NBSIR 77-1250

Analysis of Housing Data Collected in a Lead-Based Paint Survey in Pittsburgh, Pennsylvania Part I

Douglas R. Shier
and
William G. Hall

Applied Mathematics Division
Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

Report Prepared for

Center for Building Technology
Institute for Applied Technology, NBS

March 1977
Issued May 1977

Sponsored by the
Office of Policy Development and Research
Department of Housing and Urban Development
Washington, D.C. 20410

NBSIR 77-1250

ANALYSIS OF HOUSING DATA COLLECTED IN A LEAD-BASED PAINT SURVEY IN PITTSBURGH, PENNSYLVANIA PART I

Douglas R. Shier
and
William G. Hall

Applied Mathematics Division
Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

Report Prepared for

Center for Building Technology
Institute for Applied Technology, NBS

March 1977

Issued May 1977

Sponsored by the
Office of Policy Development and Research
Department of Housing and Urban Development
Washington, D.C. 20410

U.S. DEPARTMENT OF COMMERCE, Juanita M. Kreps, *Secretary*

Dr. Sidney Harman, *Under Secretary*

Jordan J. Baruch, *Assistant Secretary for Science and Technology*

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Acting Director*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

ACKNOWLEDGEMENT

The authors would like to acknowledge the constructive and helpful suggestions of Dr. Irwin H. Billick, Program Manager, Lead-Based Paint Poisoning Prevention Research, Department of Housing and Urban Development. Thanks are also expressed to Harvey W. Berger, Program Manager, Housing Technology Programs, National Bureau of Standards, for frequent and valuable discussions during the preparation of this report.

ANALYSIS OF HOUSING DATA COLLECTED IN A LEAD-BASED PAINT SURVEY IN PITTSBURGH, PENNSYLVANIA

Part I

Abstract

This report is a companion document to a previous report (NBSIR 76-1024) on blood lead levels of children tested during a lead-based paint survey in Pittsburgh, Pennsylvania. The emphasis in this report is on the methodology used and types of housing-related information collected by the survey. Through the use of portable x-ray fluorescence lead detectors, measurements were taken from a variety of surfaces within rooms of the dwelling unit as well as at locations exterior to the unit. Analyses of these x-ray fluorescence measurements established that older housing units exhibit considerably greater lead levels than newer housing units. In addition, wet rooms (kitchens and bathrooms) have higher levels than other (dry) rooms, rooms with a poor surface/substrate condition have higher levels than rooms with a good surface/substrate condition, and trim surfaces (e.g., doors, windows, baseboards) have higher levels than walls. Also, exterior surfaces show higher readings than functionally similar interior surfaces. While the present report concentrates on the housing aspect of the survey, subsequent processing of the Pittsburgh data is under way to determine possible relationships among blood lead levels, socioeconomic variables and housing-related characteristics.

Key Words: Children; data analysis; housing; lead paint; lead poisoning; surveys; x-ray fluorescence.

1. INTRODUCTION

This report describes the design, methodology, analysis and results of a large-scale housing survey conducted in Pittsburgh, Pennsylvania. The primary purpose of the survey was to assess (using a rather detailed series of measurements on the sampled dwelling units) the extent, magnitude and distribution of lead-bearing surfaces within the units. In addition, selected measurements of lead levels were recorded for exterior surfaces of the dwelling units.

The Department of Housing and Urban Development (HUD) has been mandated by Title III of PL 91-695, the Lead-Based Paint Poisoning Prevention Act of 1971, to determine the nature and extent of the lead paint poisoning hazard to children throughout the nation. HUD has also been directed to promote the development of lead detection devices, and has been charged with assessing the performance and usefulness of materials and systems for abatement of lead paint hazards.

The National Bureau of Standards (NBS) has provided technical assistance to HUD in each of the above areas of responsibility. In particular, the HUD sponsored housing survey detailed in this report was performed in response to the first of these responsibilities. In 1973, HUD requested that NBS develop a survey procedure to update and refine its previous estimates [1]* of the number of dwelling units containing lead-based paint. While housing surveys for quantitatively assessing lead sources have been conducted in other cities [2, 3, 4, 5, 6], this survey is believed to be the first major effort at

*Figures in brackets indicate references listed in Section 5 (pp. 64-65) collecting substantial data on lead in housing in a systematic way. In contrast to earlier surveys, the survey described here was a large-scale, statistically designed effort aimed at characterizing (using an extensive series of measurements) the lead levels in dwelling units comprising an entire city, and not simply a high-risk target area. Previous NBS reports [7, 8, 9] have dealt with other aspects of this survey, including preliminary testing of the survey methodology.

The primary objectives of the survey were:

1) To determine the distribution of representative measures of lead paint content with respect to dwelling type.
2) To determine the distribution of such measures, within dwelling type, according to surface type, surface condition and room type.
3) To examine, using statistical techniques, relationships between measures of lead levels in dwellings and blood lead levels of children resident in those dwellings*.
4) To demonstrate a feasible methodology for collecting housing data and to construct a data base for support of other specific tasks within the NBS and HUD Lead Paint Research Programs.

There were, as well, a number of secondary objectives of the survey. One was that of providing a "clean" data base of collected and derived information to support future research and guide policymakers in the general area of lead-associated hazards. Another secondary objective was to explore, to the maximum extent practicable, possible relationships among

*Preliminary results of such statistical analyses have been previously reported [8].

the data gathered and relevant information collected from other sources.

The housing survey methodology developed by NBS included procedures for sample construction, data collection and statistical analysis. The methodology was pre-tested in Frederick, Maryland during the spring of 1973. Procedures were then revised in light of this pre-test experience and used during a field test in Washington, D.C. later in 1973 [7]. Further refinements were incorporated into the methodology, and computer software for editing and analyzing the data was also developed.

The city selected to carry out the survey was to be typical of the larger northeastern cities in terms of age, climate and demographic characteristics. It was also desirable that the city have an appropriate agency with the technical and institutional capabilities for operating in both the housing and public health areas. The city eventually selected for this survey was Pittsburgh, Pennsylvania, and the Allegheny County Health Department (ACHD) was designated as the agency to perform the survey. Pittsburgh is centrally located within the dense population belt defined by the Great Lakes basin and the Northeast Corridor. It is a large older city, but one which is sufficiently compact so as not to require excessive travel times between inspected dwelling units. The ACHD is responsible for both the housing and public health functions exercised by the city. Furthermore, prior to the survey, in 1974, the ACHD had an ongoing lead screening program, an experienced and able staff, and an enthusiastic desire to carry out the survey. These characteristics, plus its relative proximity to Washington (which made collaboration and liaison much easier), served to make Pittsburgh and the ACHD ideal for the conduct of this survey.

Approximately 4000 dwelling units were visited by two-person teams of

ACHD housing inspectors during the course of this survey (summer 1974 - spring 1975). On the average, a single inspection team visited 5.6 units per day. Roughly 3300 of these units received a complete inspection. If, during inspection of a unit, children under seven years of age were identified as residents, a medical technician team subsequently visited the dwelling unit in order to collect additional socioeconomic data and a blood sample for lead analysis. Such data were gathered from a total of 456 children representing some 296 households out of 552 households with children under seven. The resulting data base (incorporating identification, housing data, blood lead results and socioeconomic data) contains more than 2.5 million characters of information.

The remainder of this report is divided into three sections and three appendices. Section 2 deals with design aspects of the survey and includes a discussion of alternative measures for characterizing lead levels in the dwelling unit as a whole (or in part). This section also discusses important housing characteristics expected to be related to the presence and amount of lead-based paint. Section 3 defines the data elements collected during the housing inspections and the subsequent blood collection effort. Quality control procedures relevant to the collection and analysis of data are also discussed here, as well as the construction of a "clean" data base using appropriate computer programs for data editing. Section 4 reports the results of analyses performed to determine the distribution of lead-based paint both within and among dwelling units. A preliminary discussion of observed relationships between children's blood lead levels and various measures of lead levels in their dwelling units has been reported [8], and a more thorough analysis is in preparation.

Appendix A describes the procedures used for completing the housing inspection form, while Appendix B specifies the actual formats for data elements comprising the data base. Appendix C provides a table of contents to the various analyses performed using the data collected in the housing survey.

The multiplicity of measurements taken and of additional data items collected makes it impractical to present the results of all analyses performed using the information in the data base. The salient results of such analyses are, however, presented and discussed in this report (which is Part I of a two-part document). Additional detailed tables (which are representative of the rather voluminous tabulations that have in fact been made) are presented separately in Part II of this document. The tables and figures used in Part I are for the most part derived from the tabulations of Part II, supplemented by information contained in another NBS report [7].

2. SURVEY DESIGN

This section begins by describing the strategy used to select those dwelling units included in the Pittsburgh survey. The next portion of this section discusses certain housing-related characteristics thought to be relevant to the lead levels of surfaces in dwelling units. Such considerations were important in devising the contents of the housing inspection form. In addition, the issue of how actually to aggregate the lead levels found in dwelling units is addressed in the final portion of this section.

2.1 Housing Selection Strategy

The strategy used to select housing units for this survey was based on the desire to obtain data characterizing the entire Pittsburgh area, and not simply certain areas where high lead poisoning incidence is suspected. For this reason, a representative random sample of the entire Pittsburgh area was needed. This was accomplished by drawing a simple random sample from the Polk [10] city directory for Pittsburgh. A sequence of computer-generated random numbers was used to perform the sample selection, as described in [9].

The advantage of a simple random sample is that it carefully balances the cost of performing the survey against the accuracy of results required by the survey. That is, this procedure can achieve quite accurate results, while still being much more efficient and much less costly than a complete enumeration of the target population. Since in a simple random sample, each unit has an equal likelihood of being selected, this method is not expected to produce serious statistical biases in one direction or the other.

In the present survey, a sample size of 4000 dwelling units was chosen. This sample size was large enough to ensure the required accuracy (within 5% at a 95% level of confidence) for each of a number of subpopulations defined by housing age and occupancy class. The sample size value is consistent with the calculations presented in [9], as applied to these subpopulations. In addition, the sample size was large enough to include sufficient information about various categories of characteristics considered important in the survey. A discussion of those characteristics follows.

2.2 Important Characteristics in Survey Design

The design of the survey form used in the housing inspections was based upon housing-related characteristics believed to be important in describing the potential lead-based paint hazard of a dwelling unit. Any such potential hazard clearly depends on the lead levels of various painted surfaces within the dwelling unit as well as the accessibility of paint on these surfaces to a child. On the one hand, the lead levels on surfaces in a dwelling unit are expected to depend on the unit's construction characteristics and the historical maintenance (or decoration) practices to which the unit has been subjected. On the other hand, the accessibility of paint on a surface is related to its location (e.g., room type and surface type) and the condition of the substrate (holes, cracks, etc.) or the paint film itself (cracked, peeling, etc.).

The types of construction characteristics relevant to surface lead levels deserve further amplification. In particular, the following construction characteristics were thought likely to be important in determining lead levels on surfaces.

(1) Age of Housing. The age of the units, measured in this survey using the date of construction, was expected to be a reliable predictor for the existence of lead-based paint. Indeed, the age of housing has often been cited as a major determinant for the presence and amount of lead-based paint [2, 3, 11, 12], and a careful analysis of this relationship is accordingly one of the aims of this report. There are several objective reasons for expecting such a relationship. First, the age of a unit is likely to be

a good observable indicator (at least within construction and material types) for the thickness of existing paint films [13]. A second reason is related to the fact that paints marketed over the years have had decreasing lead contents [13, 14]. This situation has resulted from changes in paint technology and increased regulation of paint usage. Prior to roughly 1940, better quality paints had high lead content and were commonly used in dwelling units; in the era from 1940 to 1960, evolving paint technology produced good paints containing less lead; since 1960, there has been an intensive effort to reduce further the lead content of paints.

