63	0.43	0.0000	4.80	0.00	0.1075	0	1	0	0	0
186	2622007	63	0.25	0.0000	4.80	0.00	0.0085	0	1	0	0	0
187	2622007	70	0.11	0.0000	4.80	0.00	0.0275	0	1	0	0	0
188	2622007	71	0.06	0.0000	0.33	0.00	0.1000	0	1	0	0	0
189	2651600	65	1.06	0.0000	0.67	0.00	1.5750	0	1	0	0	0
190	2710101	64	0.25	0.0000	0.67	0.00	0.3736	0	1	0	0	0
191	2730700	63	0.15	0.0000	4.80	0.00	0.0372	0	1	0	0	0
192	2751402	62	934.94	0.0002	4.80	114017.35	232.7204	0	0	1	0	0
193	2831008	66	0.07	0.0000	1.80	0.00	0.0050	0	1	0	0	0
194	2832004	64	0.20	0.0000	0.67	0.00	0.2985	0	1	0	0	0
195	2840403	64	1.51	0.0000	0.33	0.00	4.5300	0	1	0	0	0
196	2841203	66	10.10	0.0000	1.80	0.00	10.1000	0	1	0	0	0
197	2841401	67	0.11	0.0000	1.80	0.00	0.1000	0	1	0	0	0
198	2841500	64	0.14	0.0000	0.67	0.00	0.2145	0	1	0	0	0
199	2841500	67	0.12	0.0000	1.80	0.00	0.1100	0	1	0	0	0
200	2980102	64	0.31	0.0000	1.33	0.00	0.2230	0	1	0	0	0
201	3011103	65	2147.11	0.0004	0.67	30000.30	2230.0630	0	0	0	1	0
202	3011103	67	4730.04	0.0005	1.80	83664.06	4730.0360	0	0	0	1	0
203	3011805	63	0.50	0.0000	4.80	0.00	0.1257	0	1	0	0	0
204	3020401	61	0.00	0.0740	4.80	0.00	0.0000	0	1	0	0	0
205	3020401	64	0.30	0.0000	1.80	0.00	0.3667	0	1	0	0	0
206	3051600	62	1.80	0.0000	4.80	0.00	0.3702	0	1	0	0	0
207	3051600	63	0.44	0.0000	4.80	0.00	0.1001	0	1	0	0	0
208	3051600	67	0.29	0.0000	1.80	0.00	0.2985	0	1	0	0	0
Total								52	112	5	23	49

Notes:

1 WET_WIPE = 1 if the sample was noted as a wet wipe in the documentation.

2 MSNG_DATA_FLG = 1 if PPM = 0 or SD_LOAD = 0. Missing values for PB_ADJ, SAMP_WGT, SD_VACU, PPM, and SD_LOAD were recorded as zero.

3 PPM_FLG = 1 if PPM >100,000.

4 LOAD_FLG = 1 if SD_LOAD > 2,000.

5 WGT_FLG = 1 if SAMP_WGT > 20 for floor samples, SAMP_WGT > 5 for sill samples, or SAMP_WGT > 8 for well samples.

TABLE C-2

SOIL SAMPLES WITH GREATER THAN 2,6000 PPM OF LEAD, DROPPED FROM ANALYSIS FILE

LABCODE	LBP_ID	PB_ADJ	SAMP_WGT	SD_ID	PPM
M	1010909	1630	0.5465	81	2,983
C	1831106	3530	1.0007	82	3,528
C	1831304	3530	1.0007	82	3,528
M	1820802	1890	0.5054	82	3,740
M	1820802	2580	0.5370	81	4,804
M	1011709	2500	0.5026	81	4,975
C	950402	5900	0.9965	81	5,921
C	2441509	6006	1.0006	83	6,002
M	1010503	4040	0.5650	81	7,150
C	3011905	8610	1.0397	82	8,281
M	1011501	6260	0.5079	81	12,325
M	1011501	10900	0.5131	82	21,243
M	1010503	22000	0.5102	82	43,120

APPENDIX D

XRF Validation Data and the
Distribution of the
Adjusted XRF Readings

FIGURE D-1

BASELINE VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #34 ON STEEL SUBSTRATE

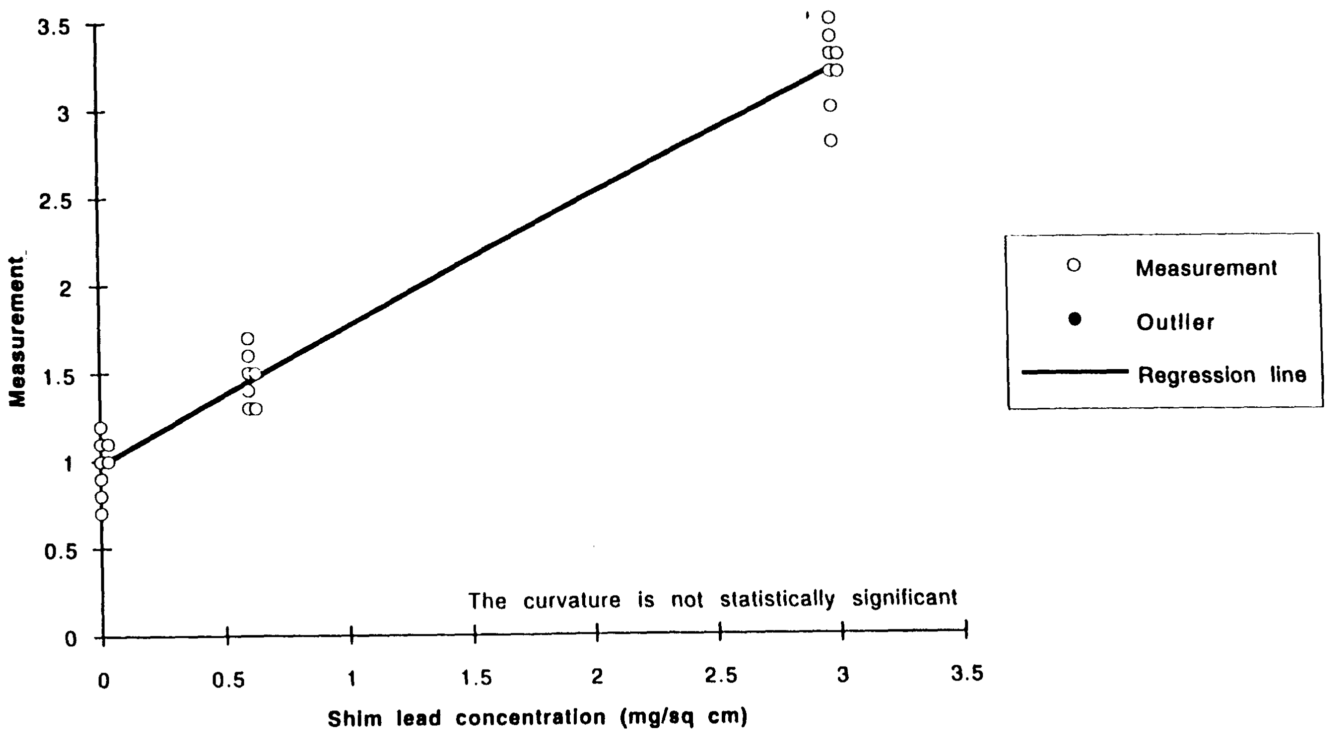


FIGURE D-2

BASELINE VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #35 ON STEEL SUBSTRATE

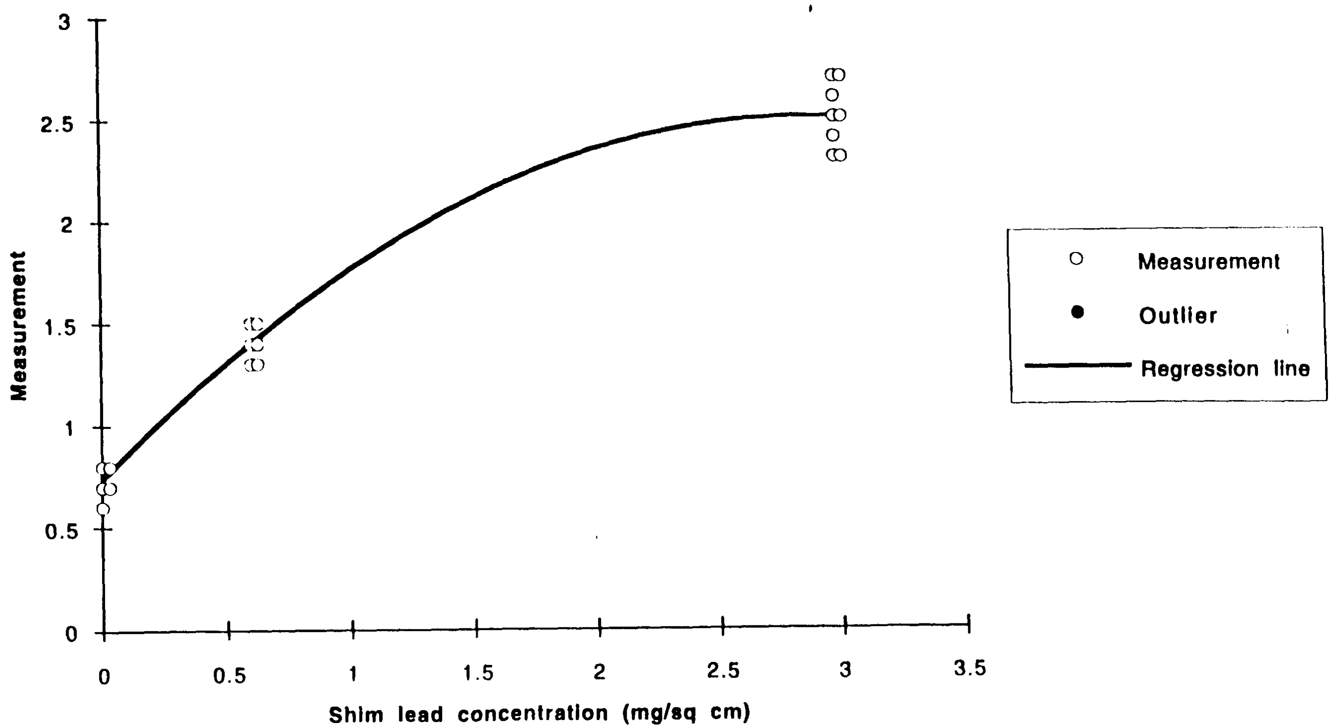
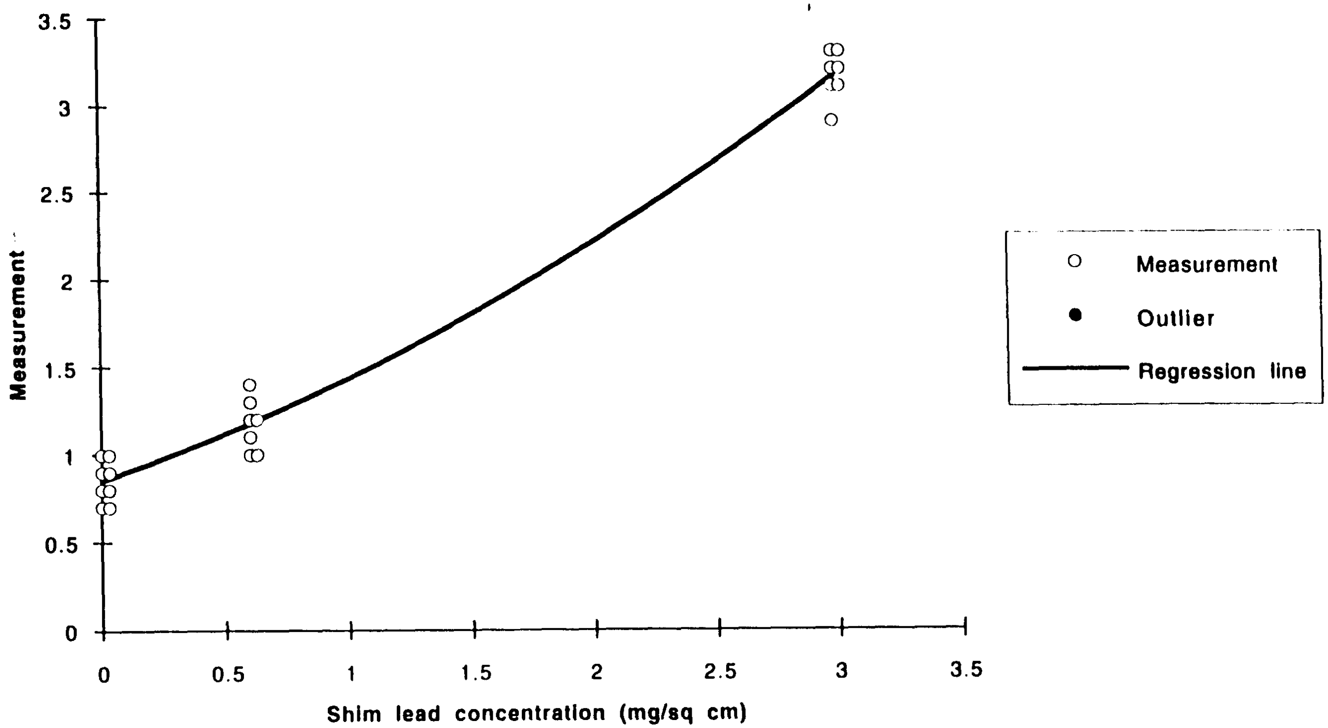


FIGURE D-3

BASELINE VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON STEEL SUBSTRATE



D-3

FIGURE D-4

CLOSEOUT VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #34 ON STEEL SUBSTRATE

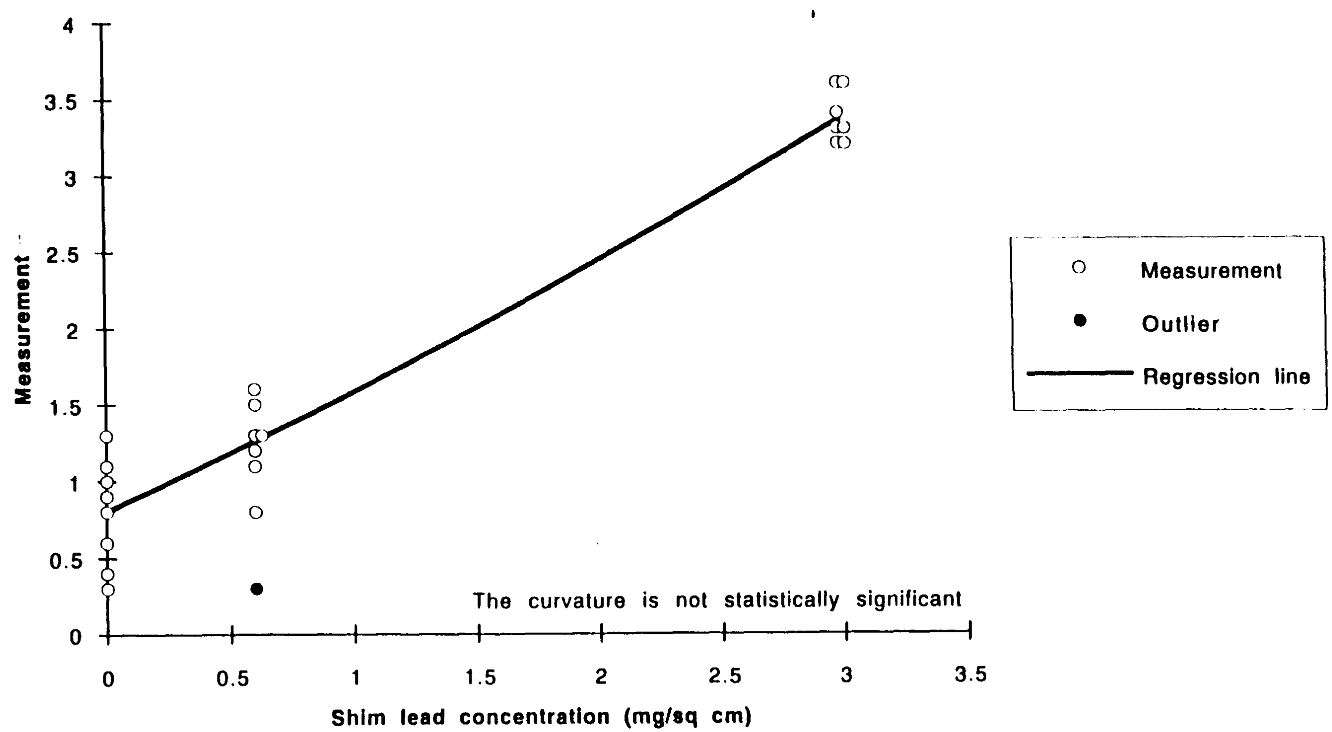


FIGURE D-5

CLOSEOUT VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #35 ON STEEL SUBSTRATE