(2) Construction Type and Materials. The type of construction used in the dwelling unit (e.g., frame or concrete) as well as the type of base material used certainly have an impact on the type of paint used, if any. Units having construction types or base materials that do not require painting, for either functional or aesthetic reasons, are likely to have been painted less frequently than other units, if at all.

(3) Occupancy Characteristics. Still another class of factors on which lead levels may depend is related to the past or present occupants. These variables include ownership status of the unit (owner-renter), income, family composition, educational level and related socioeconomic indicators. The survey conducted in Pittsburgh was only able to gather information about present (and not past) occupants. However, because Pittsburgh has a relatively stable population compared with cities of similar size

and demography, major differences are not expected to exist between the characteristics of past and present occupants.

In addition to the lead levels of surfaces, the accessibility of such surfaces to children is also an important factor in determining the potential risk of lead-based paint ingestion. For example, a surface with a peeling, powdering or flaking paint film or a deteriorating substrate can cause whatever paint is present to be more immediately available than a sound surface with a tightly adhered paint film. Similarly, a chewable surface such as a door edge is more accessible to a child than a ceiling. Accordingly, the housing inspection was designed to provide information about the location and condition of various surfaces within the dwelling unit. This location aspect is described both in terms of surface type (walls, ceilings, windows, doors, etc.) and room usage (bathroom, kitchen, etc.). The condition aspect is described by means of specific condition codes developed to characterize the substrate as well as the surface of interest. Since the condition of both substrate and surface will normally deteriorate with the passage of time, the age of a unit may also serve as a general indicator of condition.

2.3 Measurement of Lead Levels

The various characteristics just discussed are for the most part both observable and quantifiable. For example, the age categories "Pre 1940", "1940-1959" and "1960-1975" were used in the housing inspections. These categories can be determined with reasonable accuracy (by householder's knowledge or inspector's estimate), are compatible with classifications used in Census Bureau housing publications, and correspond well to the eras in which different paint technologies and usages were prevalent [13, 14]. Section 3 discusses in greater detail the actual contents of the housing inspection form and the specific observations or measurements taken in each dwelling unit.

One of the most important characteristics of a dwelling unit is a description of the lead levels of various surfaces in the unit. The measurement of lead levels in this study was made using a portable x-ray fluorescence (XRF) lead detector. Characteristics of such devices are described at greater length elsewhere [15, 16, 17]. Measurements, in units of milligrams of lead per square centimeter, were taken at various locations in each of several rooms within the dwelling as well as at various locations exterior to the dwelling.

While the interpretation of any individual measurement on a specific surface is straightforward, the aggregation of these individual measurements into a measure of lead level for the entire dwelling is a more difficult matter. There are several desirable properties that any such aggregated measure should possess, namely:

1) The measure should be easily derivable from characteristics that can be observed and quantified.

2) The measure should be conceptually simple and amenable to an unambiguous interpretation.

3) The measure should take into consideration existing standards which have been used by various political subdivisions in legislation and in enforcement programs.

4) The measure should be amenable to various stratifications. In practice this means that the measure should be defined by a continuous

variable (i.e., one assuming a continuum of values).

An aggregated measure which simply indicated that one dwelling unit (on an overall basis) contains more than some specified level of lead* while another does not, would not satisfy criterion 4 above. Such a binary (YES-NO) measure would represent a great oversimplification and could not be used to provide a gradation of dwelling units from "containing more lead" to "containing less lead." In Fact, a number of alternative continuous measures or statistics were used in describing the lead levels in a dwelling unit (or possibly a room within the dwelling): namely, the highest lead level present, the mean lead level, the median lead level, and the fraction of surfaces that exceed some specified level. The highest lead level present was considered important both because it is an estimate of the maximum possible lead level which can be associated with a unit and because it is consistent with the language typical of legislation and abatement policies [18, 19]. The other aggregated measures were used because they also are reasonable quantifications of the amount of lead present in a dwelling unit; in addition, they can accommodate the classification of surfaces according to a range of lead levels. Since the specification of a fixed level in the last of these measures is somewhat arbitrary, a sequence of fixed levels was actually used, thus providing a series of distinct (but related) aggregated measures.

3. DATA COLLECTION PROCEDURES

This section describes the various data elements gathered during the dwelling unit inspection and the subsequent collection of blood samples

---

*Lead levels exceeding 2.0 mg Pb/cm^2 are considered by some health authorities to pose a potential "hazard."

samples together with socioeconomic information. Quality control procedures were instituted during all phases of the survey to ensure the integrity of data collected. Such procedures are also briefly discussed in this section.

As previously mentioned, the housing data were collected for approximately 4000 dwelling units selected at random using a city directory for Pittsburgh. Two-person teams of ACHD housing inspectors collected these data during the course of inspecting both the interior and exterior of a unit; such an inspection typically required 30-40 minutes to complete. If it was determined that a child aged six years or less was resident in the unit, a subsequent medical technician team from the ACHD was sent to obtain a blood sample and to gather socioeconomic (as well as child-related) data.

The housing data collected consisted of observations (e.g., the number and condition of surfaces) made by the team leader, lead level measurements obtained on given surfaces using a portable XRF device, and (for a limited number of items) responses from the occupant of the dwelling. The team leader would record such information on the data collection form (DCF), which is shown in Figure 1; the second member of the housing inspection team assisted in operating the XRF device. All XRF readings were indicated as milligrams of lead per square centimeter (mg Pb/cm^2) and a single reading was obtained for each of a number of specified surfaces*. Detailed instructions for performing this data collection, together with

*Of course, there is no assurance that a single XRF measurement will fully characterize a given surface. Practical considerations, however, precluded repeated measurements on the same surface.

OMB NO. 014R 2800: APPROVAL EXPIRES JUNE 30, 1975

NBS-744
(REV. 7-74)

U.S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

SURVEY OF LEAD-BASED PAINT HAZARD IN DWELLING UNITS

SECTION I - IDENTIFICATION

ADDRESS

SERIAL NUMBER	CODE	TRACT	BLOCK	ZIP CODE	V	STREET NAME AND NUMBER
1 - 6	7 - 8	9 - 14	15 - 17	18 - 22	23	24 - 59
.	0 1					

CODE	XRF SERIAL	TEST BLOCK	ZERO READING	ZERO VERNIER	CALIBRATE READING	CALIBRATE VERNIER	PEAK READING	PEAK VERNIER	INSPECTORS		DATE MONTH	DATE DAY	
7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 59
0 2													

CODE	SEQ	TEST BLOCK	TYPE	XS	OC	YEAR	OWNER	MORT	PUBLIC	SUB	CHILD		
7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 59
0 3													

SECTION II - INTERIOR

	CODE	WALLS AND CEILING COND	READING WALL 1	READING WALL 2	READING WALL 3	READING WALL 4	READING CEILING	TRIM COND	WINDOW NUMBER	WINDOW READING	DOOR NUMBER	DOOR READING	BASE-BOARD	OTHER FLOOR	RADIA-TOR	CABINET	FIRE PLACE
	7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 47	48 - 50	51 - 53	54 - 56
LIVING	0 4																
DINING	0 5																
FAMILY	0 6																
KITCHEN	0 7																
BATH 1	0 8																
BATH 2	0 9																
BED 1	1 0																
BED 2	1 1																
BED 3	1 2																
STAIRWAY	1 3																
HALL	1 4																
BASEMENT	1 5																
	1 6																
	1 7																

SECTION III - EXTERIOR

	CODE	WALL COND	WALL RDG	PORCH COND	PORCH RDG	DOOR COND	DOOR RDG	WINDOW COND	WINDOW RDG	RAILING COND	RAILING RDG	FENCE COND	FENCE RDG	GARAGE COND	GARAGE RDG	EXCLUDED ROOMS BED	EXCLUDED ROOMS BATH	EXCLUDED ROOMS OTHER
	7 - 8	9 - 11	12 - 14	15 - 17	18 - 20	21 - 23	24 - 26	27 - 29	30 - 32	33 - 35	36 - 38	39 - 41	42 - 44	45 - 47	48 - 50	51 - 53	54 - 56	57 - 59
FRONT	1 8																	
REAR	1 9																	

I. Visitation (V)
Completed
1. Not dwelling unit
2. Refusal
3. No such address
4. Vacant
5. Under construction
6. Demolished

II. Type of Construction (Type)
1. Frame
2. Masonry
3. Concrete

III. Outside Surface of Building (XS)
1. Wood siding
2. Wood shingle
3. Asbestos shingle
4. Asphalt siding
5. Brick or stone
6. Stucco or cement block
7. Aluminum siding

IV. Occupancy (OC)
1. One-family detached
2. One-family attached
3. 2 to 4 families
4. 5 to 9 families
5. 10 to 19 families
6. 20 to 49 families
7. 50 or more families

V. Year Built (YEAR)
1. 1939 or earlier
2. 1940 to 1959
3. 1960 to present

VI. Owner - Renter (OWNER)
O - Owner
R - Renter

VII. Mortgage (MORT)
V - VA
F - FHA
C - Conventional
N - None
R - Renter

VIII. Public Housing (PUBLIC)
P - Public
N - Non-public

IX. Subsidized (SUB)
S - Subsidized
N - Not Subsidized

X. Children Resident (CHILD)
Number of resident children age 6 and under

XI. Material and Condition (COND)
(a) Base Material
P - Plaster
V - Vinyl
G - Gypsum board
B - Masonry
T - Tile
W - Wood
M - Metal
O - Other
(b) Base Condition
1. Good
2. Fair
3. Poor
(c) Surface
Q - Peeling paint
K - Tight paint
N - Not painted, good
X - Not painted, good

explicit definitions of the individual data elements, appear in Appendix A.

In general structure, the DCF consists of four sections: an identificat section, an interior section, an exterior section and a glossary section. Each of these sections has somewhat different characteristics and is describe briefly below.

(1) Section I (IDENTIFICATION) contains information that is associated with the dwelling unit as a whole. Also included here is information pertinent to the identity of the inspection team, the date of inspection and the particular XRF instrument used. More specifically, the dwelling unit identifiers include a unique serial number, census tract, ZIP code, and street address of the dwelling unit. All of these items appear in the first line of this section. On the second line appear the inspection team data: namely, the XRF serial number, XRF calibration parameters, inspector identification and date. Such inspection team data were used in internal management and quality control procedures.

The initial two fields of the third line are associated with XRF readings taken on a reference standard test block (of known lead content) prior to inspecting a unit. These readings were used by the inspection team to determine if an on-site recalibration of the XRF device was necessary. The remaining fields on this line are associated with the dwelling unit and its occupants; into each field is entered a code whose definition appears in the glossary section of the DCF. The fields TYPE, XS and OC (representing construction type, exterior building surface and occupancy class) are completed using the team leader's observations of the unit. Entries in the last six fields of this line are based on information supplied by the occupant.

(2) Section II (INTERIOR) contains one line (lines 04 through 15) for each room of the dwelling which was inspected; each such line follows the same format. All fields except those designated by COND (condition) or NUMBER contain XRF readings (in tenths of mg/cm^2) for the surface specified in the field heading. The field was left blank if the surface did not exist or was unpainted, and an "X" was entered if the surface was inaccessible to the inspection team. Each COND field is composed of a three-character code (from the glossary section) indicating the substrate material, substrate condition and surface condition for the specified aggregate of surfaces. The NUMBER fields contain, respectively, the number of windows and the number of doors present in the room. Finally, line 17 contains the number of times the family moved during the preceding six-year time period. Line 16 of the DCF was not used.

(3) Section III (EXTERIOR) follows the same general form as Section II except that the XRF readings and condition codes refer to specified exterior surfaces rather than interior surfaces. Moreover, for interior surfaces the COND field is used to characterize all similar surfaces within a room (e.g., all walls), while for exterior surfaces the COND field is associated with one particular surface. In addition, line 18 contains three fields that are used to record the number of rooms of the indicated types which were present in the dwelling unit but could not be inspected.

(4) Section IV (GLOSSARY) contains definitions for the eleven codes used in various lines of the data collection form.

In a separate follow-up collection of data (planned and conducted by the ACHD), blood samples and child-related data were obtained from 456 of 812 children identified as eligible for inclusion in the study. The child-related data gathered during this visit included physiological, sociological and environmental elements. Such data are subsequently referred to as PSE data. The data elements collected on the PSE form (see Figure 2) are defined clearly on the form itself. A subsequent report will separately discuss and analyze the PSE data in relation to the children's blood lead levels and the levels of lead detected in their dwelling units.

At the end of each day, a supervisor reviewed the data collection forms submitted by inspection teams in order to assure the consistency and accuracy of data obtained during dwelling unit inspections. Where information was missing or inconsistent, a follow-up data collection effort was then initiated to regather those data items which could be obtained by a telephone call or an exterior inspection of the unit. The fact that a follow-up of missing data was instituted also served to make the inspection teams more careful in collecting the data and in producing a complete housing form. Additional quality control procedures were used in the determination of children's blood lead levels, and these are described in another NBS report [8].

Further validation and correction of the data base occurred as information was keypunched and entered into a computer system. The

Figure 2
PSE Questionnaire Form

PSE Condition of Child

Serial Number ____________________ ACHD Personnel __________

Date ______________ Census Tract ________ Blood lead ________

1. Age of child (months) ________
2. Sibling rank ________
3. Height (inches) ________
4. Weight ________
5. Sex ________
6. Race ________
7. Time resident in unit (months) ________
8. Time resident in census tract (months) ________
9. Percent of waking hours child spends in house ----------Less than 25% _____
   25% - 50% _____
   50% - 75% _____
   75% or more _____
10. Does the child spend more than 25% of waking hours at a friend's or relatives?
    Yes ____ No ____ If "yes" give census tract ________
11. Does child attend pre-school or day care?
    Yes ____ No ____ If "yes", Days/Week ____ Hours/Day ____
    Census Tract ______
12. Two favorite outdoor playing areas

Surface	Spot #1	Spot #2	
Grass	________	________	
Dirt	________	________	Census tract ____
Pavement	________	________	

13. Favorite indoor playing area ______________ ______________
14. Evidence of pica ------ no ______
    possible ______
    confirmed ______

---

15. Number of siblings ----- under 6 ______ 7 - 11 ______ 12 or over ______
16. Family income ---------- 0 - 3,999 ______
    4,000 - 6,999 ______
    7,000 - 10,999 ______
    over 11,000 ______
17. Work status ----- Mother works ________ Father works ______ Both work _____
18. If both parents work, child is cared for by:
    Siblings ________
    Adult relative ________
    Adult non-relative ________
    Day care ________
19. Parental education level

	Mother	Father
Highest level achieved	_____	_____
cares not to answer	_____	_____

20. Origin of parents:

	Mother	Father
Allegheny County	_____	_____
Outside Allegheny	_____	_____

21. Comments & observations (concerning child's development, behavior, supervision, etc)
    ____
    ____
    ____
    ____
22. Number of parents in household ________
23. Number of adults in household ________

general approach was to use multi-stage validation and correction procedures The key procedure was a computer program that performed consistency and validity checks on the data submitted. Entries from the DCF or the PSE forms which were inconsistent or invalid were defaulted to a "most likely" or "unknown" category, and the defaulted value was then entered into the data base. Each field with a defaulted value generated an appropriate error message; such error messages were used in conjunction with the hard copy files for reconciliation and for constructing required corrections. The final output of this process was a "clean" data base that could be used for subsequent analyses.

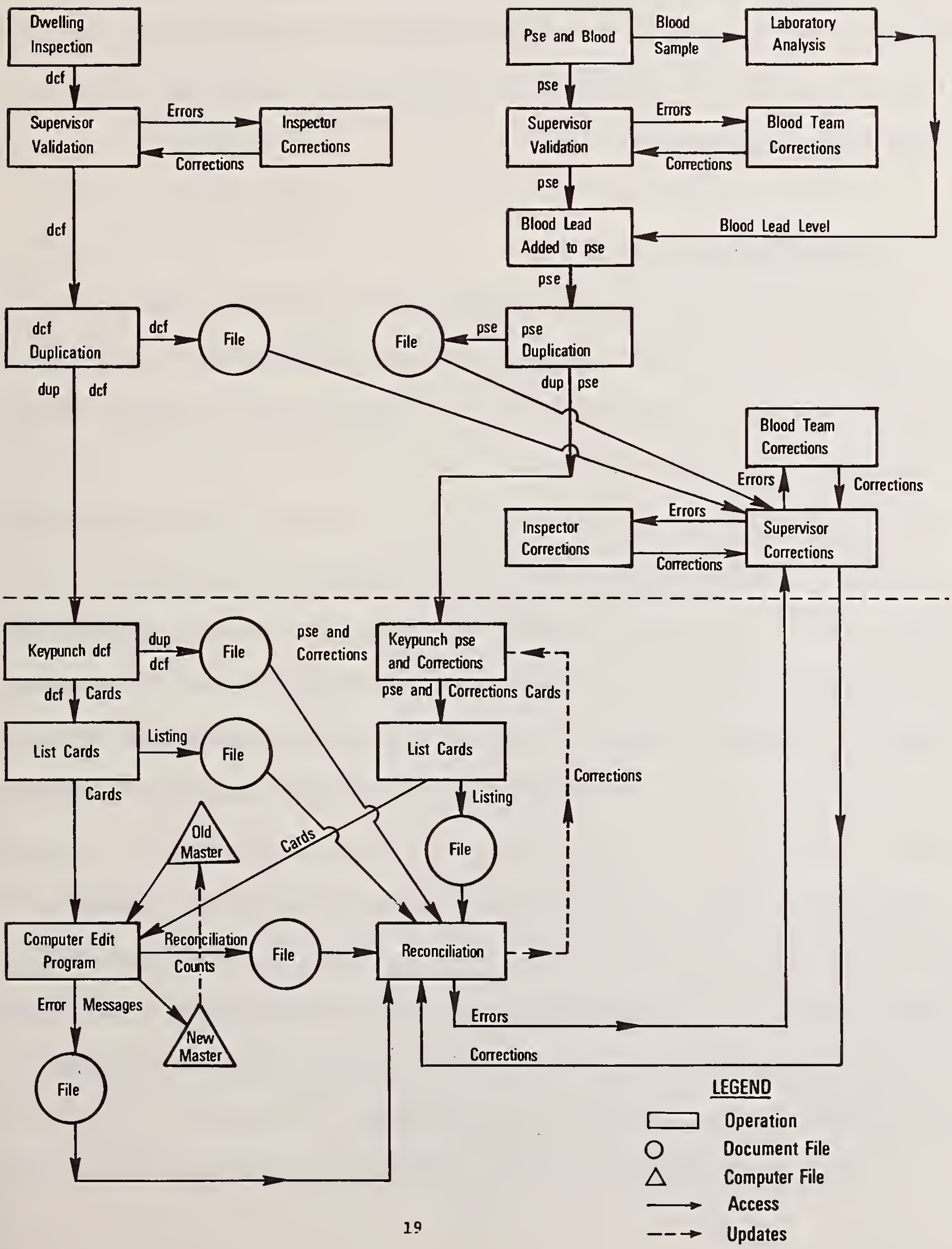
The general flow of information is shown in Figure 3 together with the appropriate quality control measures taken. In this figure, a rectangle is used to indicate a physical operation; also, a circle indicates a document file, while a triangle indicates a computer file. Annotated flow lines are used to denote flow of the specified information. An unlabelled flow line denotes access, while a dashed flow line indicates a cycle change in the data base. Operations above the dashed line were performed by the ACHD, and those below were performed by NBS.

4. ANALYSIS OF HOUSING DATA

In this section, attention is focused exclusively on characterizing (in terms of measured lead levels) the dwelling units of the sampled area. In particular, the relations of various housing characteristics (such as age of structure) to the lead levels occurring in dwelling units are investigated here. Furthermore, the distribution of lead-bearing surfaces within a typical dwelling unit is also examined.

Figure 3

Flow of Information



Subsequent processing of the current data base will concentrate on analyzing data items associated with the children for whom a blood sample was taken. In such analyses, relationships among the blood lead determinations, socioeconomic variables and characteristics of the dwelling units will be studied. A preliminary summary of relationships between blood lead levels and certain housing/demographic characteristics has already been reported [8].

4.1 Housing-Related Variables

Before proceeding to various analyses of the housing data, it will be useful to identify and describe certain housing-related variables whose influence is to be studied. Two types of general housing characteristics are most relevant here:

(1) Occupancy Class. A number of distinct (but not necessarily mutually exclusive) occupancy classes have been considered in the analyses that follow. These classes were specified in order to determine whether lead distribution patterns are indeed different among the various occupancy classes. Any such differences would be important to ascertain, since governmental authority and responsibility can vary according to occupancy class. Occupancy class can be defined by several descriptors: occupancy status (owner-renter), mortgage type (VA-FHA-conventional), number of units within the building, and extent of governmental assistance (e.g., public or subsidized housing). A large number of potential combinations are thus possible, but for the analyses conducted here the following ten categories were considered most relevant:

SINGLE FAMILY - OWNER OCCUPIED - FHA

SINGLE FAMILY - OWNER OCCUPIED

SINGLE FAMILY - RENTER OCCUPIED

SINGLE FAMILY

MULTIFAMILY - 2 to 4 UNITS/BUILDING

MULTIFAMILY - 5 or MORE UNITS/BUILDING

MULTIFAMILY (2 or MORE UNITS/BUILDING)

PUBLIC HOUSING - 2 to 4 UNITS/BUILDING

PUBLIC HOUSING - 5 or MORE UNITS/BUILDING

PUBLIC HOUSING

(2) Age of Unit. The age categories used here reflect distinguishable eras of paint technology, can be accurately estimated, and are consistent with age classifications used in Census Bureau housing publications. The three dwelling unit age categories are: PRE 1940, 1940-1959 and 1960-1975.

In addition to the above characteristics (which distinguish one dwelling unit from another), certain other characteristics are pertinent in describing the distribution of lead within a given dwelling unit. Four variables appear reasonably important in specifying collections of potential lead-bearing surfaces within the unit:

(1) Location. It is useful to distinguish those surfaces occurring interior to the dwelling unit from those occurring exterior to the unit. This distinction is considered important because of differing decoration/maintenance practices, differences in types of paint applied, and differing degrees of exposure to various environmental sources of lead (e.g., industrial or automotive emissions). The two general categories of

location considered here are denoted by INTERIOR and EXTERIOR.

(2) Room Type. This characteristic has been included since a pre-test [7] for the present housing survey indicated different lead level distributions according to room type. Moreover, paint usage and decoration practices are expected to differ by room type. It has been useful therefore to distinguish two general categories of rooms. WET rooms include kitchens and bathrooms, while DRY rooms include all rooms other than kitchens and bathrooms (e.g., bedrooms, living rooms, dining rooms).