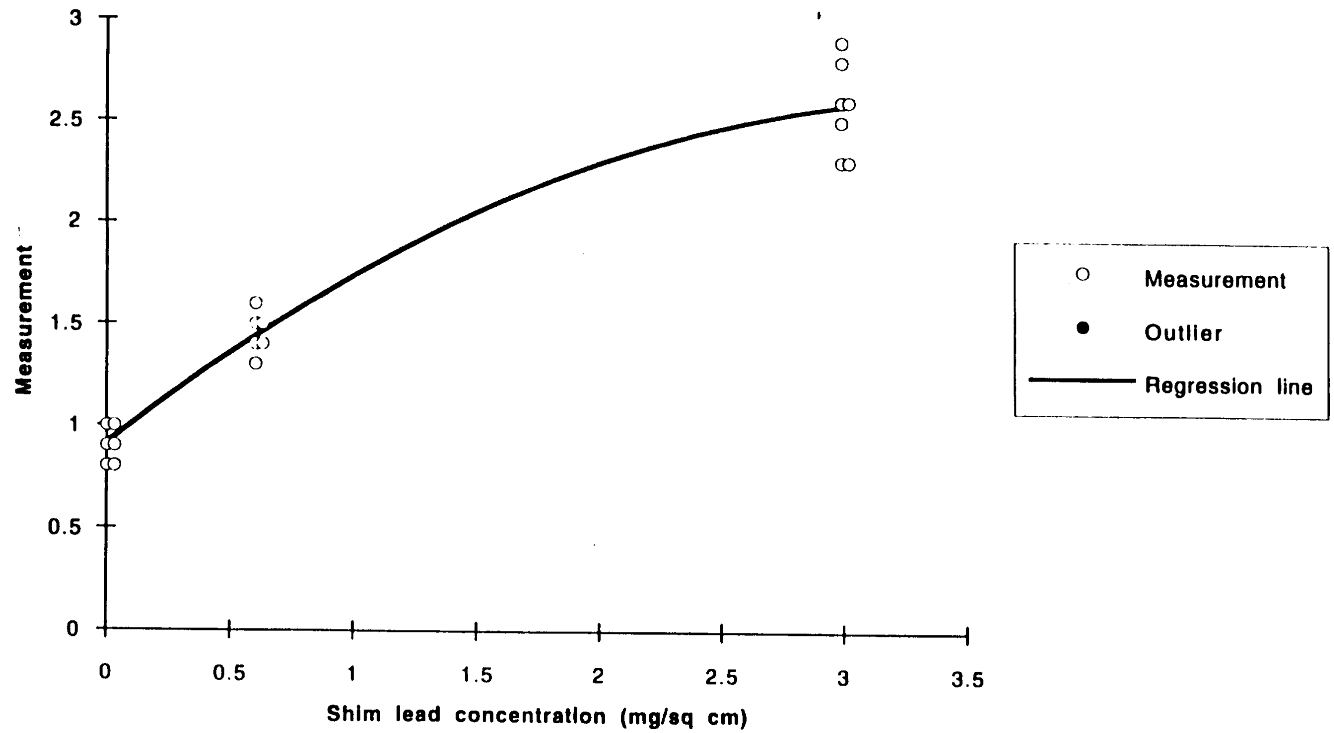


FIGURE D-6

CLOSEOUT VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON STEEL SUBSTRATE

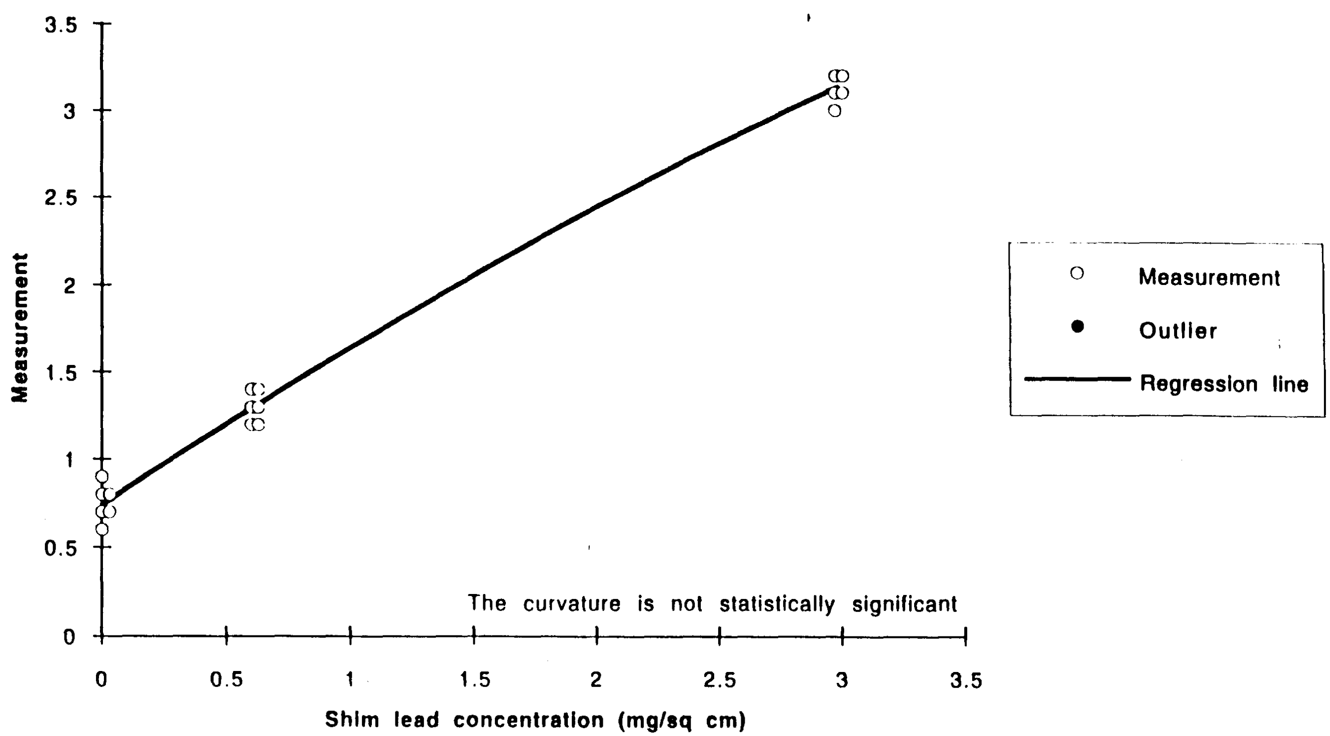


FIGURE D-7

CLOSEOUT VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #38 ON STEEL SUBSTRATE

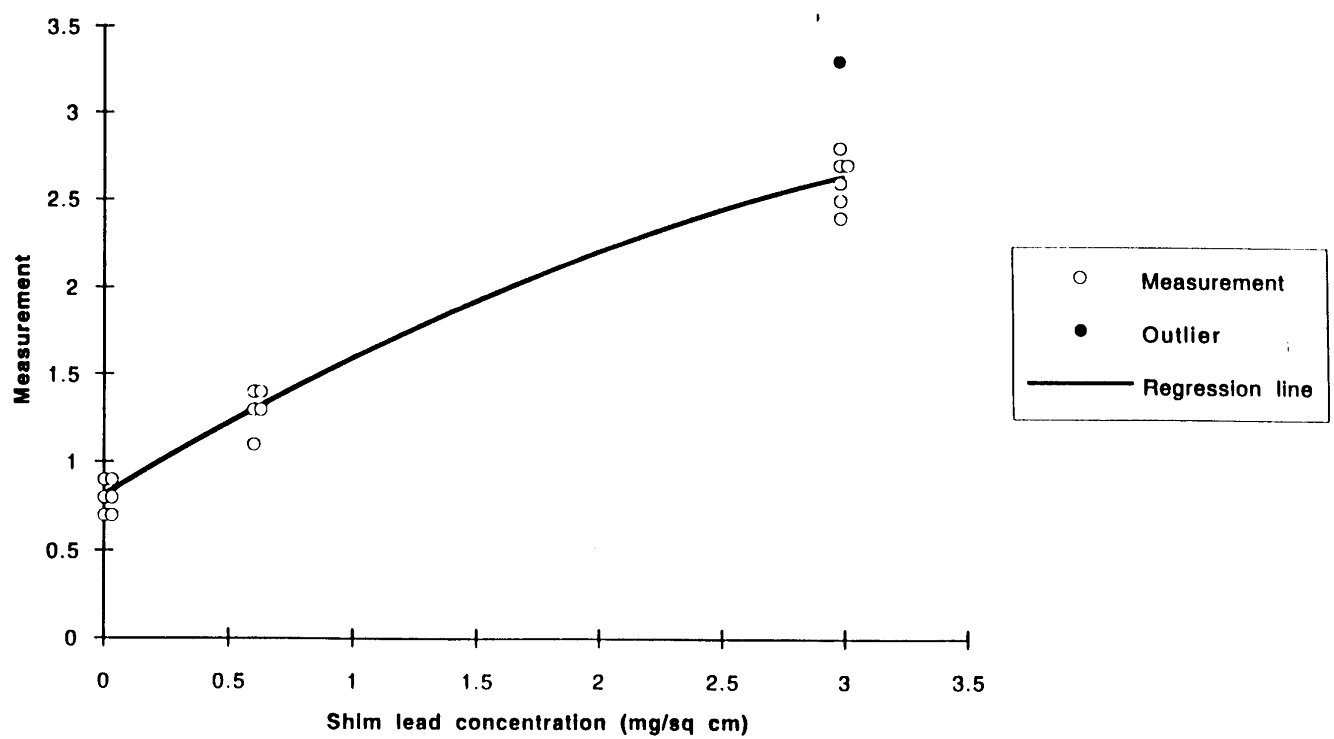
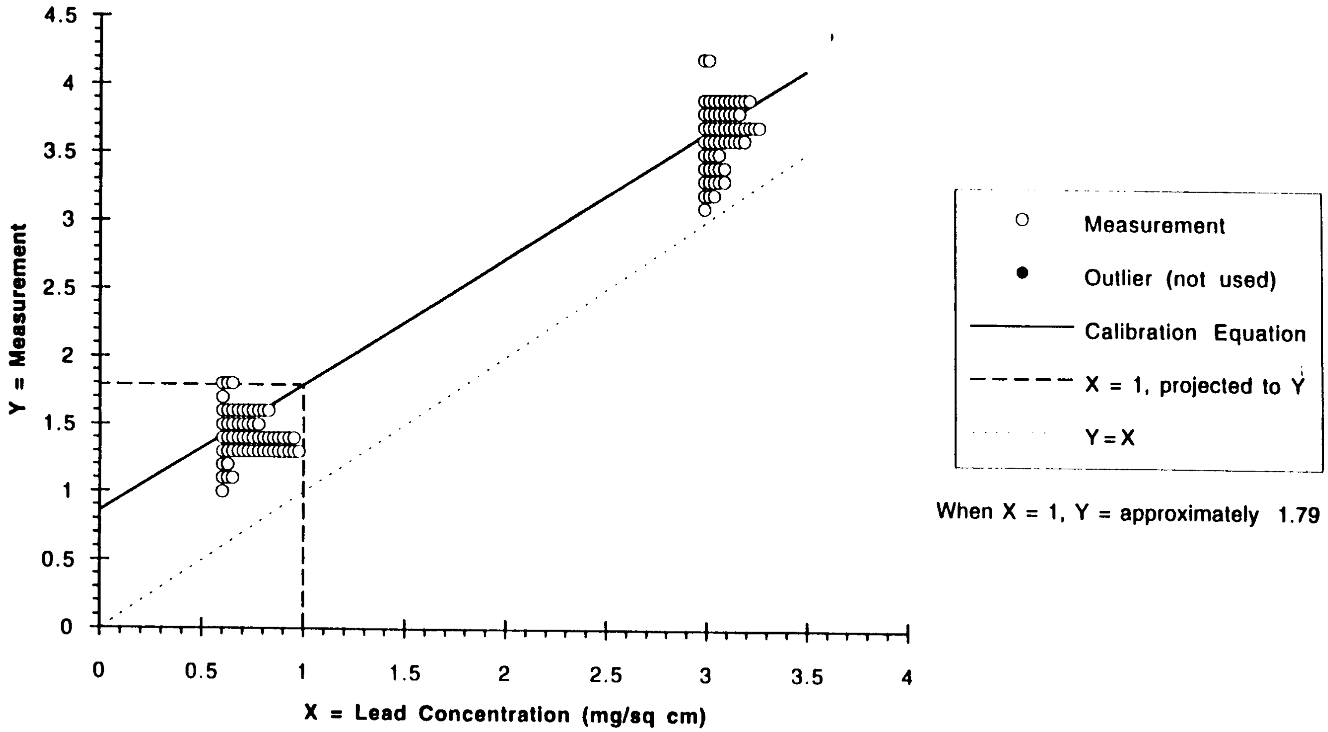


FIGURE D-8

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #34 ON STEEL

Lead concentration (mg/sq cm) = -0.9472 + 1.0748*Measurement + 0.00106*(Days since 2/1/90)



When X = 1, Y = approximately 1.79

FIGURE D-9

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #35 ON STEEL

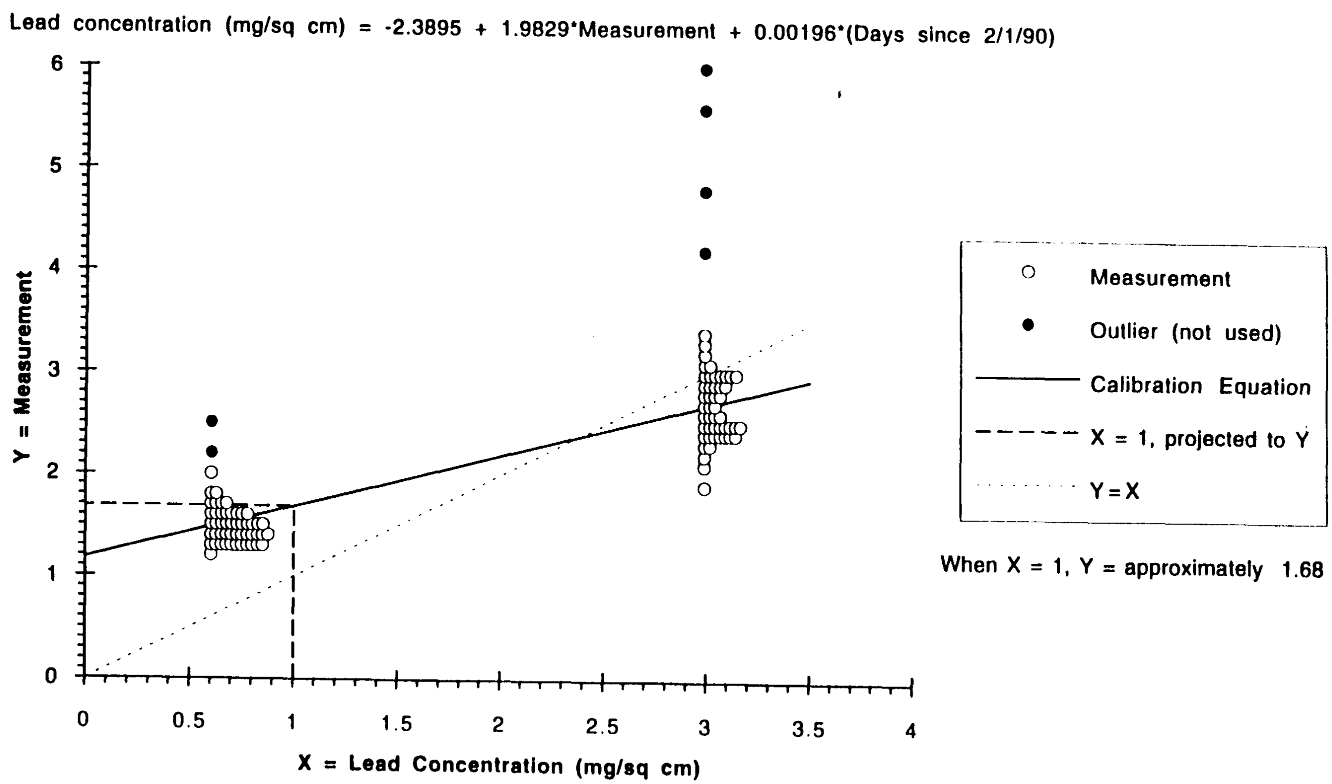


FIGURE D-10
VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #36 ON STEEL

Lead concentration (mg/sq cm) = -1.0285 + 1.0580*Measurement + 0.00105*(Days since 2/1/90)