(3) Surface Type. The pre-test also indicated that differences in lead levels exist with regard to surface type. More specifically, the different surface types analyzed here are WALLS, DOORS, and WINDOWS. In addition, the designation TRIM is used to signify the collection of doors, windows, baseboards, radiators, cabinets and fireplaces; this aggregate collection of surfaces appears meaningful, inasmuch as similar decorating practices and paint types tend to be used for all such surfaces. It is likely that surface type (as defined here) also plays an important role in describing the degree of accessibility to available paint.

(4) Surface Condition. The condition of a surface also provides a means of indicating accessibility. The categories used here are PEELING, BAD, EITHER and BOTH. The PEELING designation means that the paint film surface was peeling, flaking, blistering or powdering. BAD denotes a surface which has a bad substrate regardless of the condition of the

surface; this category includes all imperfections except hairline cracks, nail holes and the like. EITHER denotes surfaces that are either PEELING or BAD or both, while BOTH indicates surfaces that are both PEELING and BAD.

The above six variables are useful in specifying important characteristics of the dwelling unit. In conjunction with these variables, two issues arise in choosing an aggregated measure to quantify the amount of lead detected. First, it is important to specify the particular aggregate of surfaces that is under consideration. For example, one may wish to consider the collection of all surfaces in a dwelling unit, all surfaces in a room, or all surfaces of a specified surface type (e.g., all walls within the unit). Second, once the collection of surfaces (the level of aggregation) has been specified, it is then necessary to decide which particular aggregated measure is to be used. For example, it is possible to calculate either the mean, median or highest XRF reading for the specified collection of surfaces. In the analyses that follow, the influence of the six housing-related variables will be assessed using appropriate aggregated measures defined on various collections of surfaces.

4.2 Characteristics of the Sample

It is important to emphasize that any inferences drawn from the analyses in this report are based on a particular sample of dwelling units from the city of Pittsburgh (census tracts 0101 through 3204) together with the Mt. Oliver area (census tracts 4811 through 4813). It is important, then, to indicate how closely the actual sample of dwelling units obtained conforms to the underlying population of all dwelling units in the Pittsburgh/Mt. Oliver area. In particular, we will consider the closeness of representation of the collection of inspected dwelling units, relative to the entire sampled area. In this connection, an "inspected dwelling unit" is considered to be one in which at least one XRF reading was obtained. Excluded, therefore, are those dwelling units which did not exist (e.g., because of an error in the city directory or demolition of the unit), or units to which entry could not be obtained for inspection purposes.

There are several reasons why the 3342 inspected dwelling units may not be totally representative of the Pittsburgh (and Mt. Oliver) area. First, any random sample is subject to chance fluctuations (or "random errors") that are completely beyond the control of the investigator. Second, the fact that certain dwelling units could not be inspected is a potential source of statistical bias, if the selected (but uninspected) units differ markedly from selected (and inspected) units. Third, there may be actual methodological biases present in the selection of units for inspection.

An indication of how well the actual sample reflects the underlying population of dwelling units can be obtained by considering the relative occurrences of census tracts in the sample. In particular, the observed frequencies for census tracts in the sample can be compared with appropriate figures* obtained in the 1970 Census of Population and Housing [20]. Since the present study comprises some 190 individual

*These figures derive from Table H-1 (all housing units) in [20].

census tracts, it is useful to define a smaller number of census tract groups. Group 1 consists of all census tracts with leading two digits 01 or 02, group 2 consists of all census tracts with leading two digits 03 or 04, and so forth. In this way, 16 census groups are formed for the city of Pittsburgh, and a seventeenth group is defined for the Mt. Oliver area. Figure 4 presents a comparison of (1) the observed relative frequencies for census tract groups in the Pittsburgh study, and (2) the relative frequencies expected on the basis of the 1970 Census Bureau data (100% sample). With the exception of census tract groups 1, 4 and 11, the two "profiles" appear in close agreement: in all three exceptional cases, the observed relative frequencies underestimate the 1970 census values. In light of the above-mentioned potential sources of error inherent in the sampling process, the inspected units appear to be generally representative of the underlying population. Moreover, it should also be noted that 1970 census information (which itself may be affected by various sources of error in collection), is here being compared with a sample obtained in 1974-1975; that is, the underlying housing stock may have undergone change (by census tract) since 1970. When all these possible sources of error have been considered, the general agreement between the two profiles in Figure 4 provides confirmation of the survey's validity.

In addition, Table 1 permits a comparison between the sampled dwelling units and the underlying housing population according to the age of the structure. In this case, the percentages indicated for the 1970 census are estimates obtained from a 20% (and not a 100%) sample and are thus themselves subject to random variation. When this addi-

Figure 4

Percent of Sample by Census Tract Group

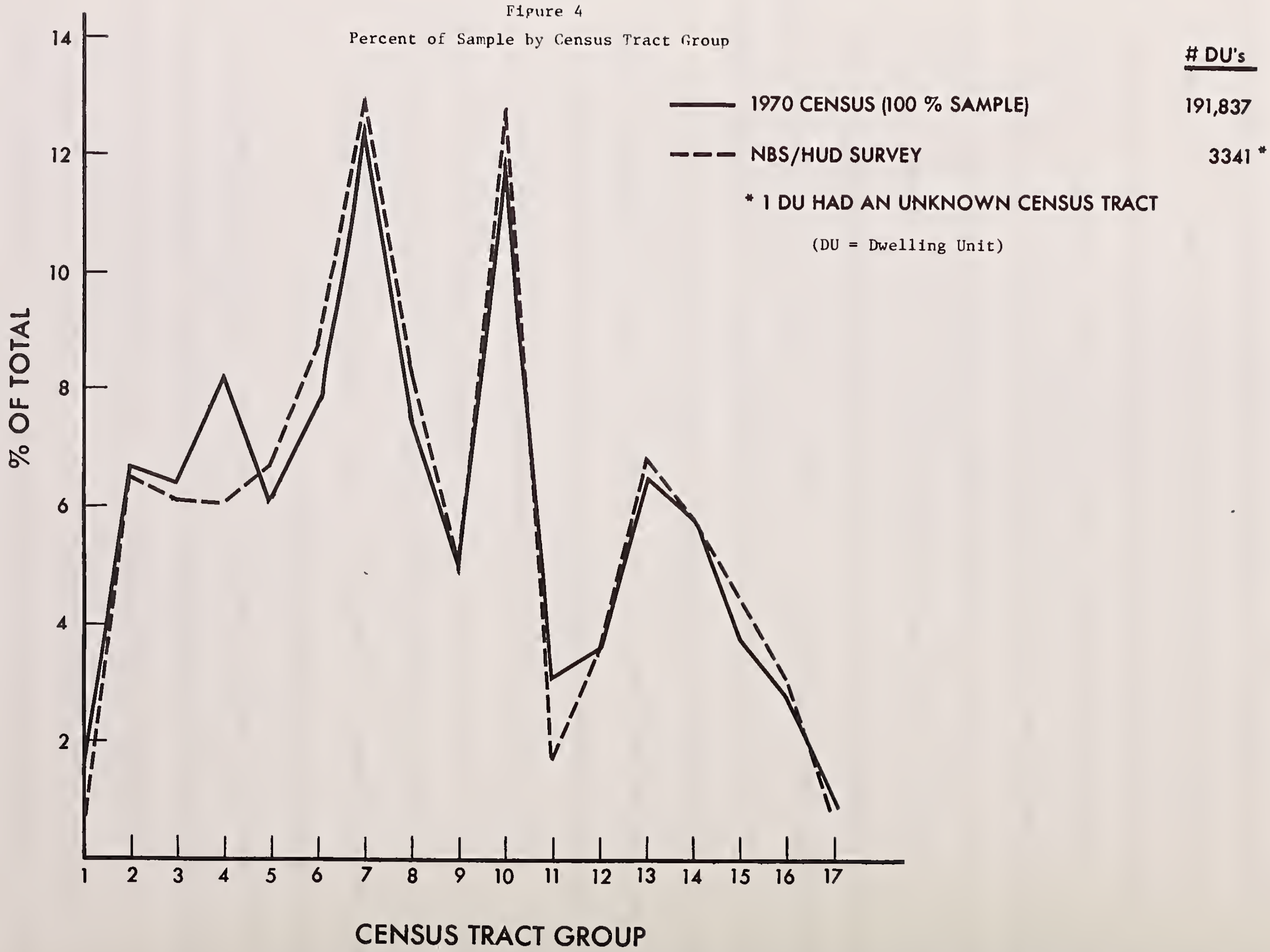


TABLE 1

Percentage of Dwelling Units by Age Category

AGE CATEGORY	NBS/HUD SURVEY	1970 CENSUS INFORMATION*
Pre 1940	76.6	74.5
1940-1959	17.3	18.3
1960-1975	6.0	7.2

* Based on Table H-2 (20% Sample), 1970 Census of Population and Housing

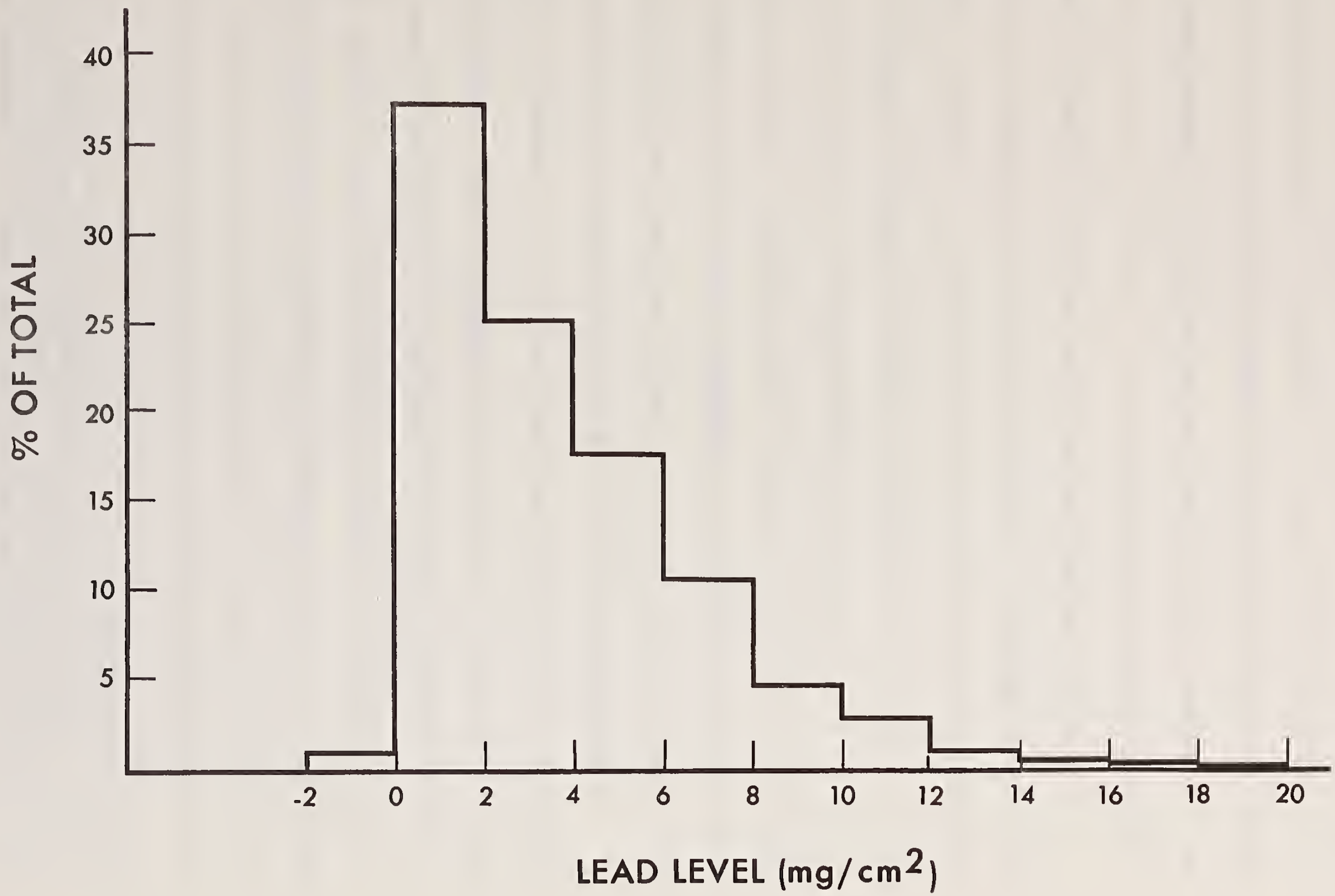
tional source of variation is added to the previously mentioned sources, the closeness of the percentages for the various age categories also tends to confirm the survey's validity.