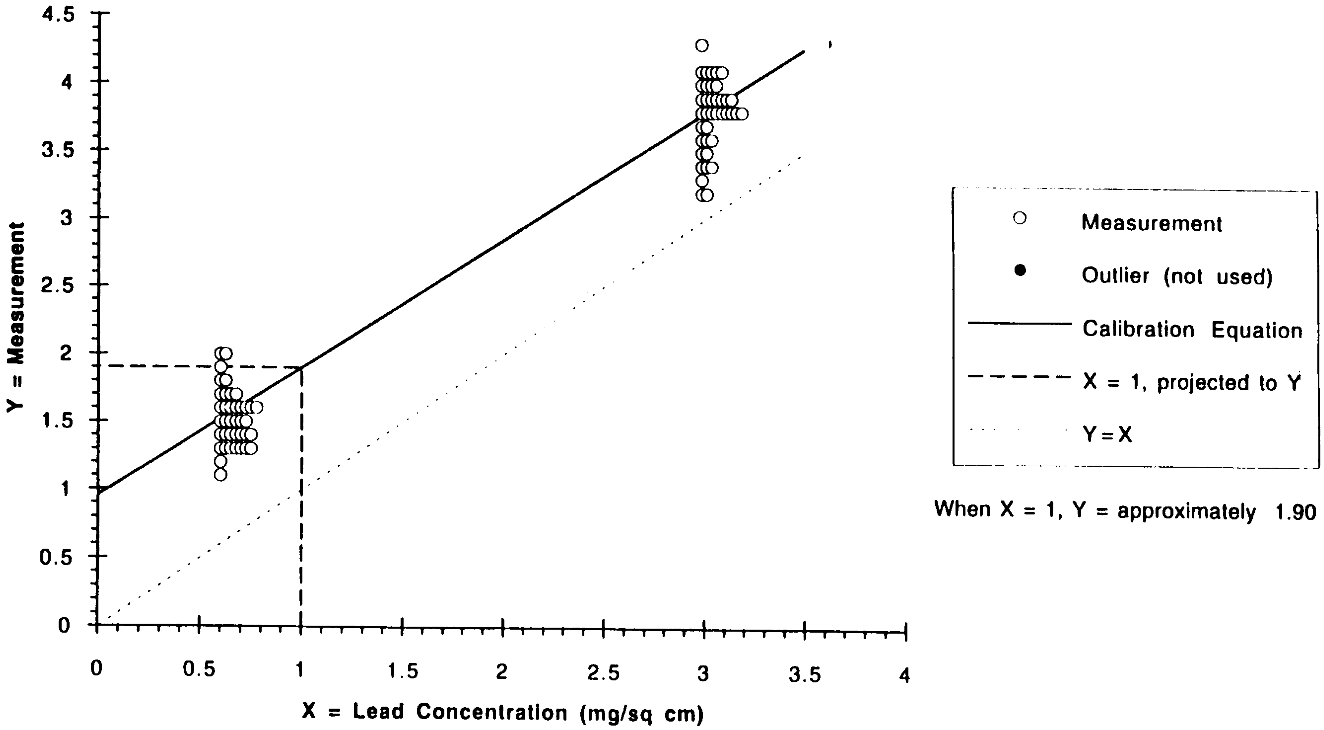
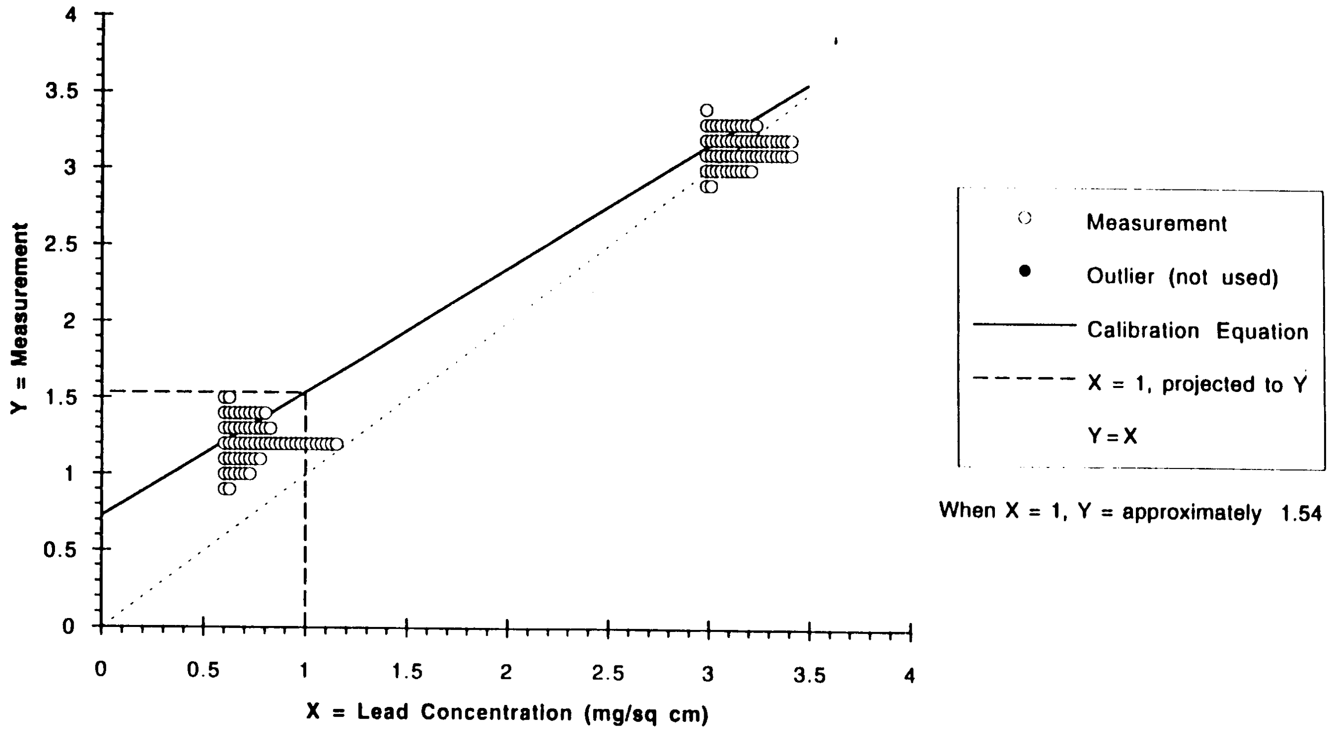


FIGURE D-11

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON STEEL

Lead concentration (mg/sq cm) = -0.9274 + 1.2351*Measurement + 0.00122*(Days since 2/1/90)

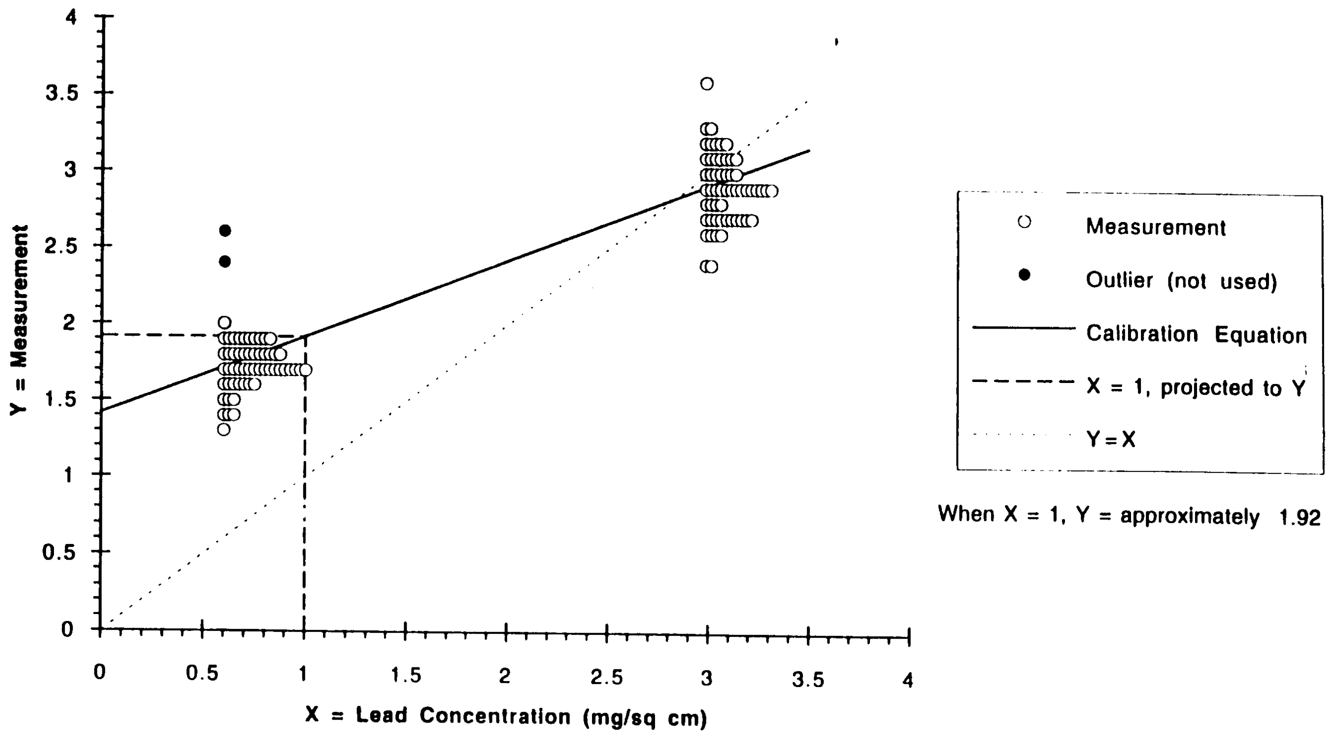


When X = 1, Y = approximately 1.54

FIGURE D-12

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #38 ON STEEL

Lead concentration (mg/sq cm) = -2.8854 + 2.0049*Measurement + 0.00198*(Days since 2/1/90)

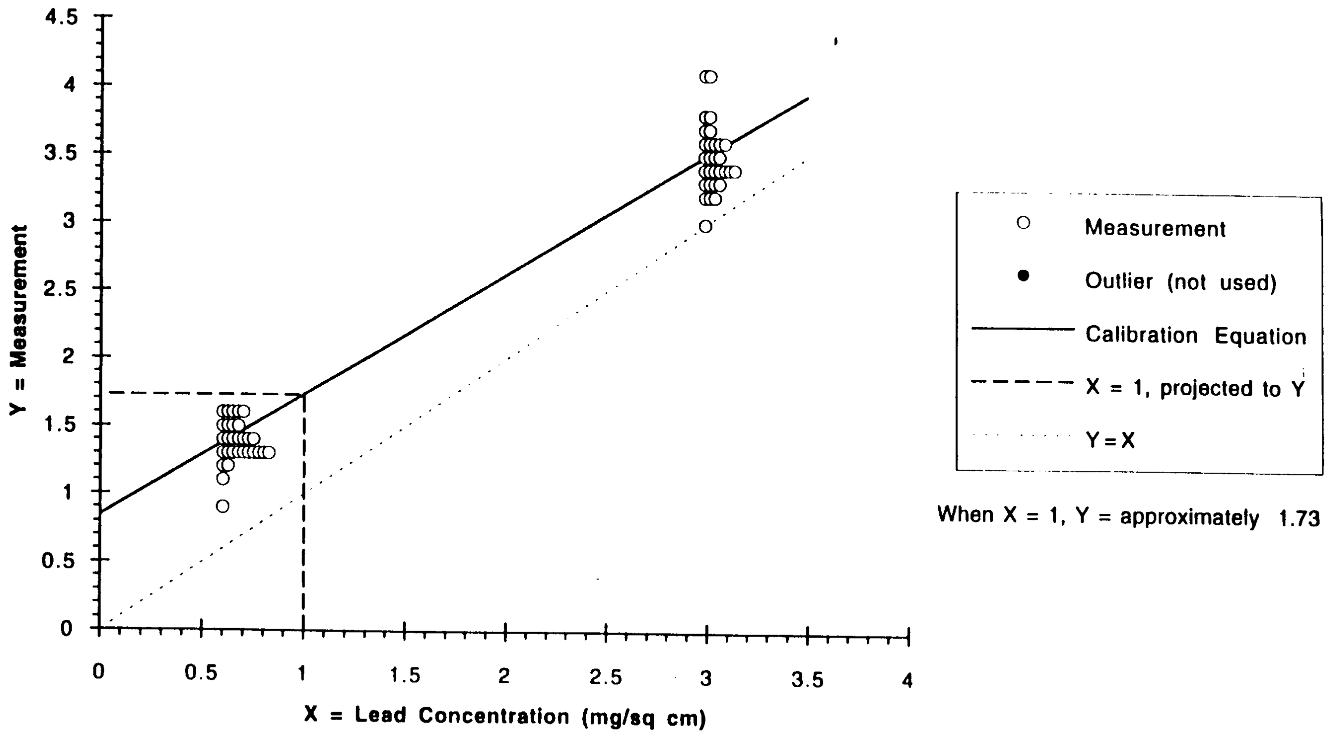


When X = 1, Y = approximately 1.92

FIGURE D-13

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #39 ON STEEL

Lead concentration (mg/sq cm) = -0.9832 + 1.1274*Measurement + 0.00112*(Days since 2/1/90)



When X = 1, Y = approximately 1.73

FIGURE D-14

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #41 ON STEEL

Lead concentration (mg/sq cm) = -0.3167 + 1.3345*Measurement + 0.00132*(Days since 2/1/90)

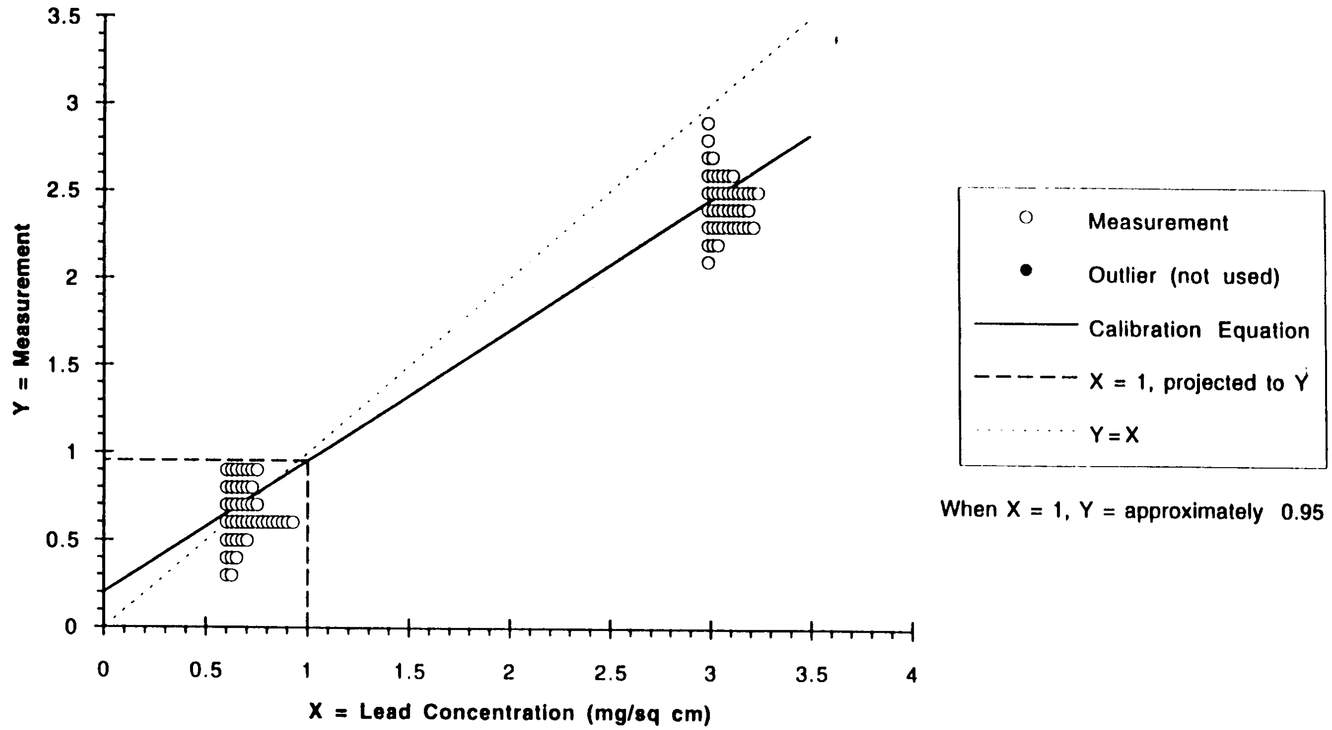


FIGURE D-15

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #32 ON WOOD

Lead concentration (mg/sq cm) = 0.5773 + 0.6748*Measurement + 0.00067*(Days since 2/1/90)

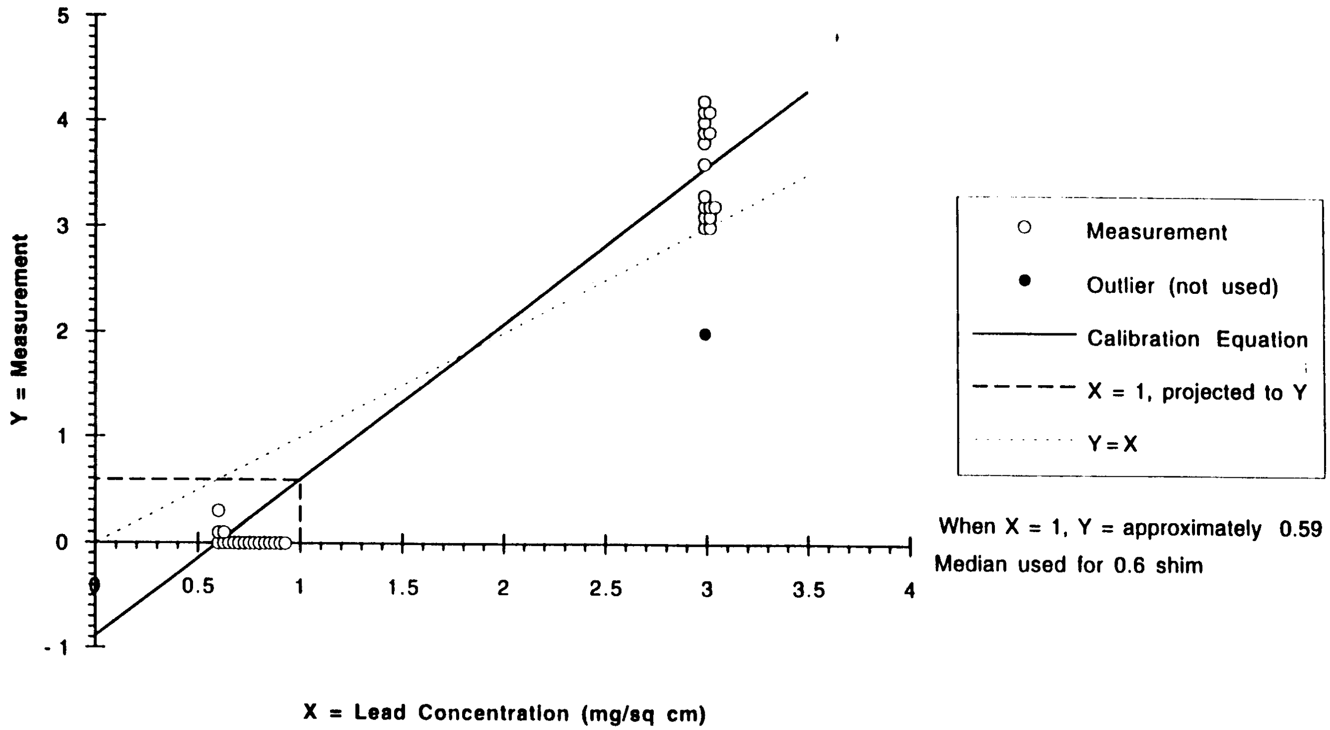


FIGURE D-16

VALIDATION MEASUREMENT BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #34 ON WOOD

Lead concentration (mg/sq cm) = 0.5108 + 0.7767*Measurement + 0.00077*(Days since 2/1/90)