4.3 Housing Survey Results

The lead levels measured on specified aggregates of surfaces can be characterized by a single quantity (such as the mean or median level) or by several such quantities. However, before computing and interpreting these statistical quantities, it is first useful to study the statistical distributions associated with the lead levels. Indeed, the actual shape of the underlying distributions will be important in deciding on the appropriateness of certain characterizing quantities (e.g., mean, median) and the validity of certain standard statistical tests. For such reasons, and in order to understand better the observed variation of lead levels, a preliminary study has been made of the statistical distributions of XRF readings. This information can be expressed by means of a frequency distribution (or histogram), such as that illustrated in Figure 5. For the purposes of this histogram, each dwelling unit in the pre 1940 age category has been characterized by the mean XRF reading taken from all measured interior trim in the dwelling unit. The proportion of mean readings falling within each of the specified lead level intervals has then been plotted on the vertical axis, yielding the histogram displayed in Figure 5. Histograms have also been graphed according to various dwelling unit age categories, surface types and aggregate measures (mean, high reading), and in general they follow shapes similar to that shown in Figure 5. That is, the distributions

Figure 5

Histogram for Mean Trim Reading/DU in Pre 1940 Units



of lead levels are highly asymmetric, typically having a long tail extending to the right. In virtually all cases considered, statistical tests indicated definite non-normality for the lead level distributions; thus, the bell-shaped normal distribution [21] does not appear to be appropriate for describing the observed lead levels*. On the other hand, both the lognormal and extreme value [21] distributions provide a reasonable fit to such data.

Because of the calibration characteristics of the XRF instruments used, it is possible to obtain negative XRF readings on a particular surface; this possibility does in fact occur in Figure 5, since the left-hand tail of the histogram extends to negative readings. One way to remedy such a situation (inasmuch as true lead contents can never be negative) is to set all negative mean readings in Figure 5 equal to zero. If this procedure is followed, then the non-normality becomes even more pronounced in the resulting frequency distributions.

In conclusion, then, it is not appropriate to use only summary statistics (such as the mean) that are based on normality in analyzing the present data. For this reason, several summary statistics -- such as the median and highest XRF reading in addition to the mean -- will be used throughout in presenting various analyses. Indeed, while the sample mean is a good estimator for moderate-tailed distributions (as the normal), it is a poor one (compared to the sample median) for long-tailed distributions [23]. It is more judicious, therefore, to base

*In fact, the distribution of individual readings must be highly non-normal, since the Central Limit Theorem [22] guarantees that mean readings conform more closely to normality than individual readings. Since the mean readings (as in Figure 5) depart considerably from normality, so must the individual readings.

any conclusions from this data set on analyses of several summary statistics.

Another way of exhibiting these statistical distributions is in the form of a cumulative distribution. For example, the histogram of Figure 5 has been redrawn as the cumulative distribution of Figure 6 (with negative mean readings being treated as zero). In the latter figure, the proportion of observations exceeding a given abscissa value (lead level) is plotted as the corresponding ordinate value; for example, Figure 6 shows that approximately 77% of the observations (mean trim readings) exceed 1.0 mg/cm^2, while some 10% of the observations exceed 7.5 mg/cm^2 in pre 1940 dwelling units. Additional cumulative distributions of this type are presented in Figures 7A and 7B. Figure 7A indicates that if the highest XRF reading in pre 1940 dwelling units is considered, about 62% of the dwellings have one or more interior walls with a lead level exceeding 2.0 mg/cm^2. If, instead, the highest XRF reading in a room is considered, then about 21% of the rooms have at least one wall exceeding 2.0 mg/cm^2, while about 13% of the interior walls (surfaces) have at least one reading in excess of 2.0 mg/cm^2. This same type of relationship is demonstrated in Figure 7B for the highest interior wall reading in 1940-1959 dwelling units. Similarly shaped cumulative distributions result when considering either the mean or the median lead measurement instead of the highest lead measurement. These graphs clearly show the pronounced effect of the degree of aggregation upon the percent of readings exceeding some fixed level. Moreover, because the cumulative distribution for individual surfaces follows the cumulative distribution for rooms much more closely than