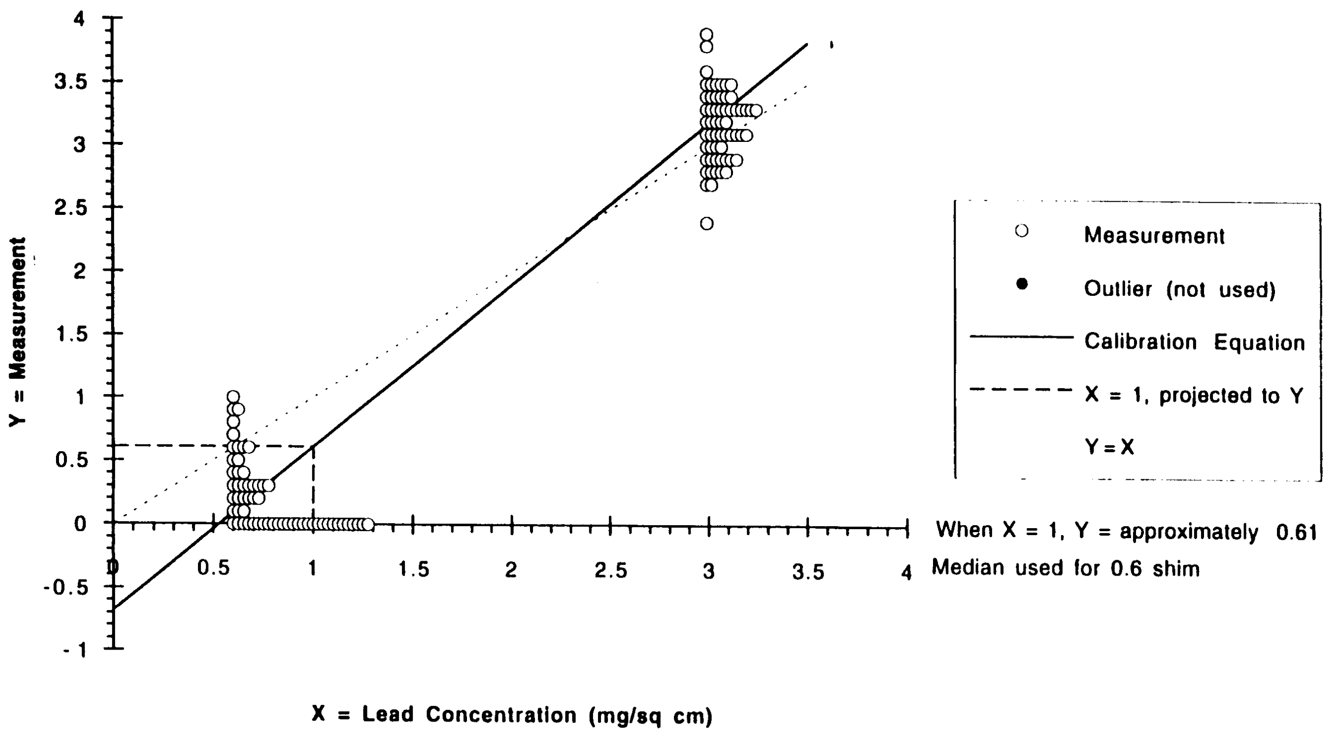


FIGURE D-17

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #35 ON WOOD

Lead concentration (mg/sq cm) = 0.5776 + 0.6842*Measurement + 0.00068*(Days since 2/1/90)

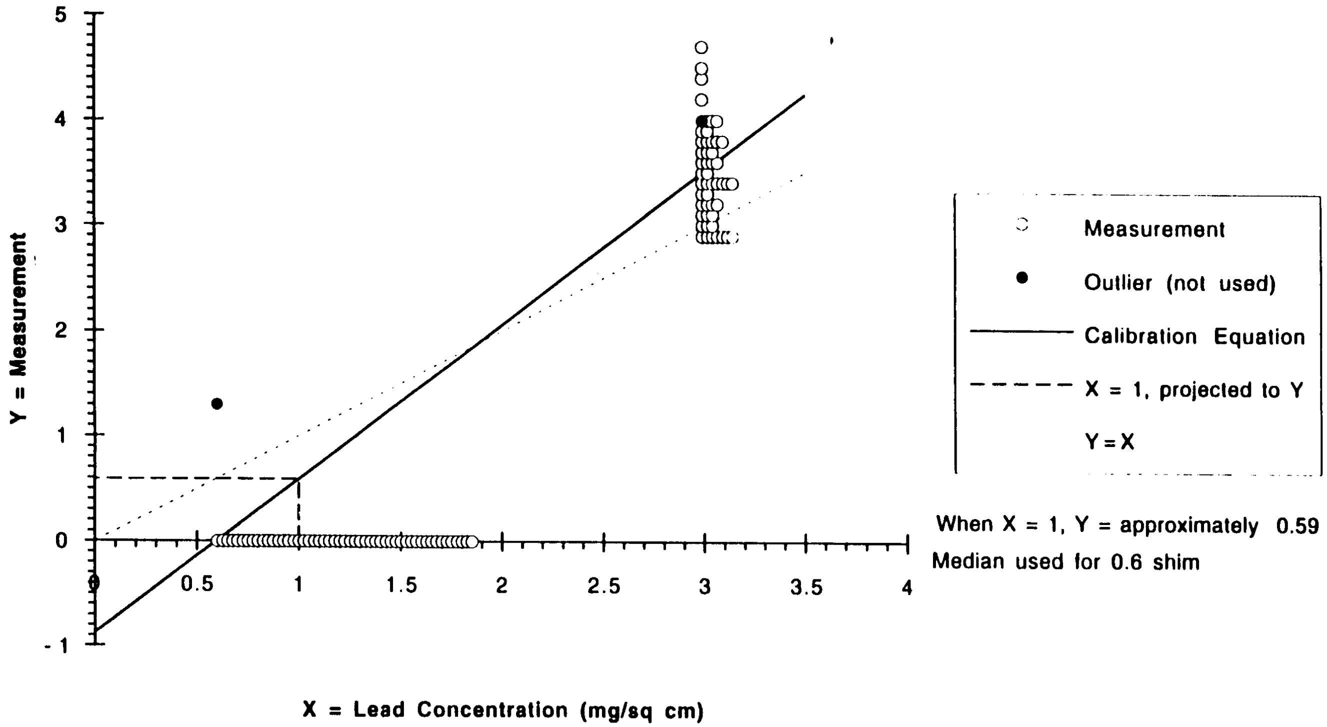
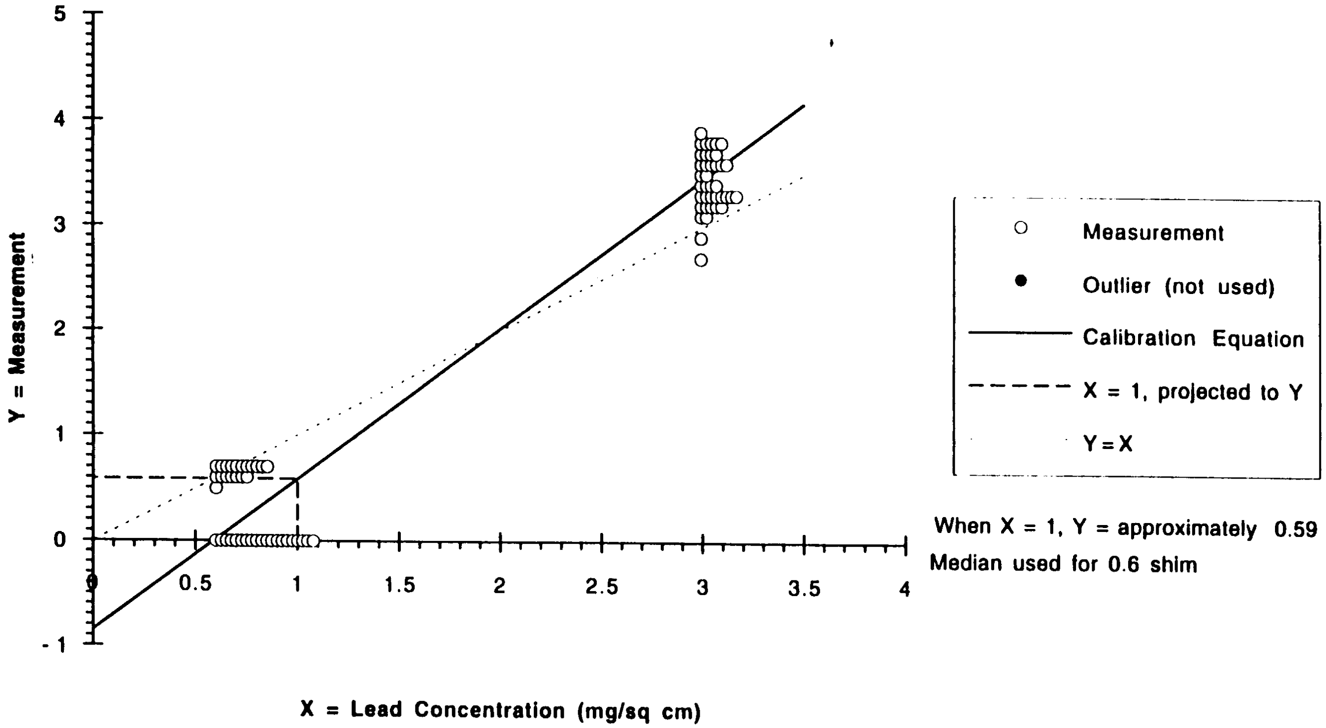


FIGURE D-18

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #36 ON WOOD

Lead concentration (mg/sq cm) = 0.5779 + 0.6977*Measurement + 0.00069*(Days since 2/1/90)



When X = 1, Y = approximately 0.59

Median used for 0.6 shim

FIGURE D-19

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON WOOD

Lead concentration (mg/sq cm) = 0.2390 + 0.8879*Measurement + 0.00088*(Days since 2/1/90)

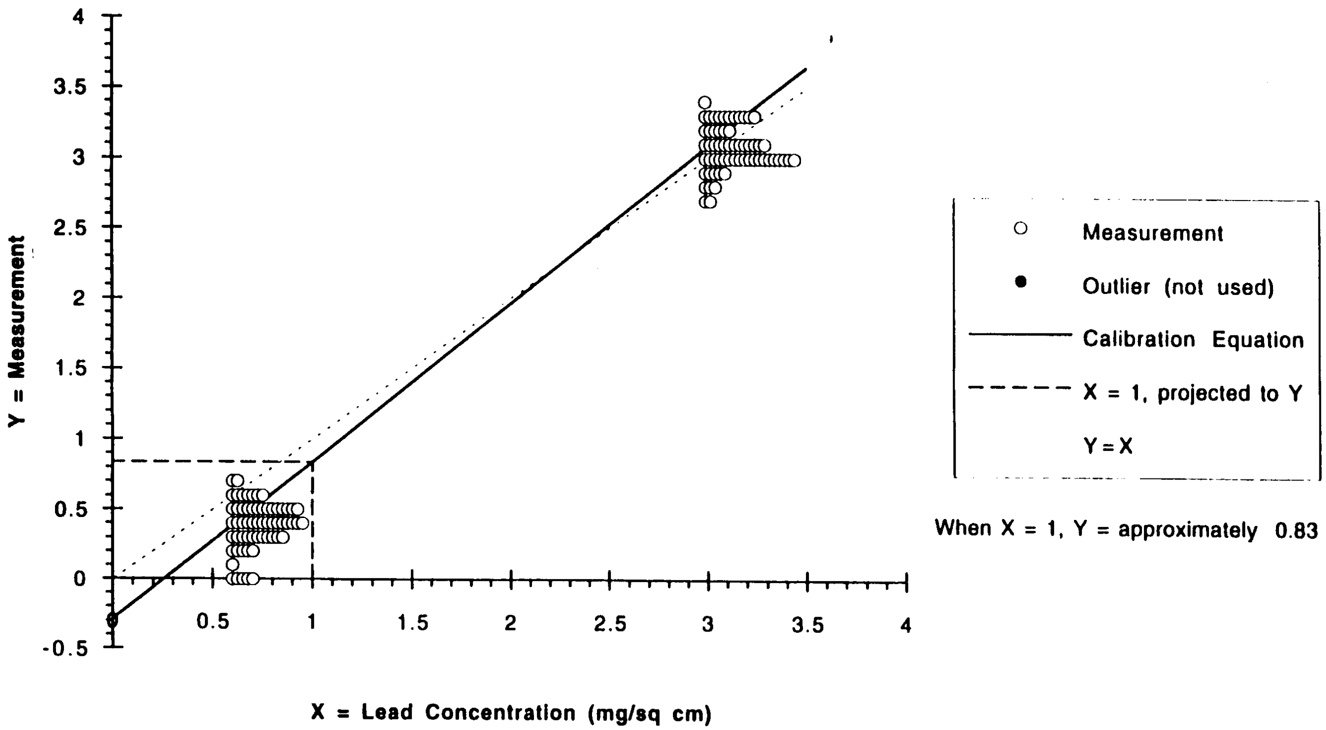
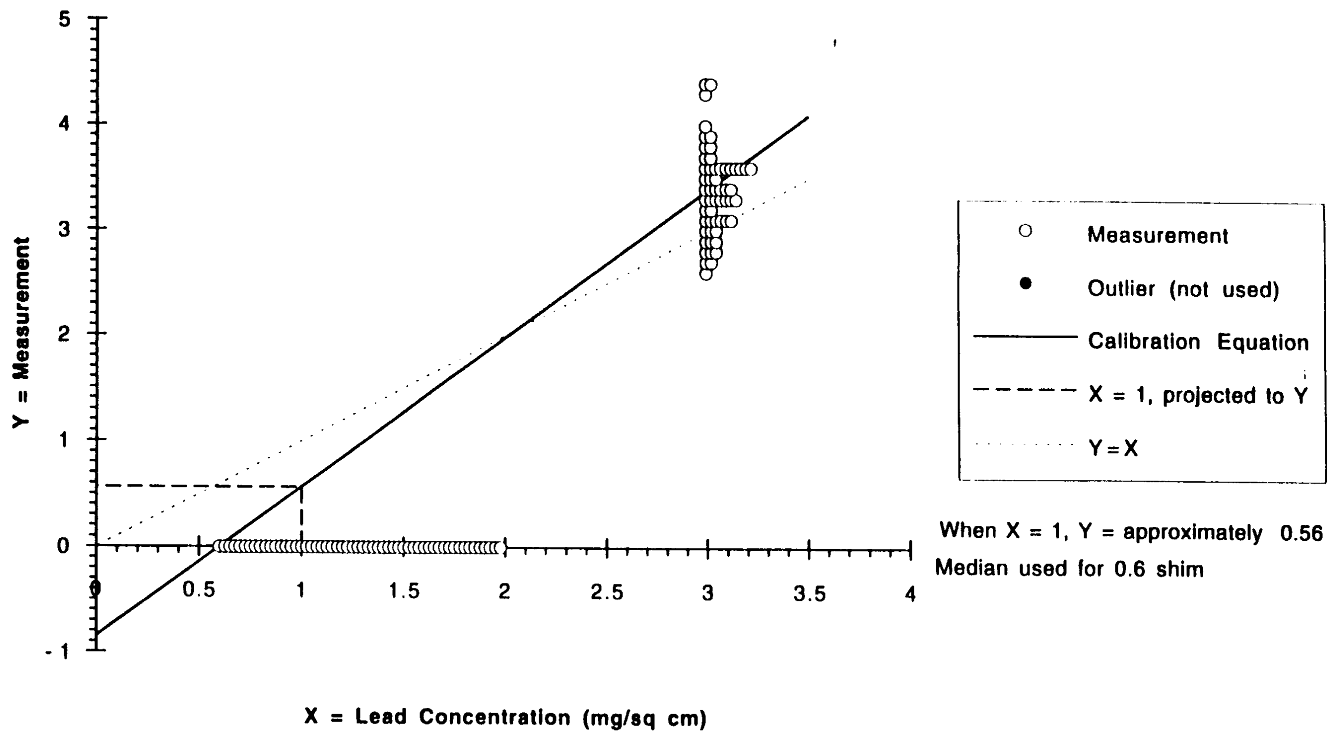


FIGURE D-20

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #38 ON WOOD

Lead concentration (mg/sq cm) = 0.5849 + 0.7081*Measurement + 0.00070*(Days since 2/1/90)



When X = 1, Y = approximately 0.56
Median used for 0.6 shim

FIGURE D-21

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #39 ON WOOD

Lead concentration (mg/sq cm) = 0.4214 + 0.7530*Measurement + 0.00075*(Days since 2/1/90)