Figure 6

Cumulative Distribution for Mean Trim Reading/DU in Pre 1940 Units

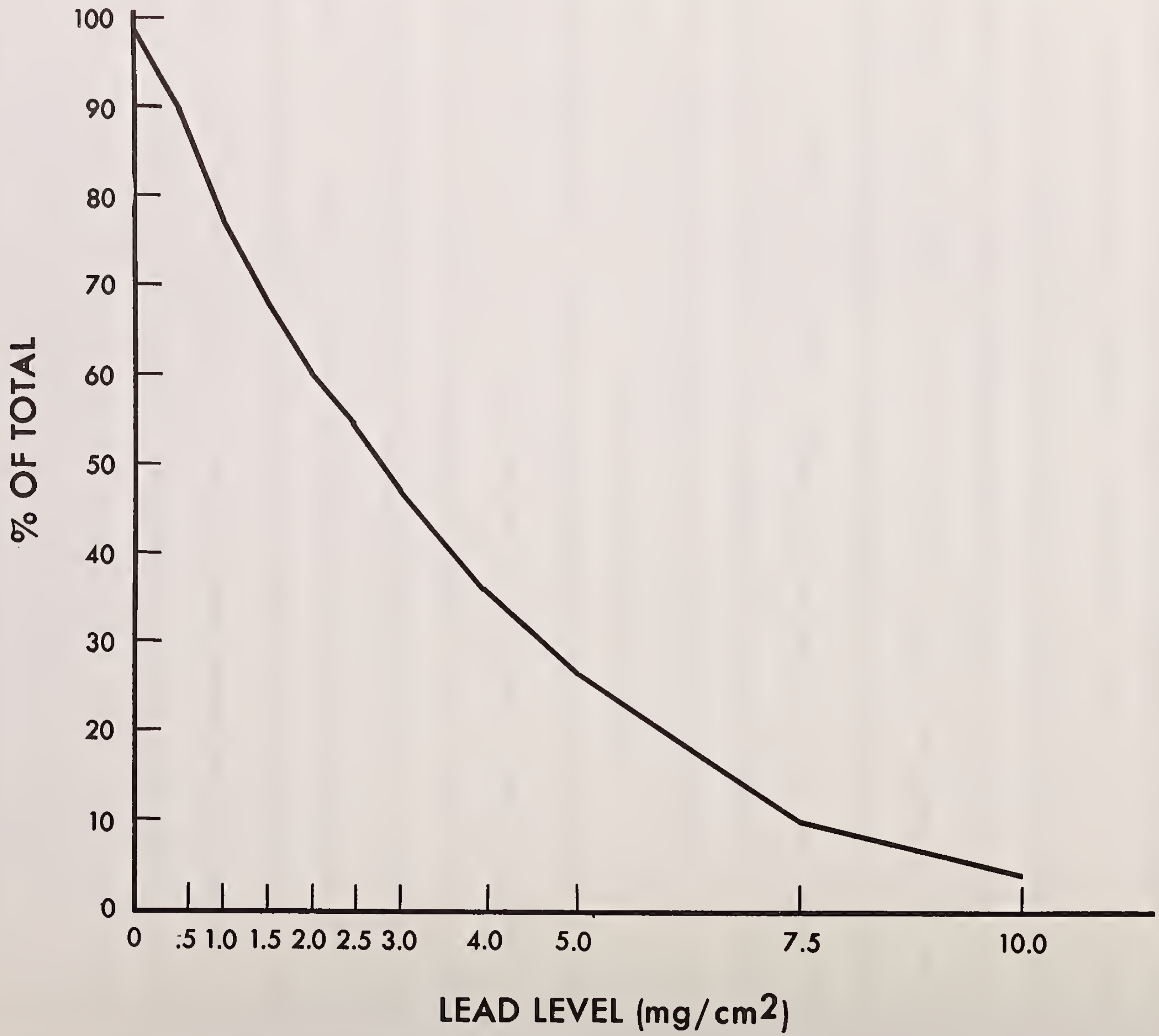


Figure 7A

Cumulative Distributions for Highest Wall Reading in Pre 1940 Units

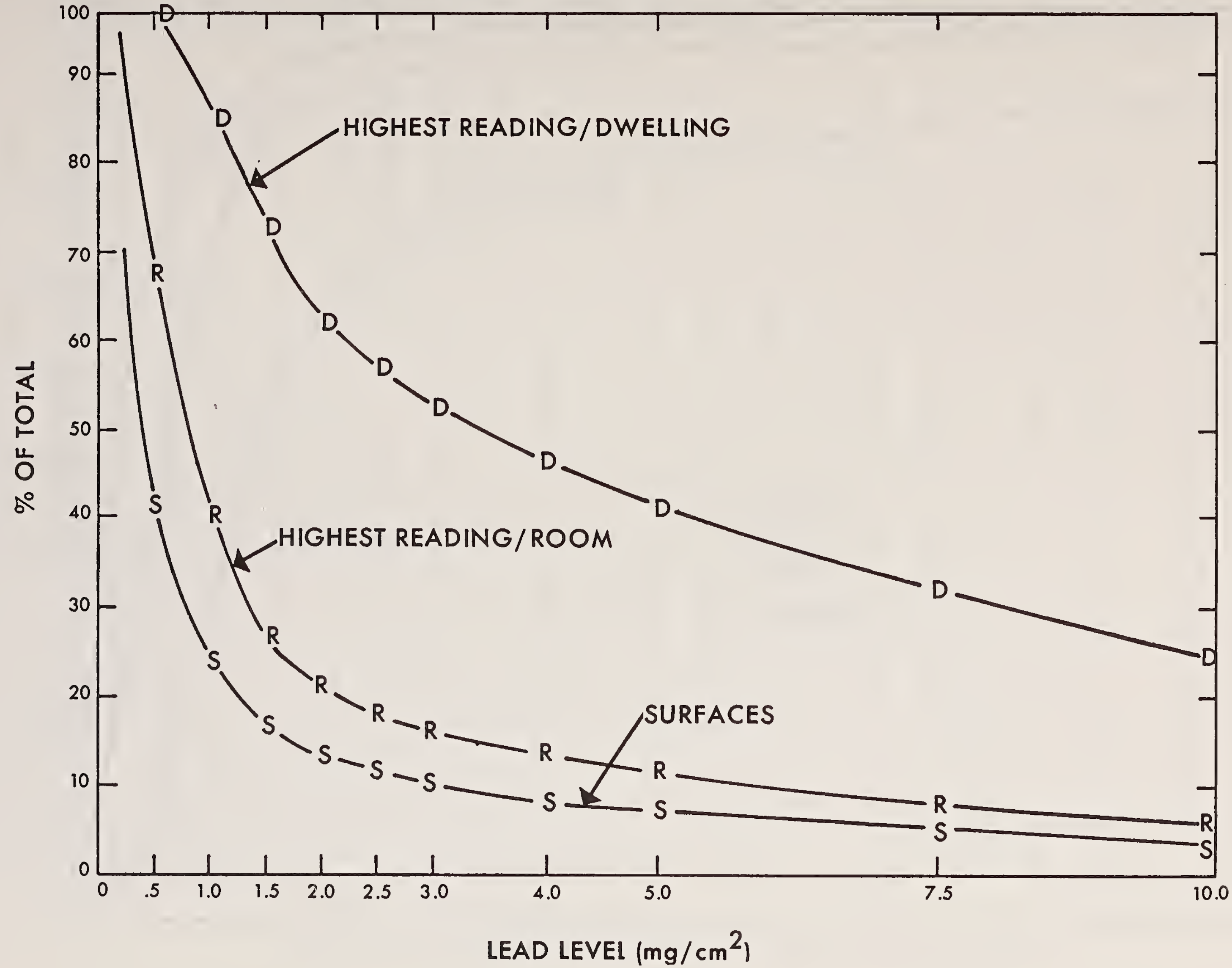
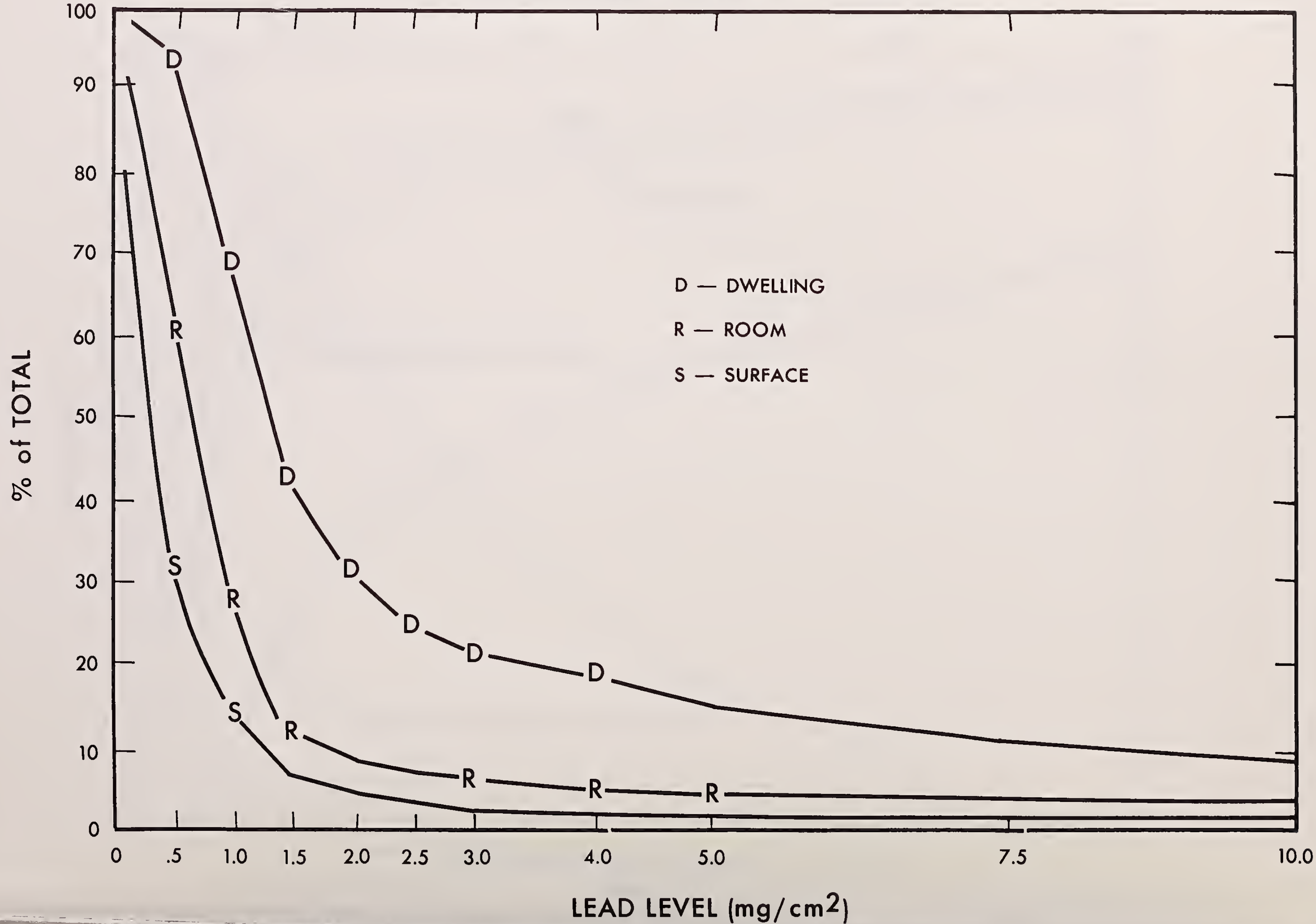


Figure 7B
Cumulative Distributions for Highest Wall Reading in 1940-1959 Units



the cumulative distribution for rooms follows that for dwelling units, it is reasonable to conclude that there is much less variation of surface readings within a room than among rooms in a dwelling unit. This conclusion is in fact borne out by further analyses of the data (not presented here).

One of the general housing characteristics whose effect is to be studied concerns the occupancy class of the dwelling unit, as defined by occupancy status (owner-renter), the number of units per building, and the extent of governmental assistance (public-private). Table 2 presents information on the percentage of various surfaces whose measured lead levels exceed certain fixed levels, according to nine occupancy class categories; such categories are consistent with classifications used in Census Bureau housing publications. In essence, this table provides an alternative representation for cumulative lead level distributions. By comparing the various columns of Table 2, it is seen that (regardless of surface type) single-family housing and multifamily housing are quite similar in terms of observed lead levels. On the other hand, public housing exhibits noticeably lower lead levels than either single-family or multifamily housing. For single-family housing, there is no substantial difference in lead levels between owner-occupied and renter-occupied housing, while for multifamily housing those buildings with 2-4 units show higher lead levels than buildings with 5 or more units.

Another important general housing characteristic is the age of housing, defined previously in terms of three age categories. Figures 8A-8C show cumulative lead level distributions for certain surfaces, according to these age categories. In particular, Figure 8A displays

TABLE 2

Percentage of Interior Surfaces Exceeding Fixed Lead Levels

Interior Surface	Level (mg/cm^2)	SINGLE			MULTI			PUBLIC		
		Owner	Renter	All	2-4	> 4	All	2-4	> 4	All*
Walls	0.5	37.7	38.2	38.2	41.4	35.2	38.1	33.3	33.8	33.5
	1.0	20.1	20.7	20.7	25.4	16.8	20.9	11.3	15.2	15.8
	2.0	10.8	11.4	11.3	15.8	7.7	11.5	5.9	6.0	7.0
	5.0	5.9	6.3	6.3	8.7	4.2	6.4	5.4	2.6	2.9
	7.5	4.2	4.4	4.4	6.3	2.9	4.5	5.4	0.9	1.4
Windows	0.5	67.7	69.5	69.4	73.4	59.9	68.0	57.6	55.5	58.5
	1.0	50.3	52.7	52.5	60.8	40.5	52.6	33.3	27.4	35.3
	2.0	34.8	37.0	36.8	45.8	24.3	37.1	30.3	7.7	15.8
	5.0	20.3	21.9	21.7	29.3	14.0	23.1	30.3	3.9	9.1
	7.5	12.4	13.5	13.4	19.2	9.5	15.3	30.3	3.4	6.3
Doors	0.5	73.4	74.6	74.5	78.8	66.5	72.9	82.5	64.6	69.8
	1.0	53.3	55.2	55.0	65.1	40.1	53.2	56.1	36.1	43.9
	2.0	33.0	34.9	34.8	47.7	17.4	33.2	31.6	9.8	16.0
	5.0	21.6	22.8	22.7	35.5	9.7	23.2	22.8	2.1	7.0
	7.5	13.4	14.2	14.2	23.1	5.8	14.8	14.0	1.3	4.2

*Note that this ALL category includes single family as well as multifamily units.