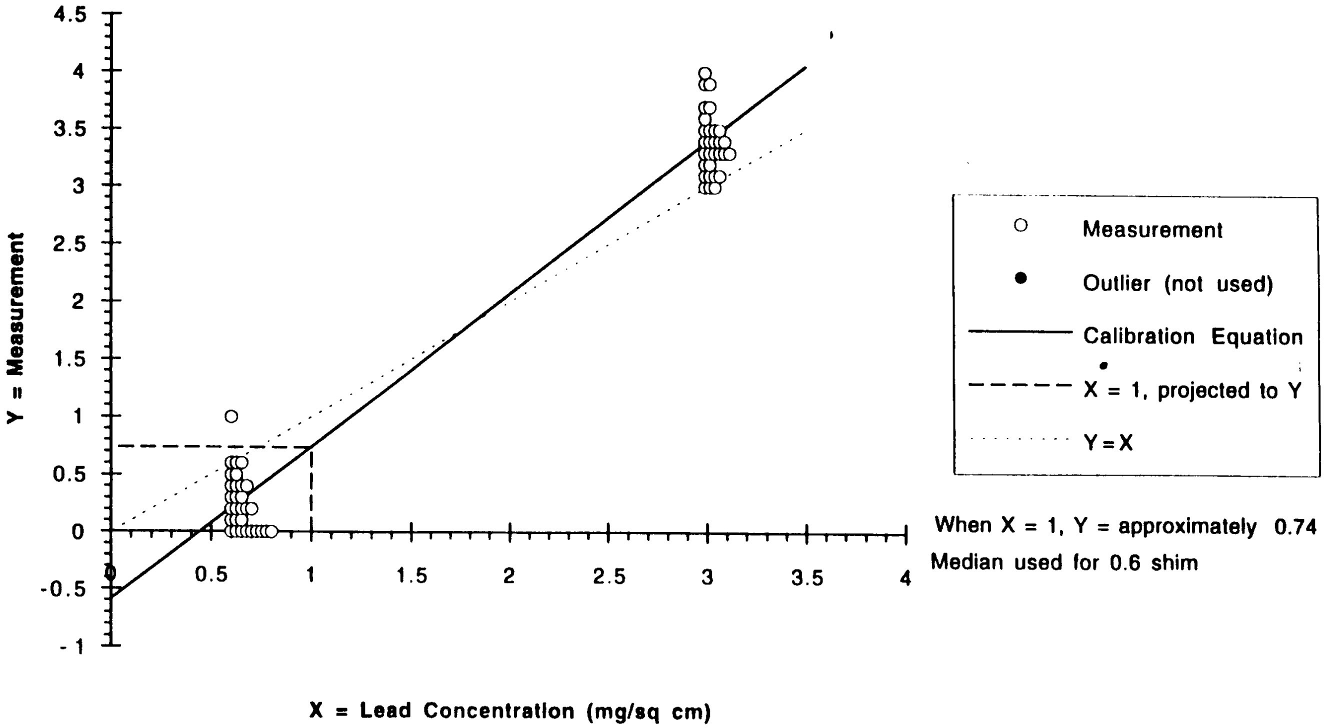


FIGURE D-22

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #41 ON WOOD

Lead concentration (mg/sq cm) = 0.0933 + 1.4646*Measurement + 0.00145*(Days since 2/1/90)

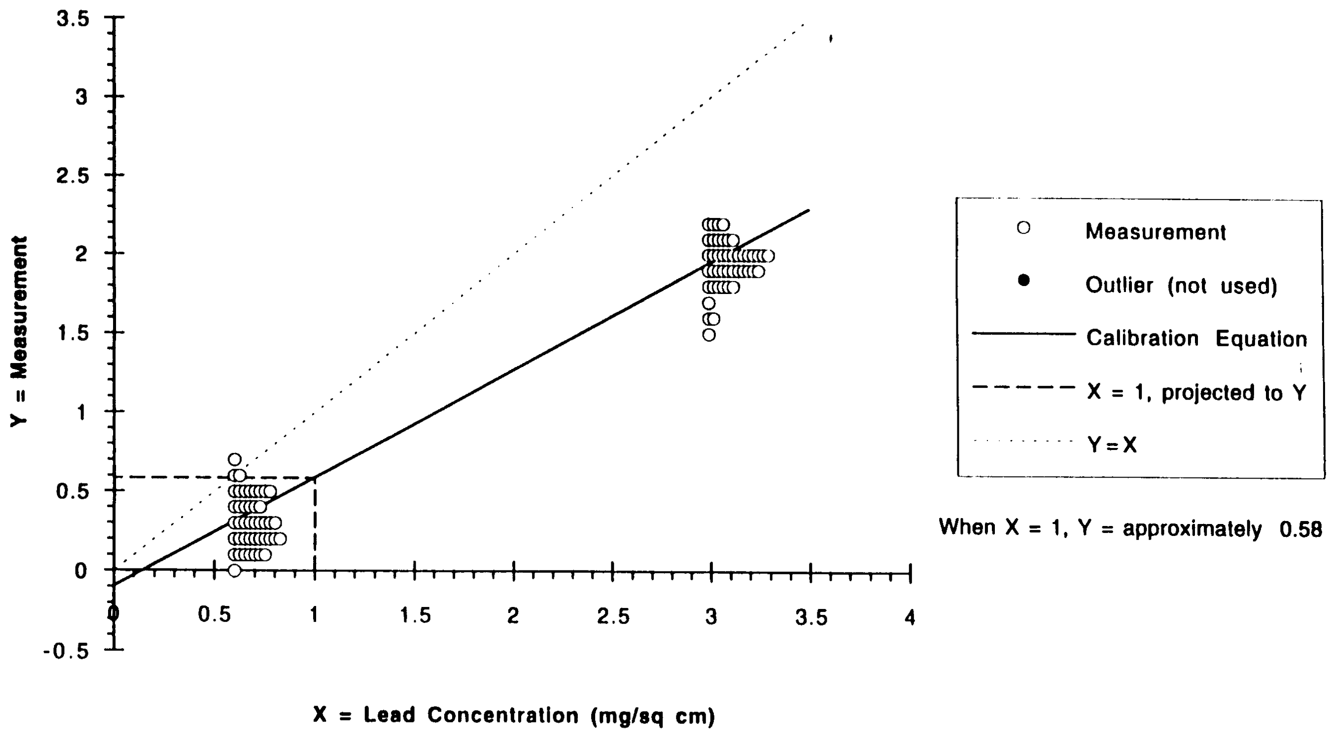


FIGURE D-23

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #32 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5440 + 1.6648*Measurement + 0.00165*(Days since 2/1/90)

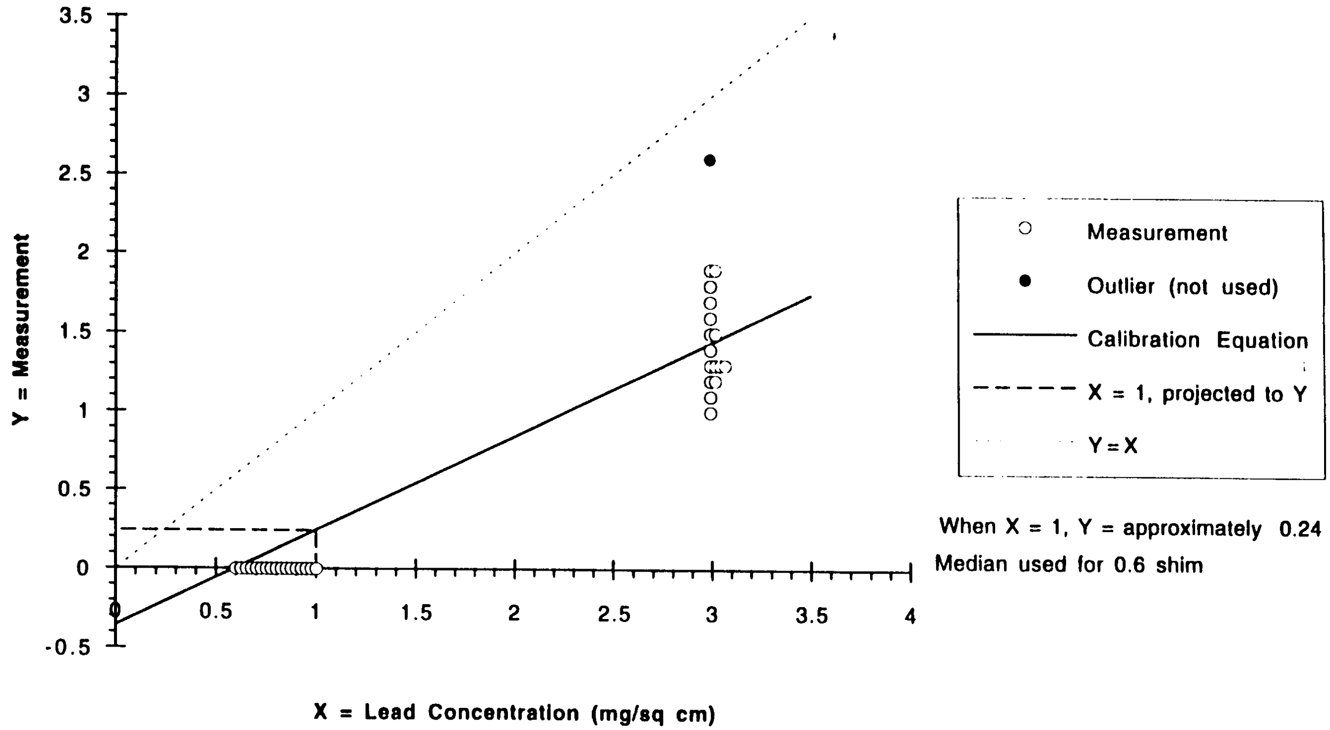


FIGURE D-24

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #34 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5529 + 1.4138*XRF reading + 0.00205*(Days since 2/1/90)

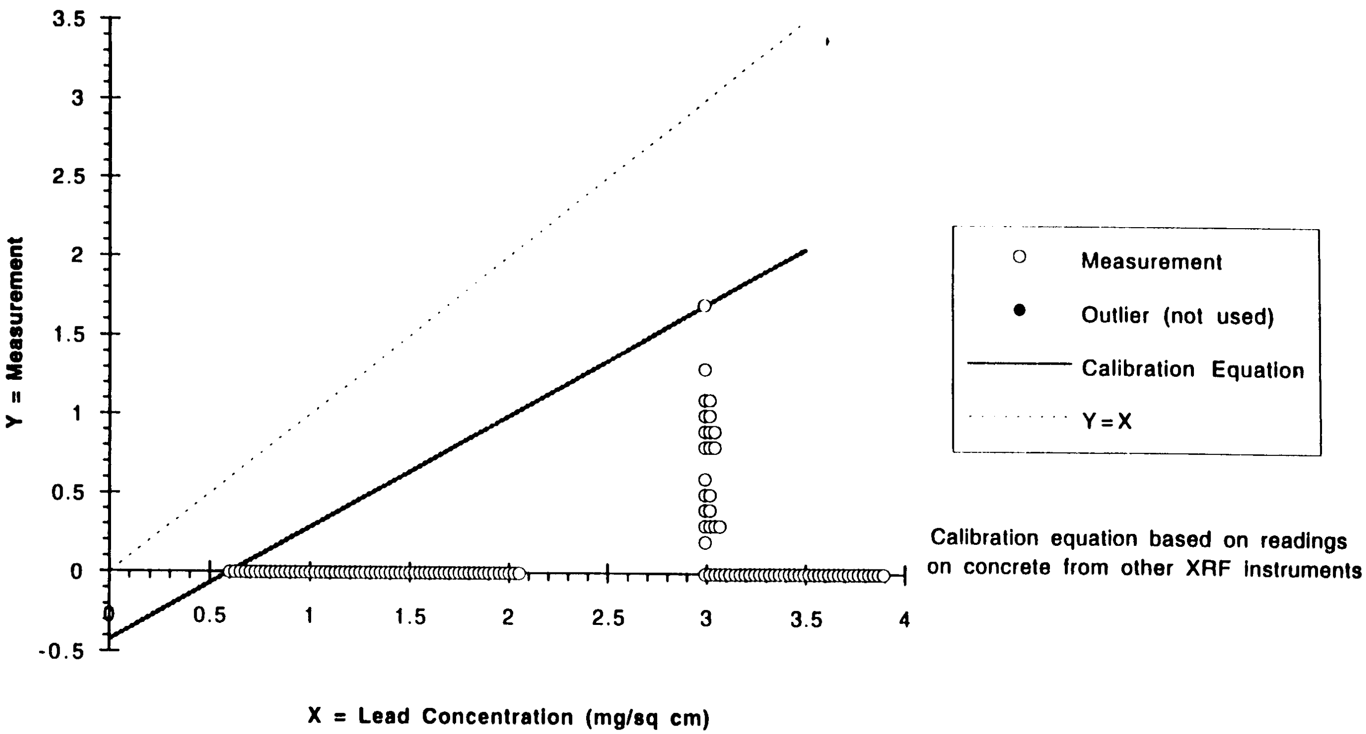
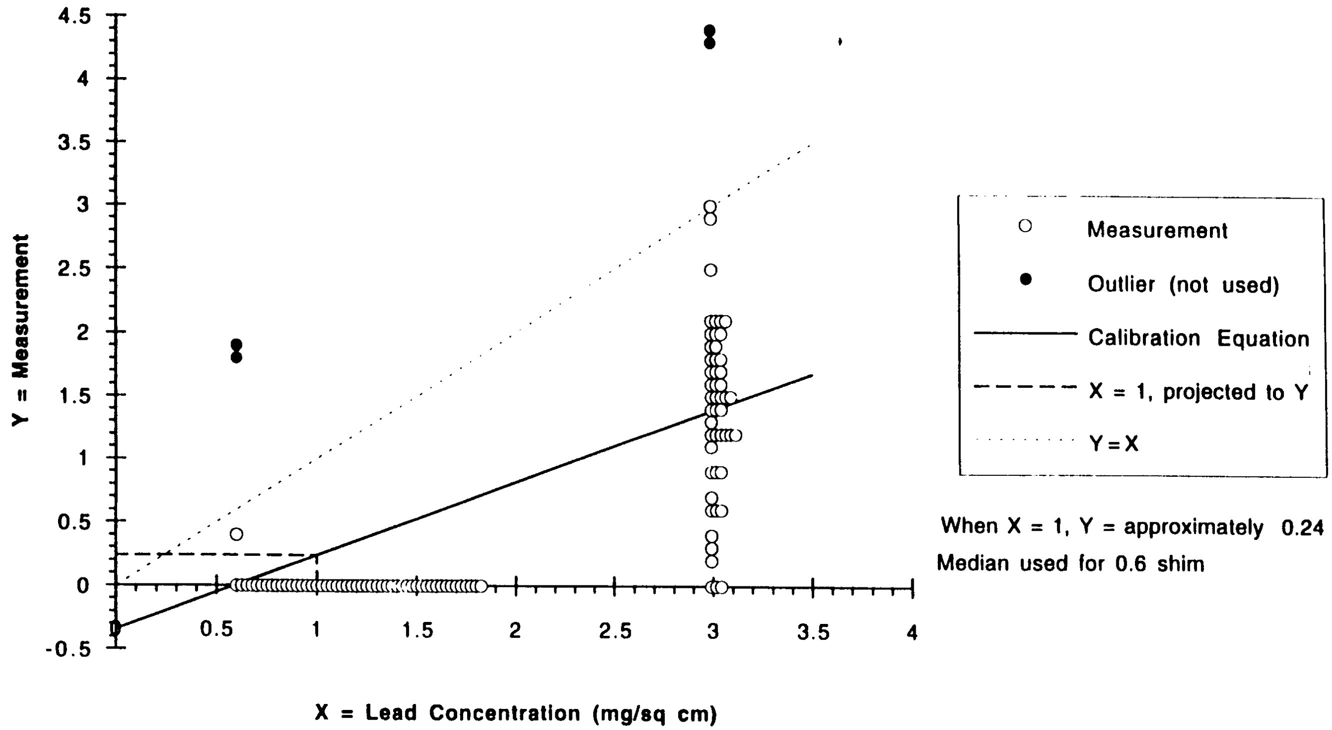


FIGURE D-25

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #35 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5435 + 1.7305*Measurement + 0.00171*(Days since 2/1/90)



When X = 1, Y = approximately 0.24

Median used for 0.6 shim

FIGURE D-26

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #36 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5692 + 1.5546*Measurement + 0.00154*(Days since 2/1/90)

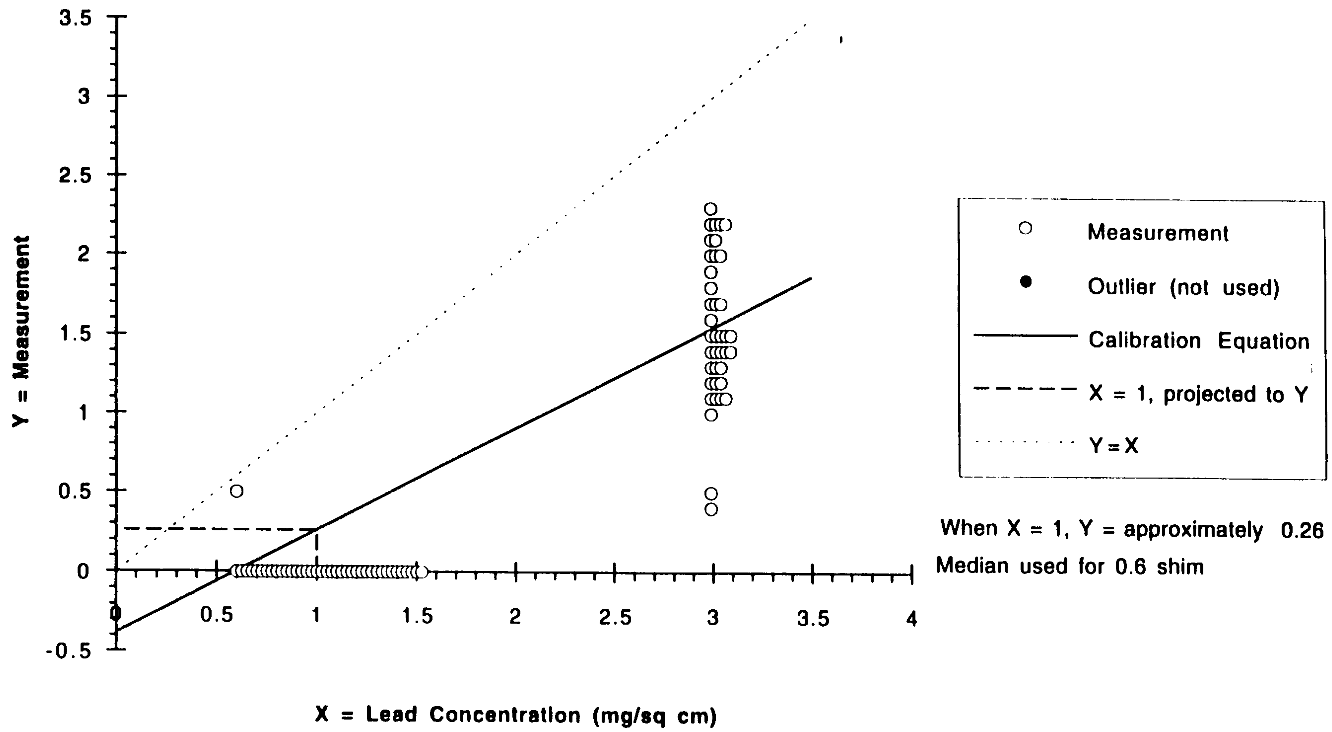


FIGURE D-27

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5475 + 2.2547*Measurement + 0.00223*(Days since 2/1/90)