Figure 8A

Cumulative Distributions for Dry Room Walls, by Age Category

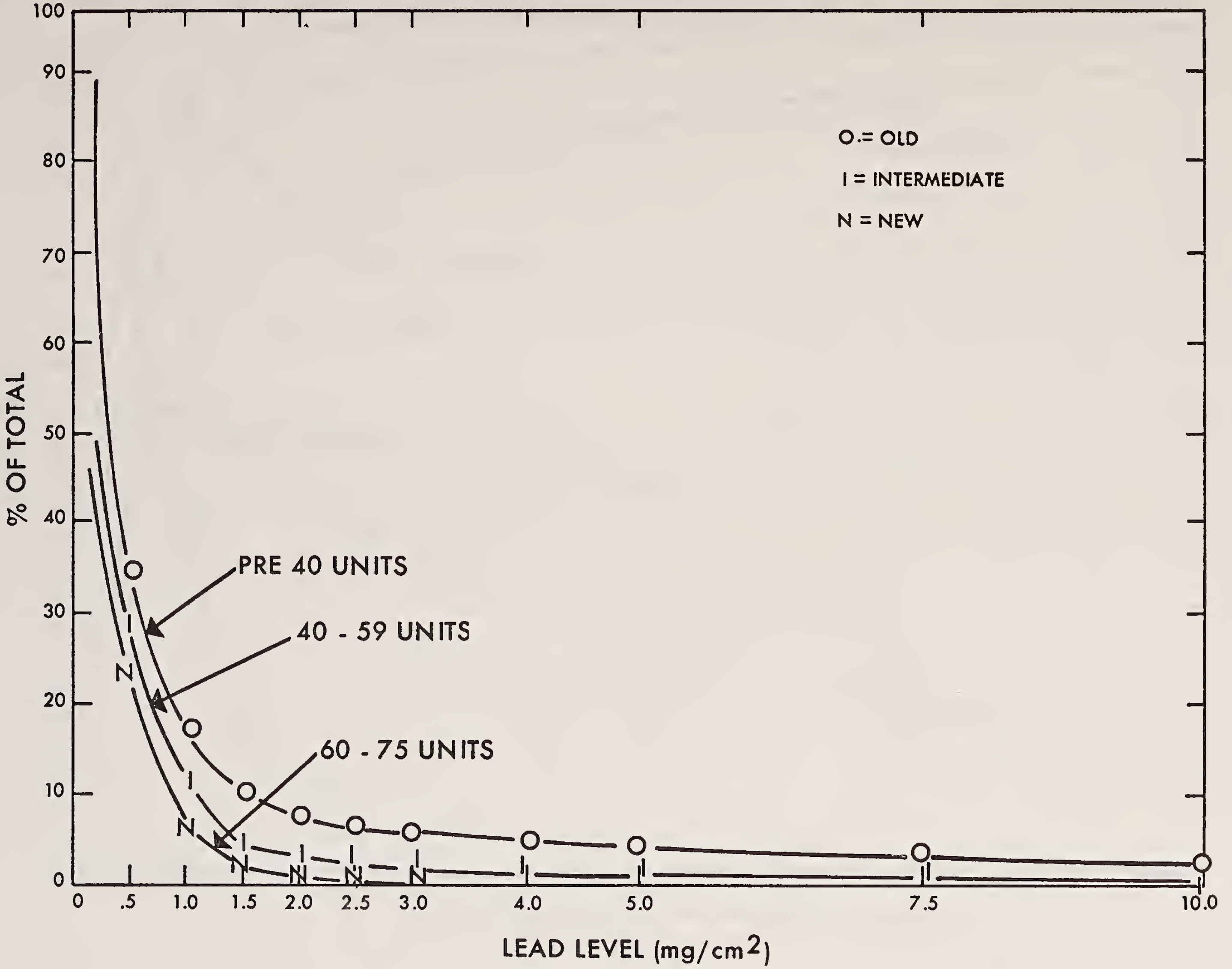


Figure 8B

Cumulative Distributions for Wet Room Walls, by Age Category

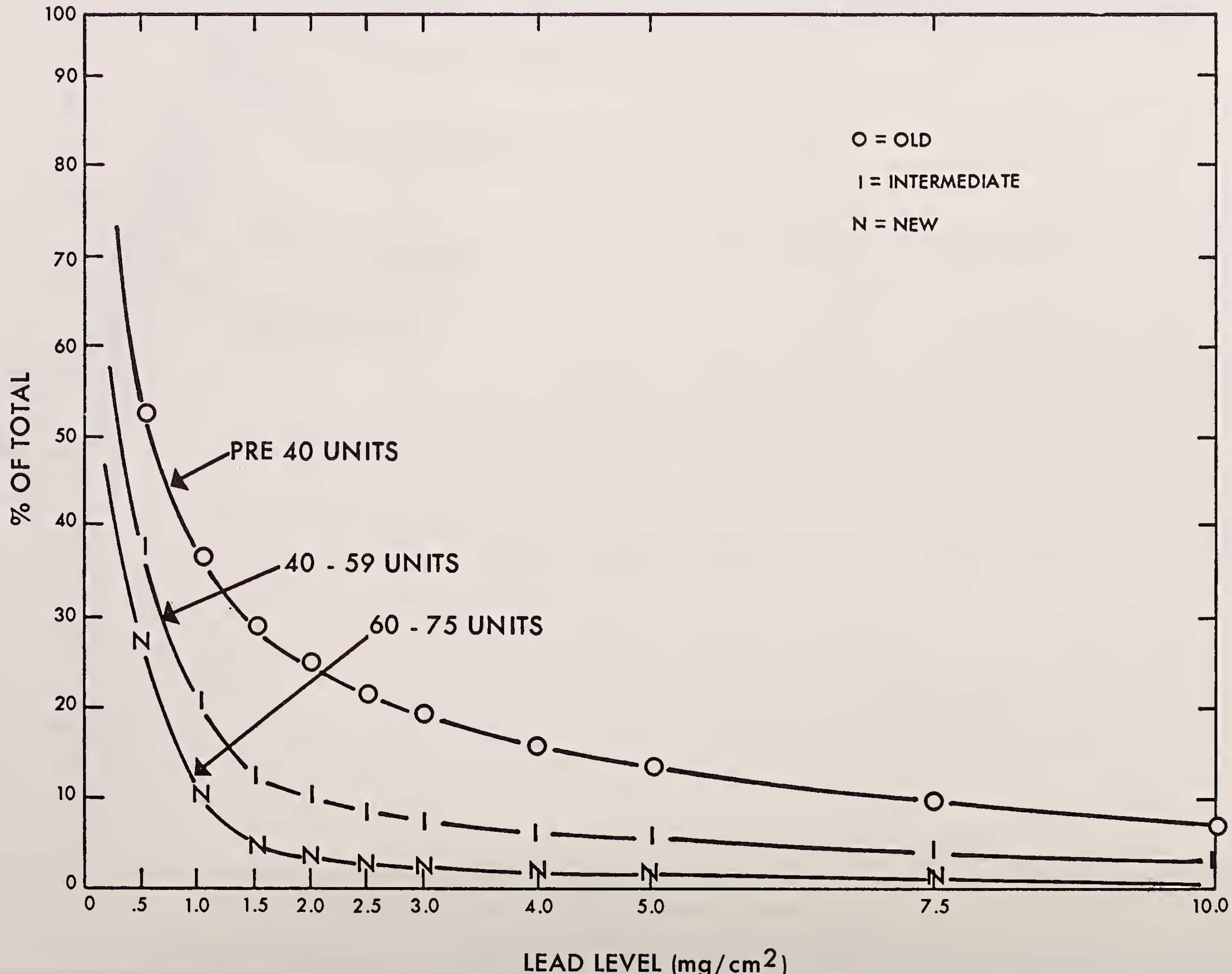
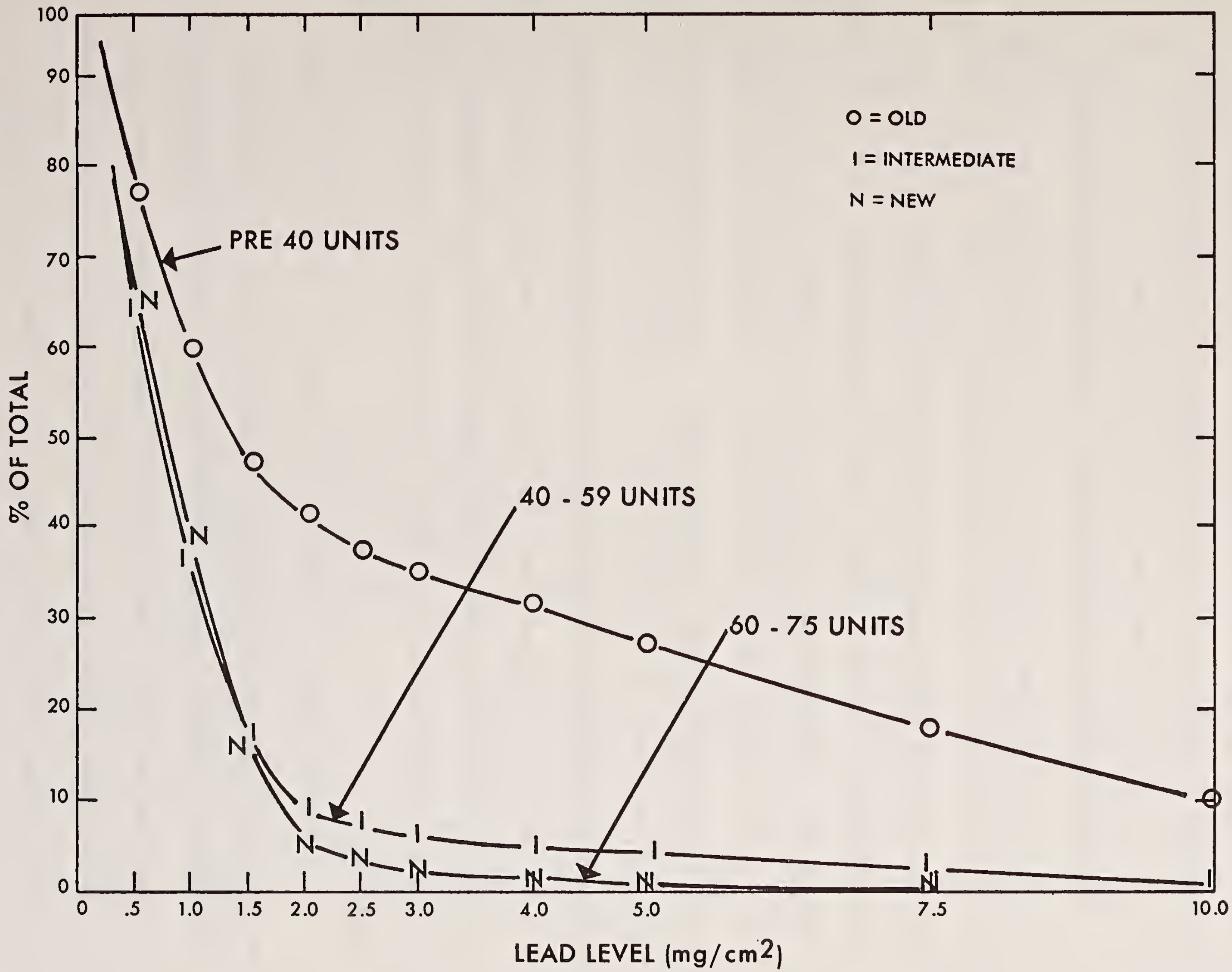


Figure 8C

Cumulative Distribution for All Doors, by Age Category



cumulative distributions for walls occurring in dry rooms, grouped by the age categories: pre 1940, 1940-1959, 1960-1975. This figure shows that pre 1940 units contain higher lead levels than 1940-1959 units, which in turn contain higher lead levels than 1960-1975 units. This general observation of greater lead levels in older housing is also borne out in Figure 8B (wet room walls) and Figure 8C (all doors).

The relationship of housing age to lead levels can likewise be studied using aggregate measures such as the mean, median and highest lead levels present in rooms of dwelling units. Table 3 gives the average value for each of these measures over various subpopulations defined by age category, occupancy class, room type and surface type. For every combination of occupancy class, room type and surface type, and for each of the three measures considered in Table 3, the measured lead levels of pre 1940 dwelling units is the greatest, that of 1960-1975 units is the least, and that of 1940-1959 units is intermediate between these two. Typically, the pre 1940 units have lead levels which are three times those of the 1940-1959 units and five times those of the 1960-1975 units. It is also observed (see Table 4) that the pre 1940 public housing units have substantially lower lead levels than the general pre 1940 housing population.

In addition, Table 3 shows (as seen earlier) that single-family and multifamily housing are generally similar in terms of measured lead levels. However, in pre 1940 units the multifamily housing tends to have somewhat higher lead levels than single-family housing, while in 1960-1975 units single-family housing tends to have somewhat higher lead levels than multifamily housing. It is also clear from Table 3

TABLE 3

Average Lead Level Measures by Room (mg/cm^2)

Occupancy	Room	Surface	Pre 1940 Units			1940-1959 Units			1960-1975 Units			All Units*		
			Mean	Median	High	Mean	Median	High	Mean	Median	High	Mean	Median	High
Single	Dry	Walls	.76	.66	1.59	.25	.23	.82	.07	.08	.57	.65	.58	1.42
		Trim	3.17	2.91	4.79	.94	.91	1.45	.82	.83	1.17	2.73	2.51	4.12
	Wet	Walls	2.08	1.83	3.49	1.05	.93	1.90	.28	.28	.89	1.84	1.68	3.13
		Trim	3.55	3.24	5.24	1.25	1.21	1.84	.98	.91	1.48	3.07	2.79	4.54
	All	Walls	1.17	1.02	2.17	.49	.44	1.14	.14	.14	.67	1.01	.91	1.94
		Trim	3.29	3.02	4.93	1.03	1.01	1.57	.88	.86	1.28	2.83	2.60	4.25
Multi	Dry	Walls	.79	.69	1.66	.29	.28	.85	.0	.0	.50	.59	.52	1.36
		Trim	3.46	3.19	5.34	1.03	1.02	1.43	.55	.54	.83	2.63	2.42	4.03
	Wet	Walls	2.51	2.14	4.06	.71	.63	1.50	.10	.09	.61	1.79	1.61	3.06
		Trim	3.62	3.43	5.42	1.08	.98	1.56	.59	.58	.88	2.68	2.48	4.00
	All	Walls	1.44	1.23	2.56	.46	.42	1.12	.04	.04	.54	1.05	.94	2.01
		Trim	3.52	3.28	5.37	1.05	1.00	1.48	.57	.56	.85	2.65	2.45	4.02
All	Dry	Walls	.77	.68	1.60	.26	.24	.83	.04	.04	.54	.64	.57	1.41
		Trim	3.23	2.97	4.89	.96	.94	1.45	.70	.70	1.02	2.71	2.50	4.11
	Wet	Walls	2.19	1.97	3.62	.93	.85	1.76	.19	.18	.74	1.82	1.65	3.09
		Trim	3.57	3.29	5.28	1.19	1.13	1.74	.78	.74	1.18	2.97	2.71	4.39
	All	Walls	1.22	1.09	2.25	.48	.44	1.14	.09	.09	.61	1.03	.93	1.96
		Trim	3.34	3.08	5.02	1.04	1.00	1.55	.73	.72	1.08	2.80	2.57	4.21

*Note that this age category also includes units for which the age of the structure was unknown.