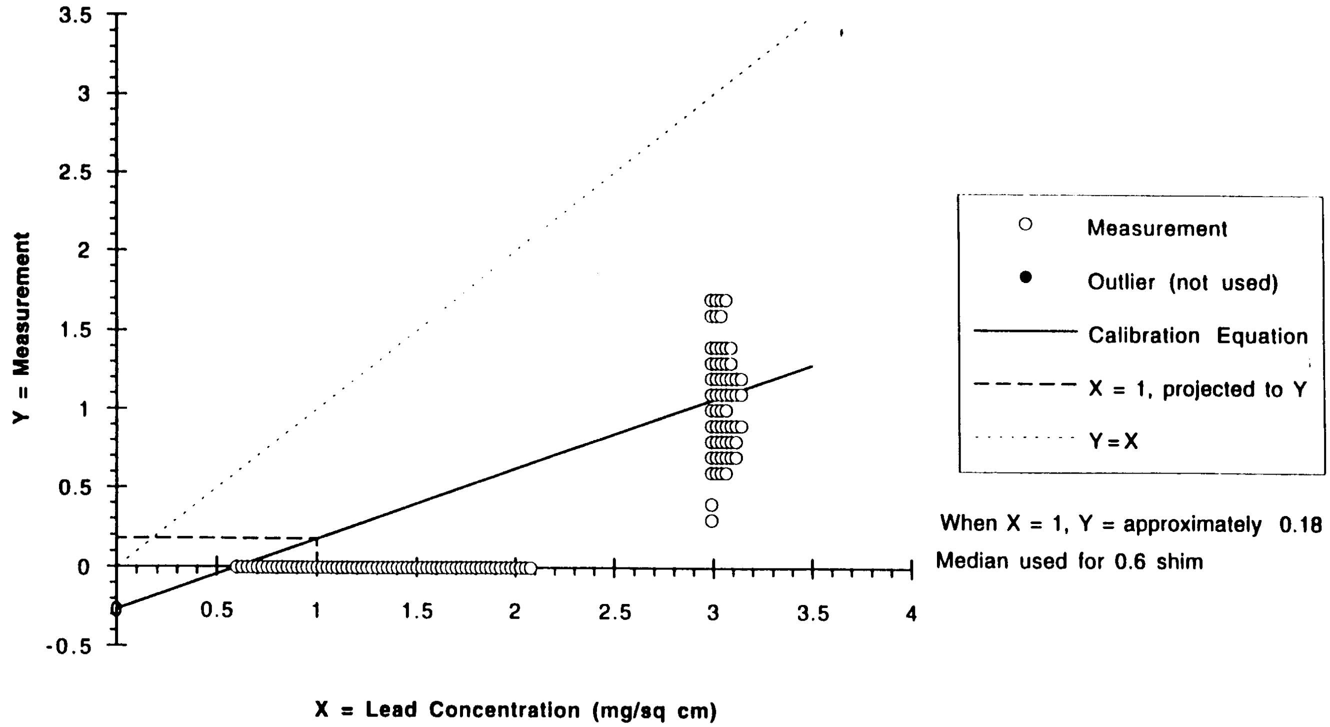
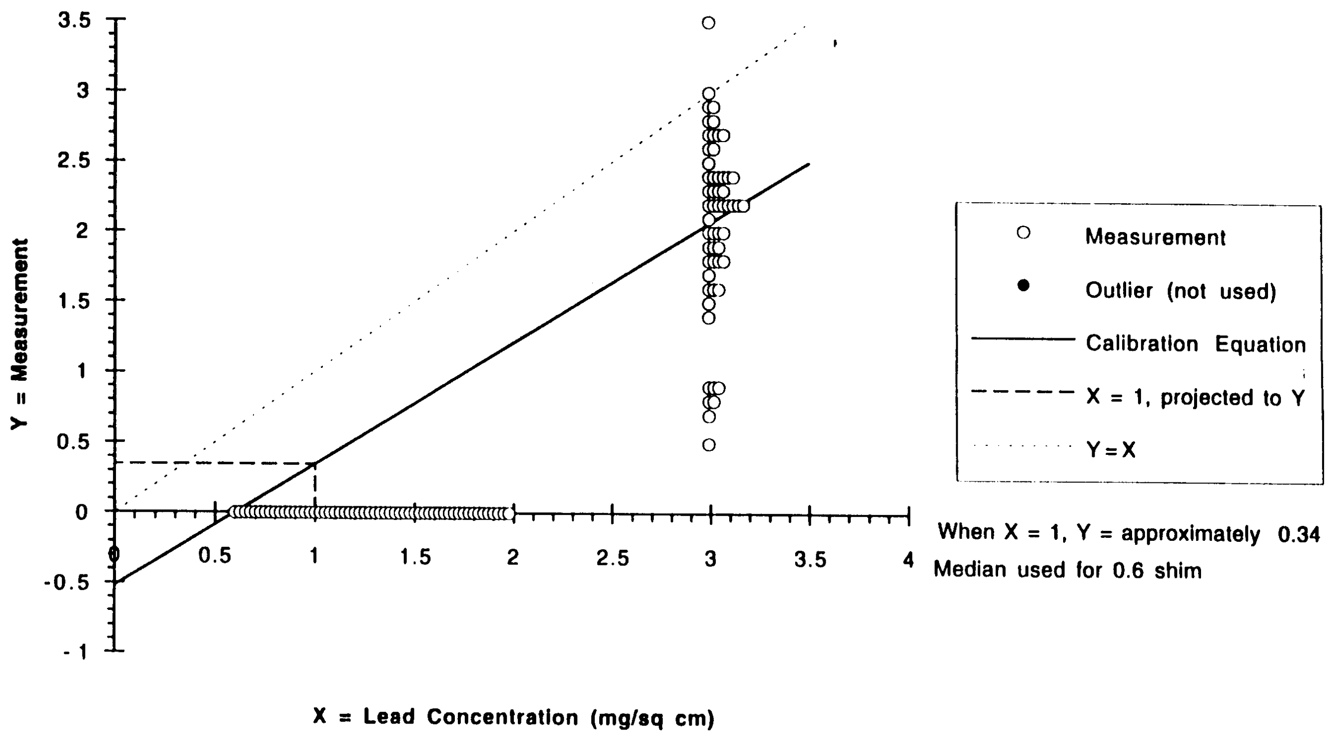


FIGURE D-28

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #38 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5753 + 1.1586*Measurement + 0.00115*(Days since 2/1/90)

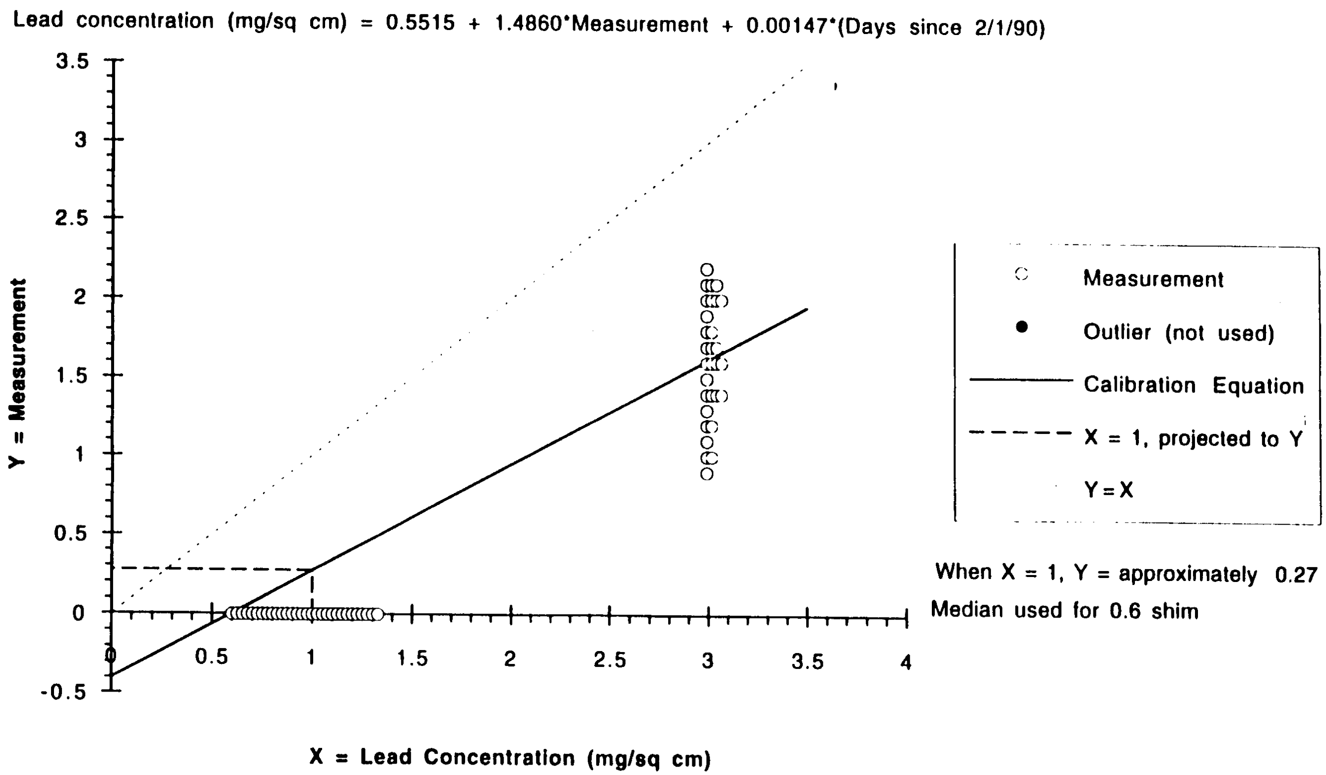


When X = 1, Y = approximately 0.34

Median used for 0.6 shim

FIGURE D-29

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #39 ON CONCRETE



When X = 1, Y = approximately 0.27
Median used for 0.6 shim

FIGURE D-30

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #41 ON CONCRETE

Lead concentration (mg/sq cm) = 0.5635 + 0.8673*Measurement + 0.00086*(Days since 2/1/90)

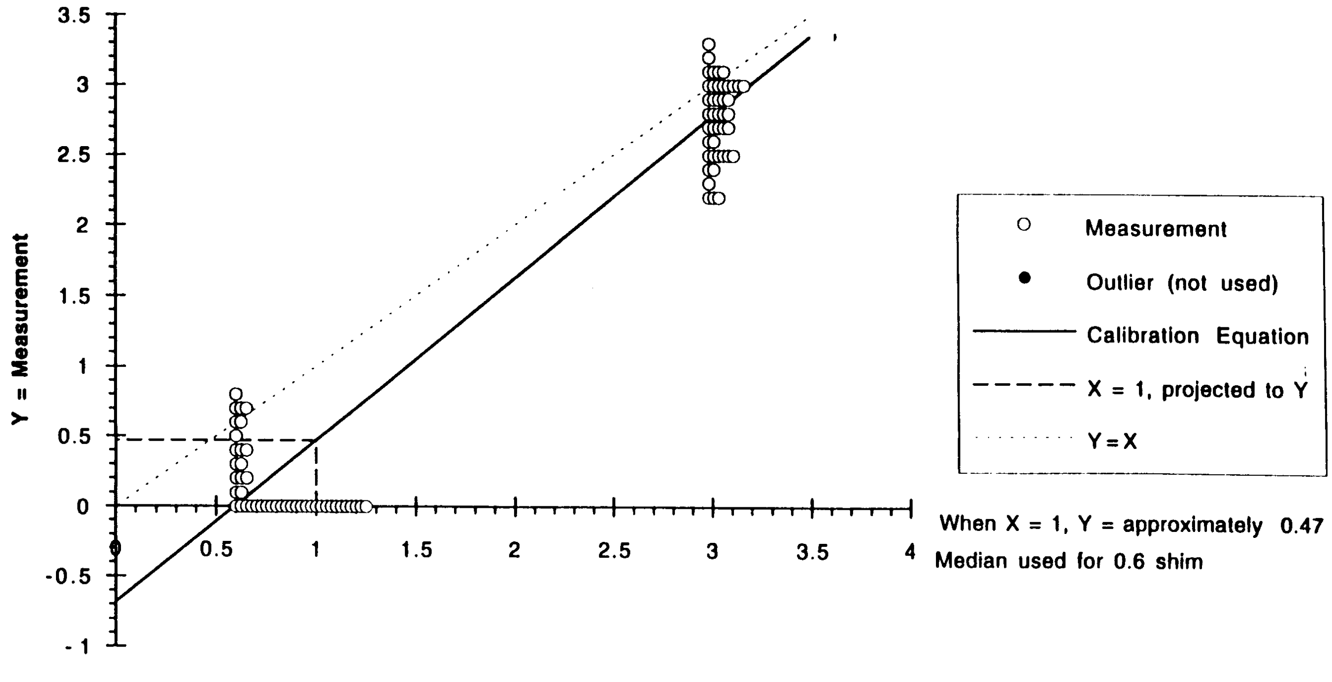


FIGURE D-31

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #32 ON DRYWALL

Lead concentration (mg/sq cm) = 0.5684 + 0.9389*Measurement + 0.00093*(Days since 2/1/90)

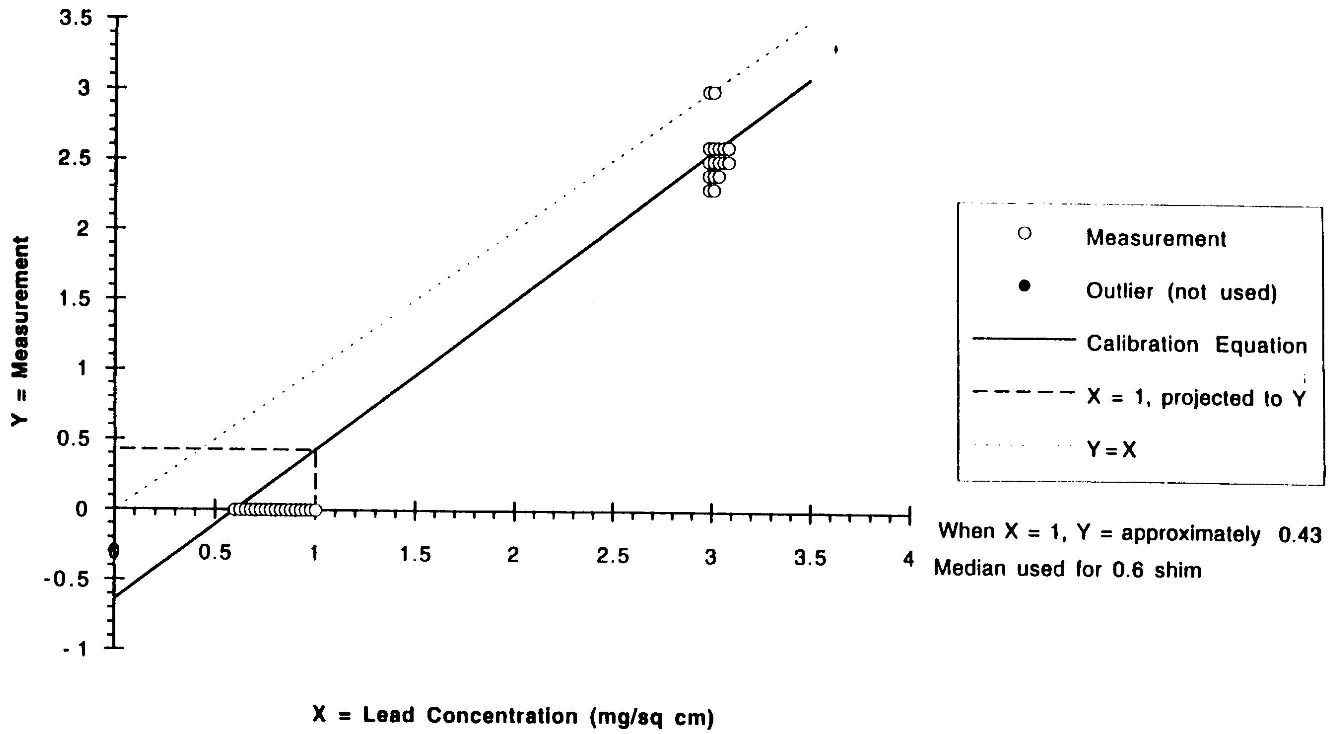


FIGURE D-32

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #34 ON DRYWALL

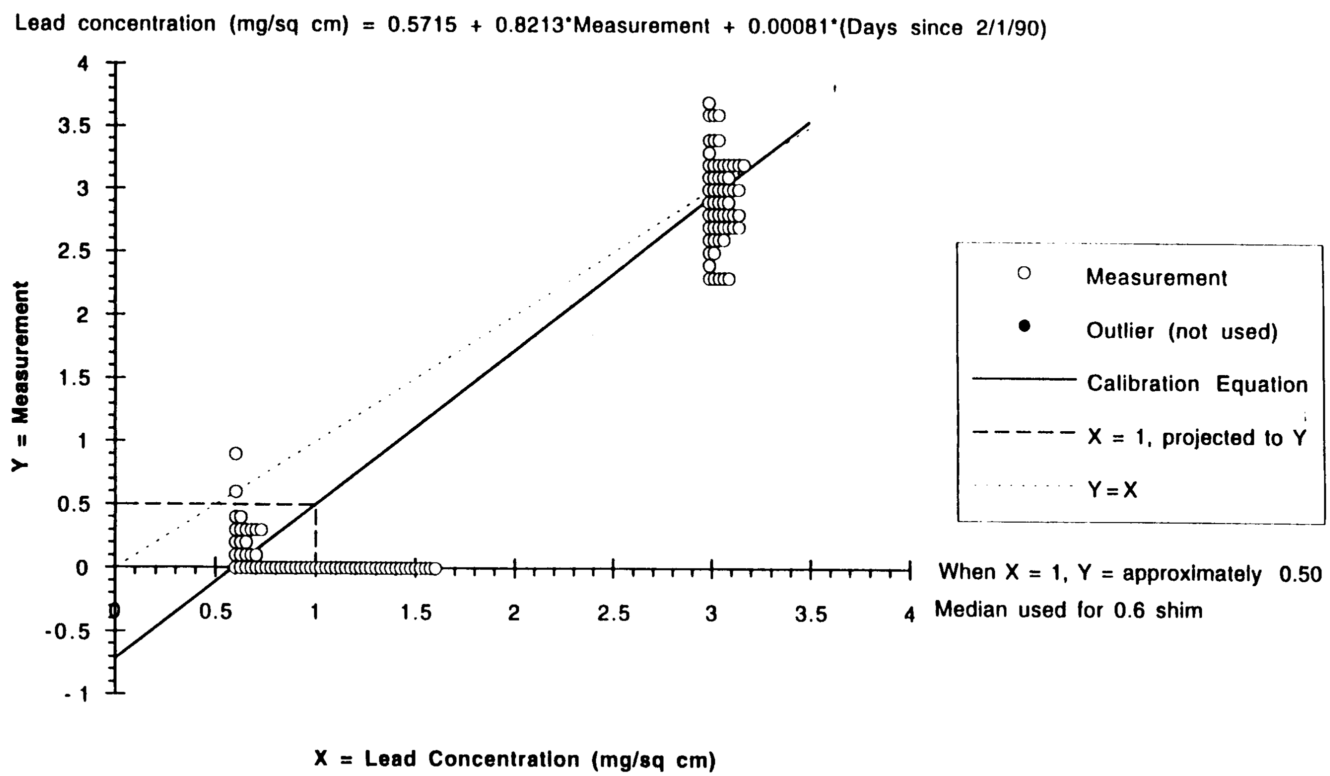


FIGURE D-33

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #35 ON DRYWALL

Lead concentration (mg/sq cm) = 0.5730 + 0.8014*Measurement + 0.00079*(Days since 2/1/90)

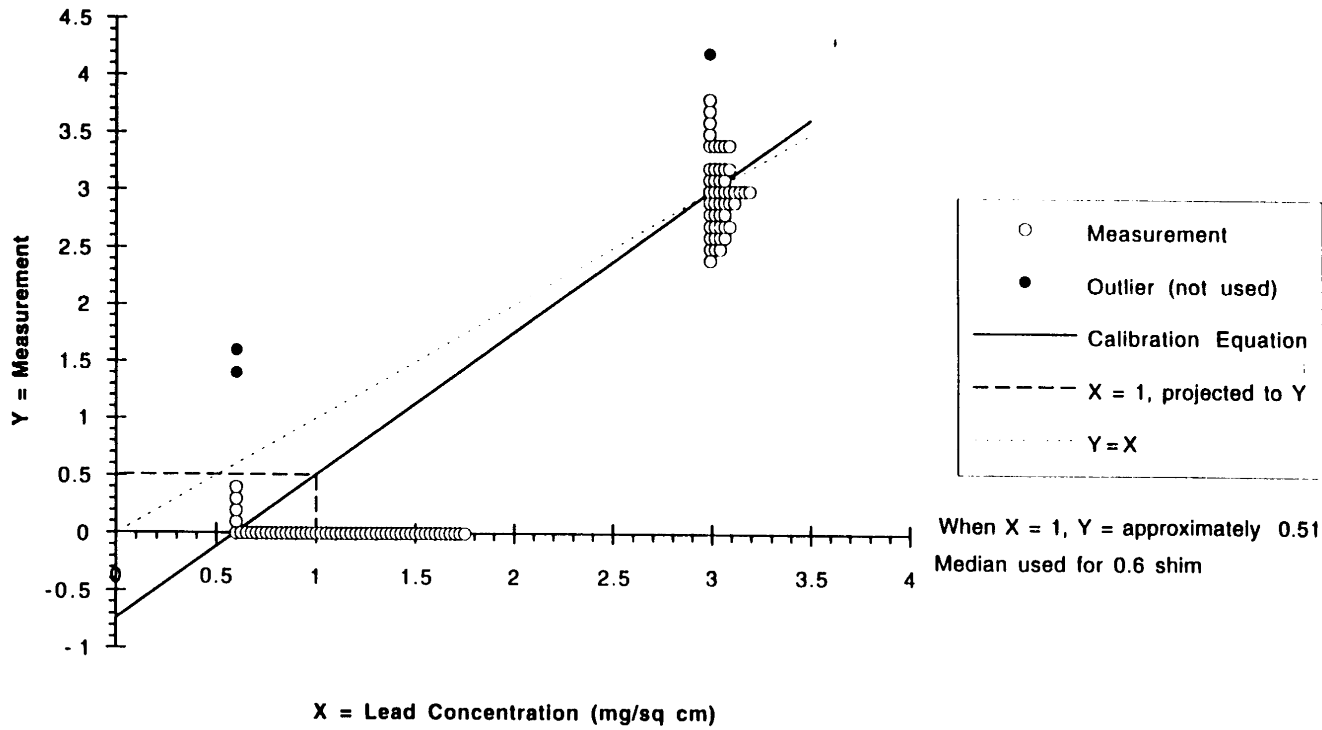
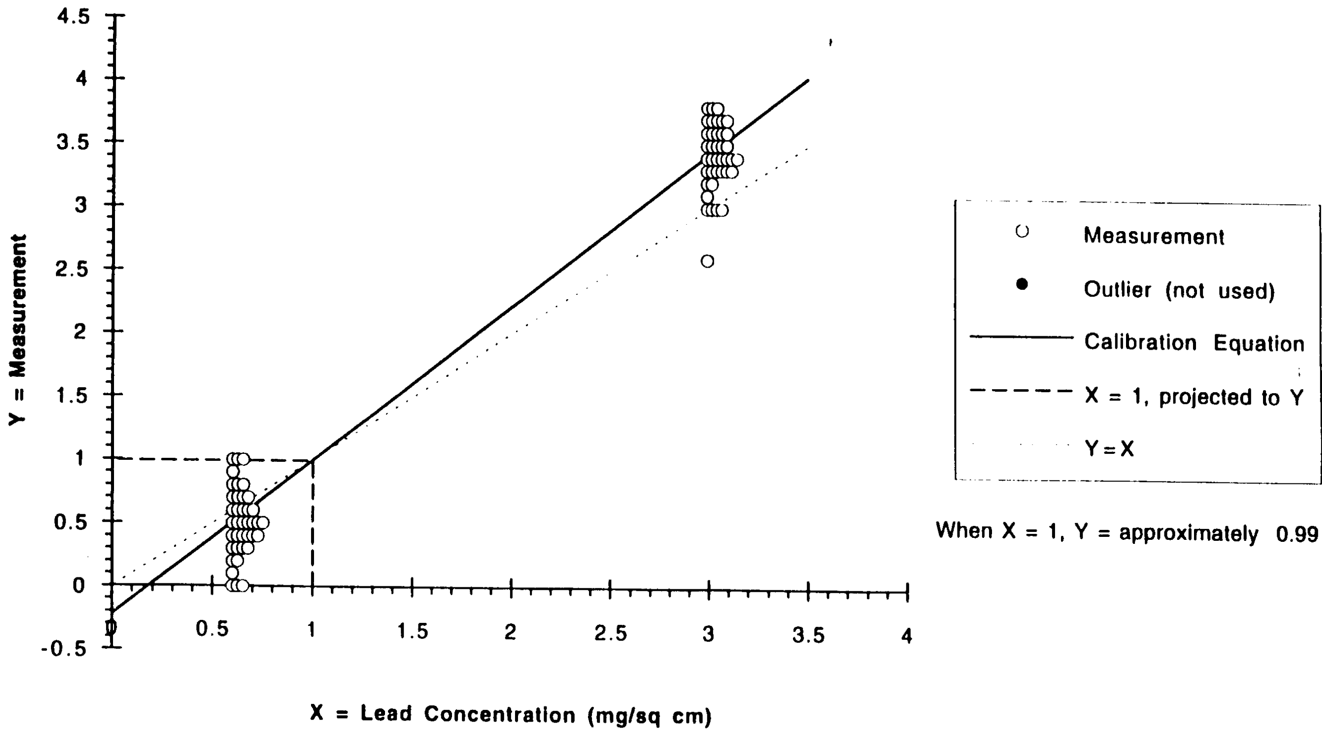


FIGURE D-34

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #36 ON DRYWALL

Lead concentration (mg/sq cm) = 0.1694 + 0.8220*Measurement + 0.00081*(Days since 2/1/90)

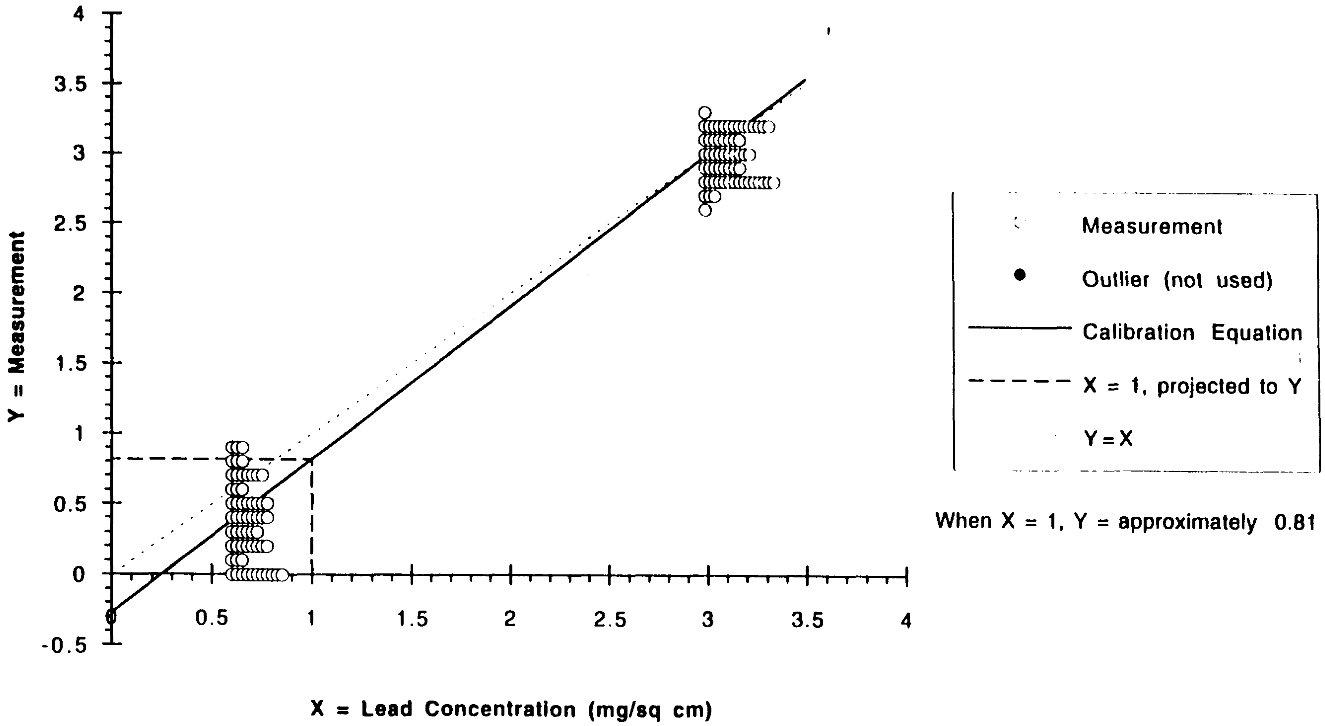


When X = 1, Y = approximately 0.99

FIGURE D-35

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #37 ON DRYWALL

Lead concentration (mg/sq cm) = 0.2311 + 0.9186*Measurement + 0.00091*(Days since 2/1/90)

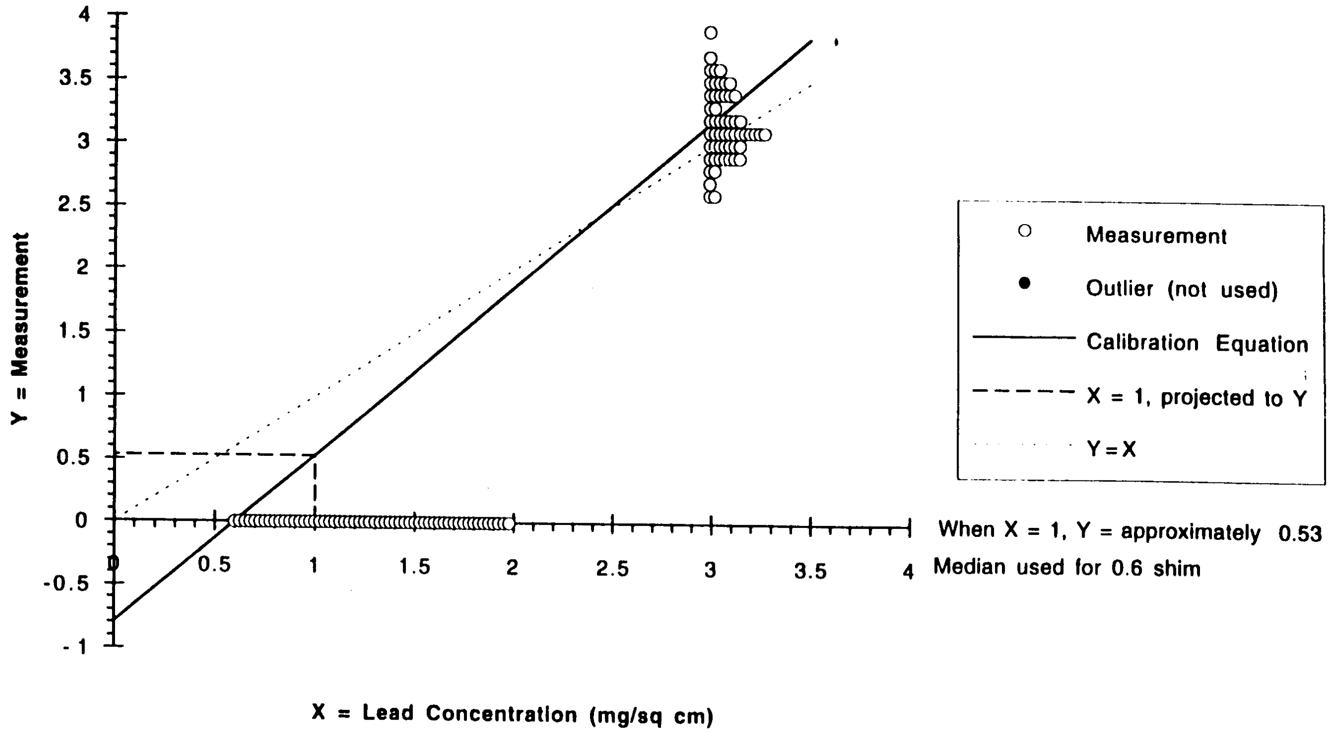


When X = 1, Y = approximately 0.81

FIGURE D-36

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #38 ON DRYWALL

Lead concentration (mg/sq cm) = 0.5839 + 0.7548*Measurement + 0.00075*(Days since 2/1/90)



When X = 1, Y = approximately 0.53
Median used for 0.6 shim

FIGURE D-37

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION
FOR XRF INSTRUMENT #39 ON DRYWALL

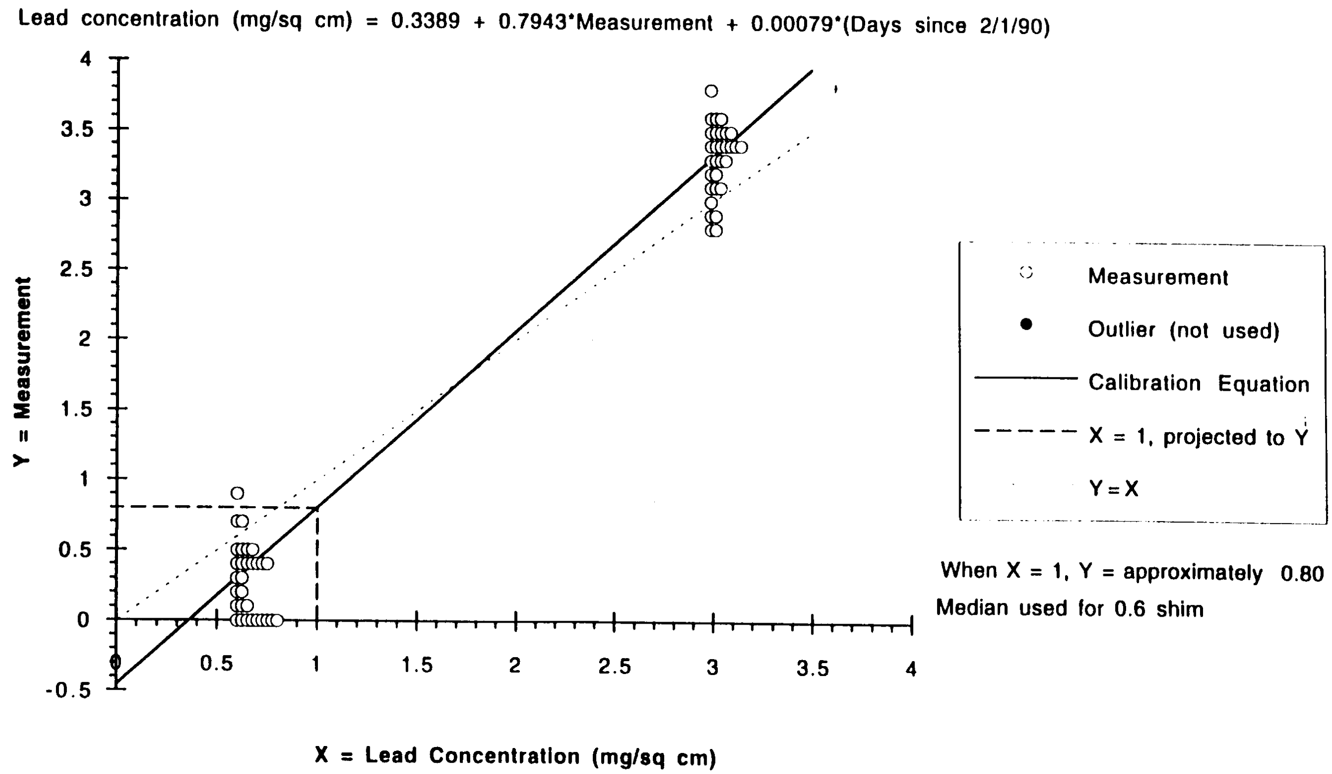
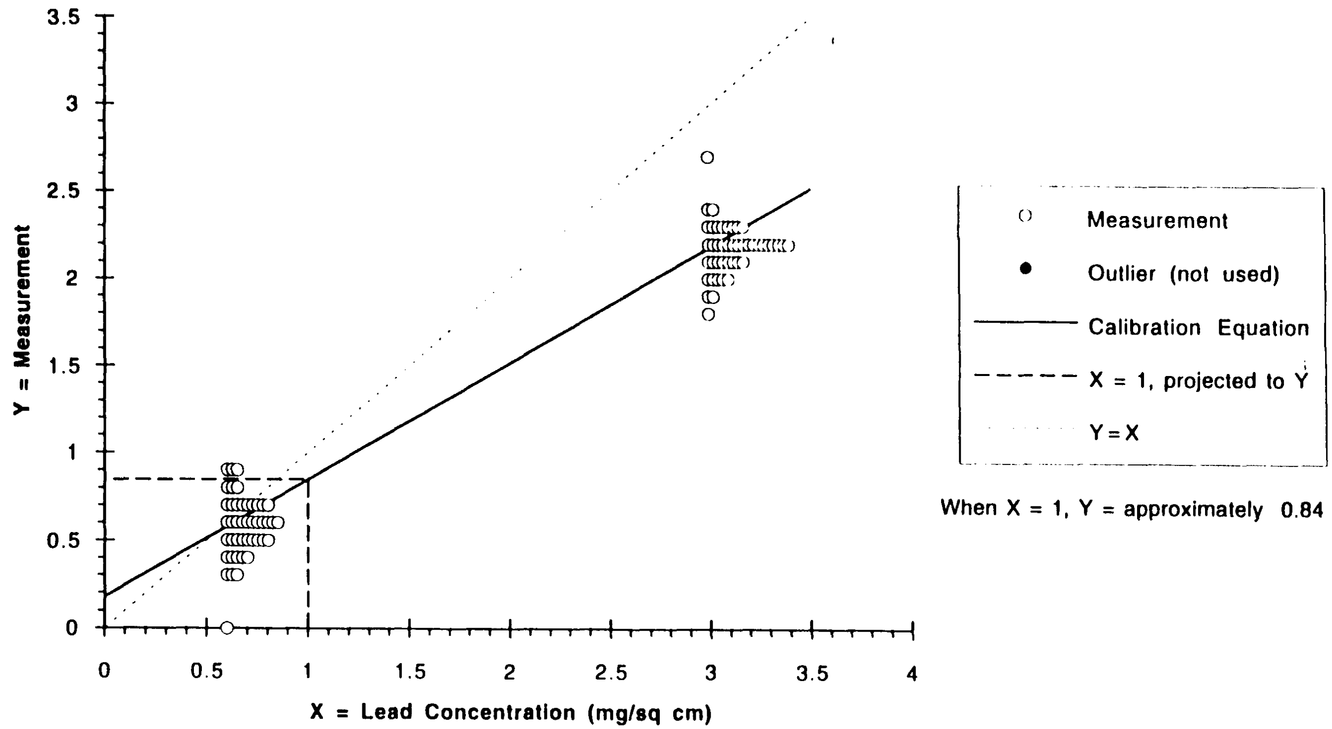


FIGURE D-38

VALIDATION MEASUREMENTS BY SHIM LEAD CONCENTRATION FOR XRF INSTRUMENT #41 ON DRYWALL

Lead concentration (mg/sq cm) = -0.3153 + 1.4959*Measurement + 0.00148*(Days since 2/1/90)



When X = 1, Y = approximately 0.84

FIGURE D-39

DISTRIBUTION OF SIMULATED RE CALIBRATED MEASUREMENTS FOR WOOD SUBSTRATES

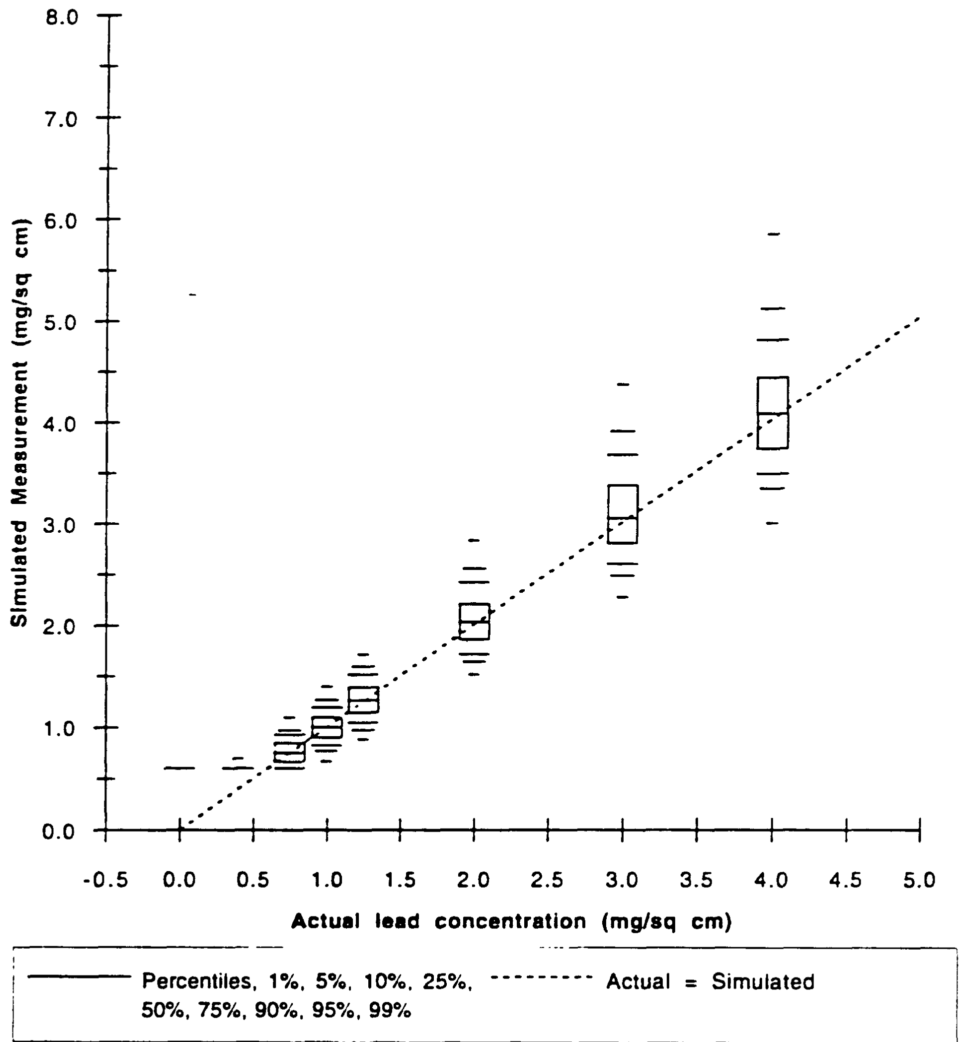


FIGURE D-40

DISTRIBUTION OF SIMULATED RECALIBRATED MEASUREMENTS FOR STEEL SUBSTRATES

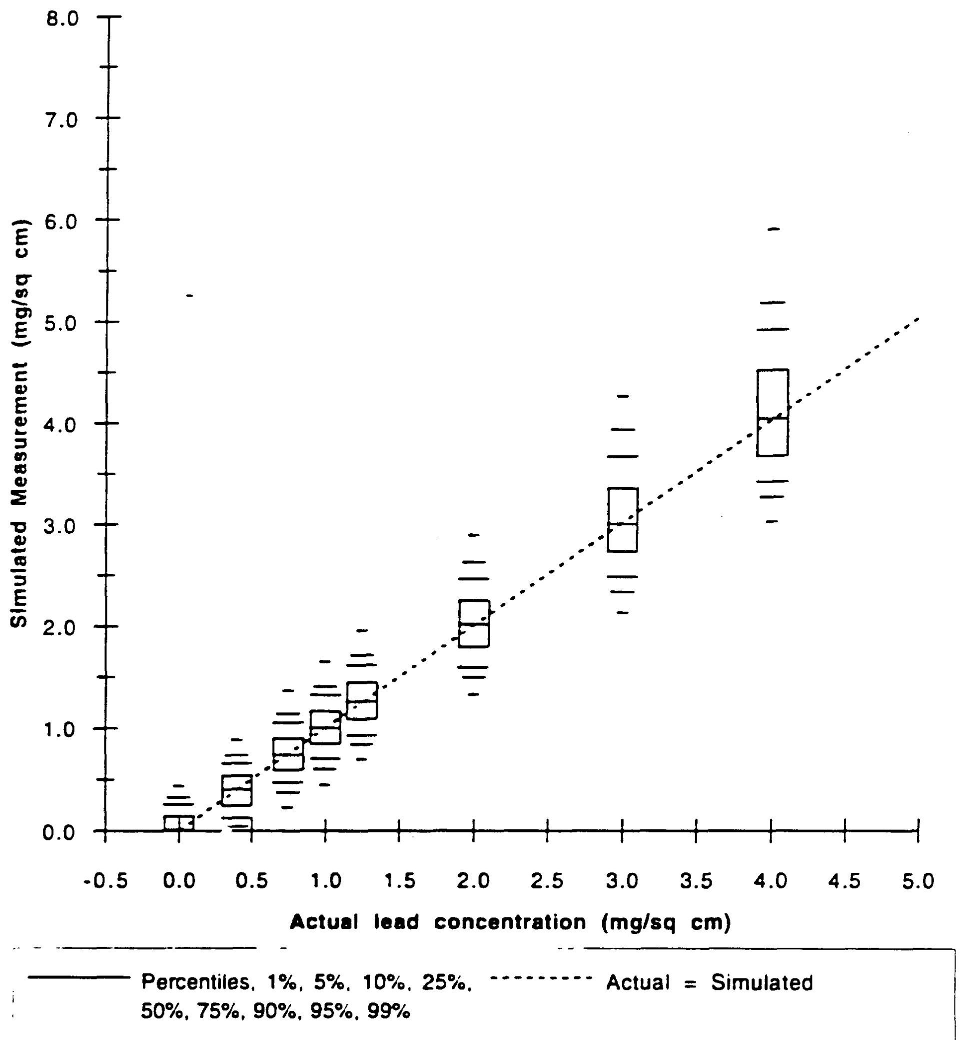


FIGURE D-41

DISTRIBUTION OF SIMULATED RECALIBRATED MEASUREMENTS FOR DRYWALL SUBSTRATES

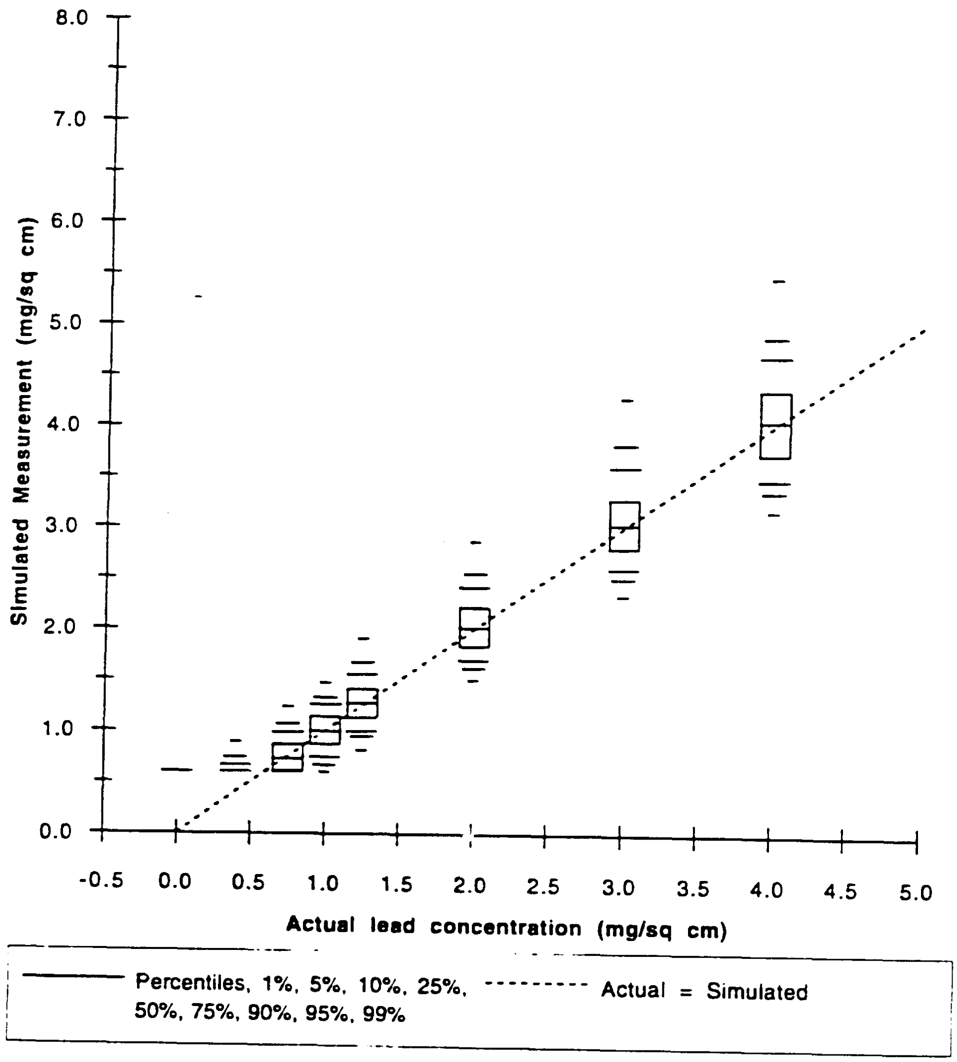


FIGURE D-42

DISTRIBUTION OF SIMULATED RECALIBRATED MEASUREMENTS FOR CONCRETE SUBSTRATES (Assumptions based on censored data)

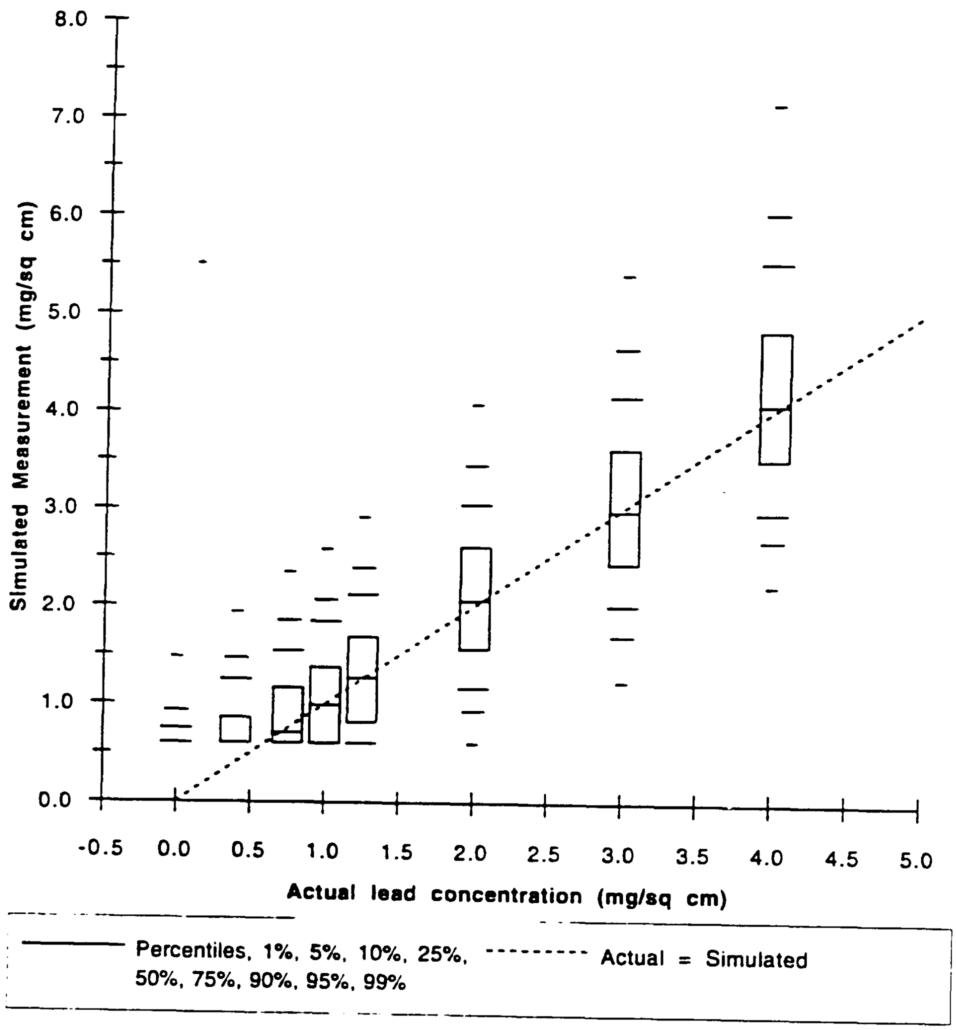


FIGURE D-43

DISTRIBUTION OF SIMULATED RECALIBRATED MEASUREMENTS FOR CONCRETE SUBSTRATES
(Assumptions ignore data from 0.6 shim, slope = 1)