TABLE 4

Percentage of Interior Surfaces in Pre 1940 Units
Exceeding Fixed Lead Levels

Occupancy	Level (mg/cm^2)	DRY			WET		
		Walls	Windows	Doors	Walls	Windows	Doors
Single	0.5	34.5	7.15	75.4	52.6	77.4	80.9
	1.0	16.6	54.9	57.2	36.1	61.6	65.5
	2.0	7.8	39.1	40.1	24.1	44.5	45.9
	5.0	4.6	23.7	26.9	13.1	25.2	29.8
	7.5	3.2	14.7	16.5	9.2	15.5	19.1
Multi	0.5	35.9	70.6	74.5	53.1	77.1	80.5
	1.0	17.7	49.9	44.1	37.3	63.2	65.7
	2.0	8.8	41.3	40.5	26.2	46.6	48.5
	5.0	4.5	26.2	30.8	15.7	29.1	33.4
	7.5	3.0	17.0	19.6	12.0	20.0	22.1
Public	0.5	33.4	64.0	70.1	43.7	68.4	68.3
	1.0	15.7	43.9	43.4	24.6	41.0	41.3
	2.0	8.2	23.0	21.2	12.3	18.8	19.8
	5.0	3.4	13.8	12.8	6.1	8.5	10.3
	7.5	1.7	7.9	8.0	3.6	7.7	7.1
All	0.5	34.8	71.4	75.3	52.7	77.3	80.8
	1.0	16.9	55.2	57.3	36.3	61.9	65.5
	2.0	8.0	39.6	40.4	24.6	44.9	46.4
	5.0	4.6	24.2	27.8	13.7	26.0	30.6
	7.5	3.2	15.1	17.2	9.8	16.4	19.8

that lead levels in wet rooms are consistently higher than lead levels in dry rooms, especially for the wall readings. This difference between wet rooms and dry rooms can alternatively be seen in the cumulative distributions graphed in Figures 9A-9F for various age categories and surface types.

The cumulative distributions of Figures 9A-9F also indicate that (within any age category) surface type is important in explaining variation in lead levels. This effect can also be clearly seen in Table 2, where the windows and doors appear distinctly different from walls in terms of the percentage of surfaces exceeding certain fixed lead levels. This observed difference between windows (doors) and walls holds across all occupancy classes for all fixed lead levels. In Table 3, a marked difference between walls and trim (doors, windows, baseboards, radiators, cabinets and fireplaces) is likewise evident: trim surfaces exhibit substantially higher lead levels than do walls, for all indicated combinations of occupancy class, room type and age category. In addition, Table 3 indicates that within each age category, dry room walls have the lowest lead levels of any surface type, followed by wet room walls, dry room trim, and wet room trim. These differences by surface and room type are much more pronounced in pre 1940 housing than in the newer units.

Alternatively, the variation in lead levels by surface type can also be calculated using <u>rooms</u> of the dwelling unit. For example, Figure 10A shows in pre 1940 dry rooms the mean trim lead level for all rooms having a mean wall lead level lying within specified intervals. Thus, although some 70.7% of the rooms have a mean wall lead level lying between 0.0 and 0.5 mg/cm^2, the trim surfaces within the same rooms

Figure 9A

Cumulative Distributions for Walls in Pre 1940 Units, by Room Type

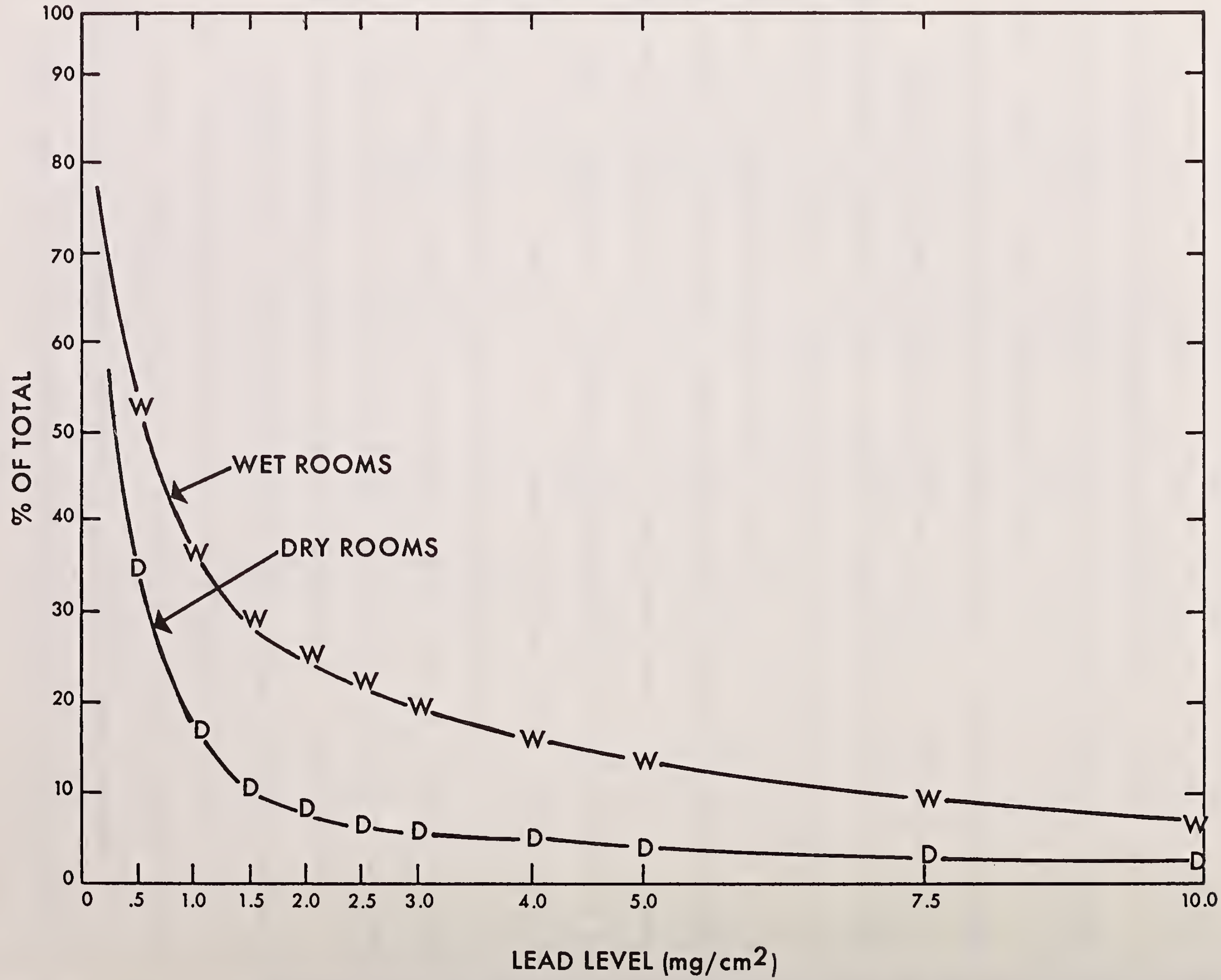


Figure 9B

Cumulative Distribution for Doors in pre 1940 Units, by Room Type

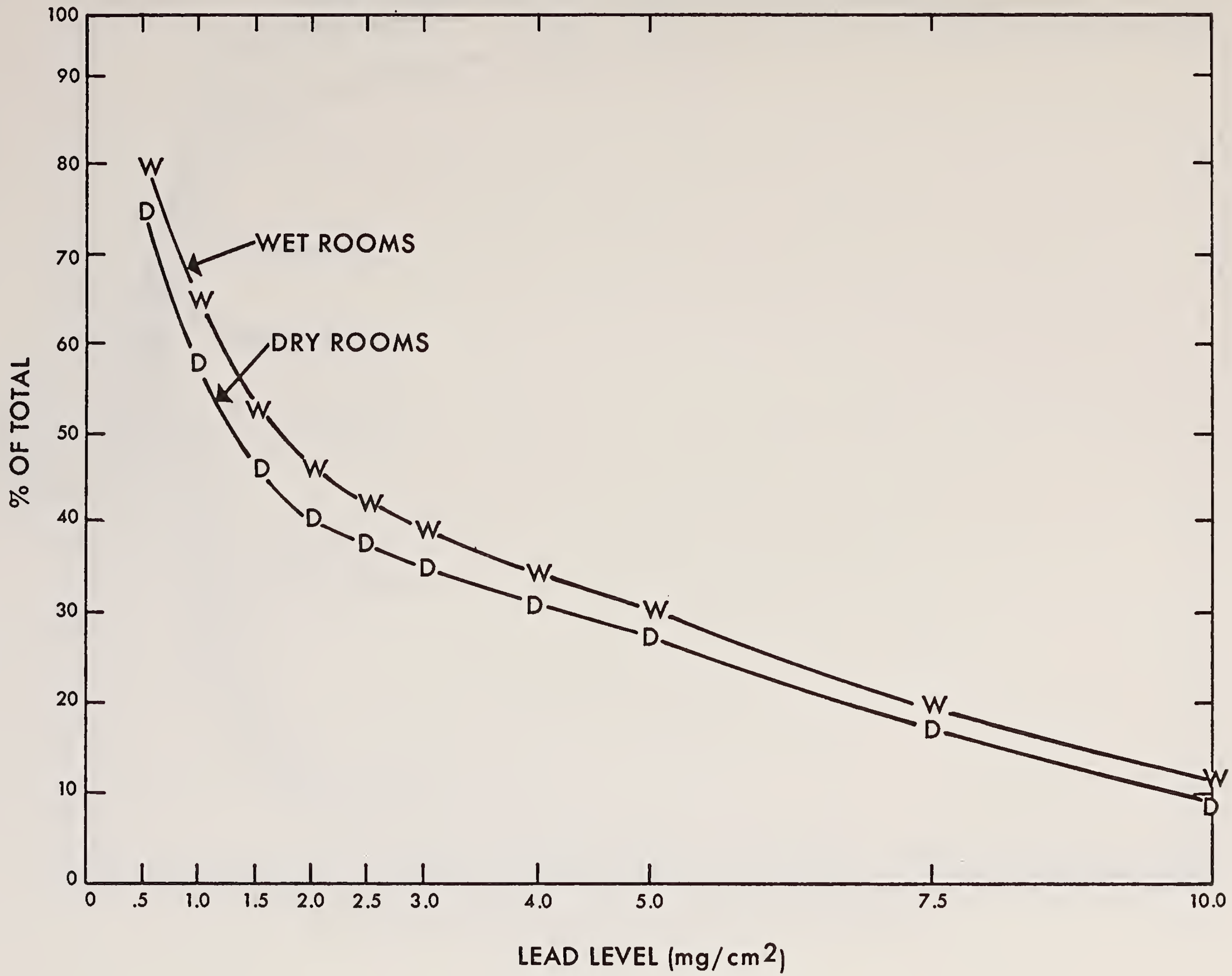


Figure 9C

Cumulative Distributions for Walls in 1940-1950 Units, by Room Type

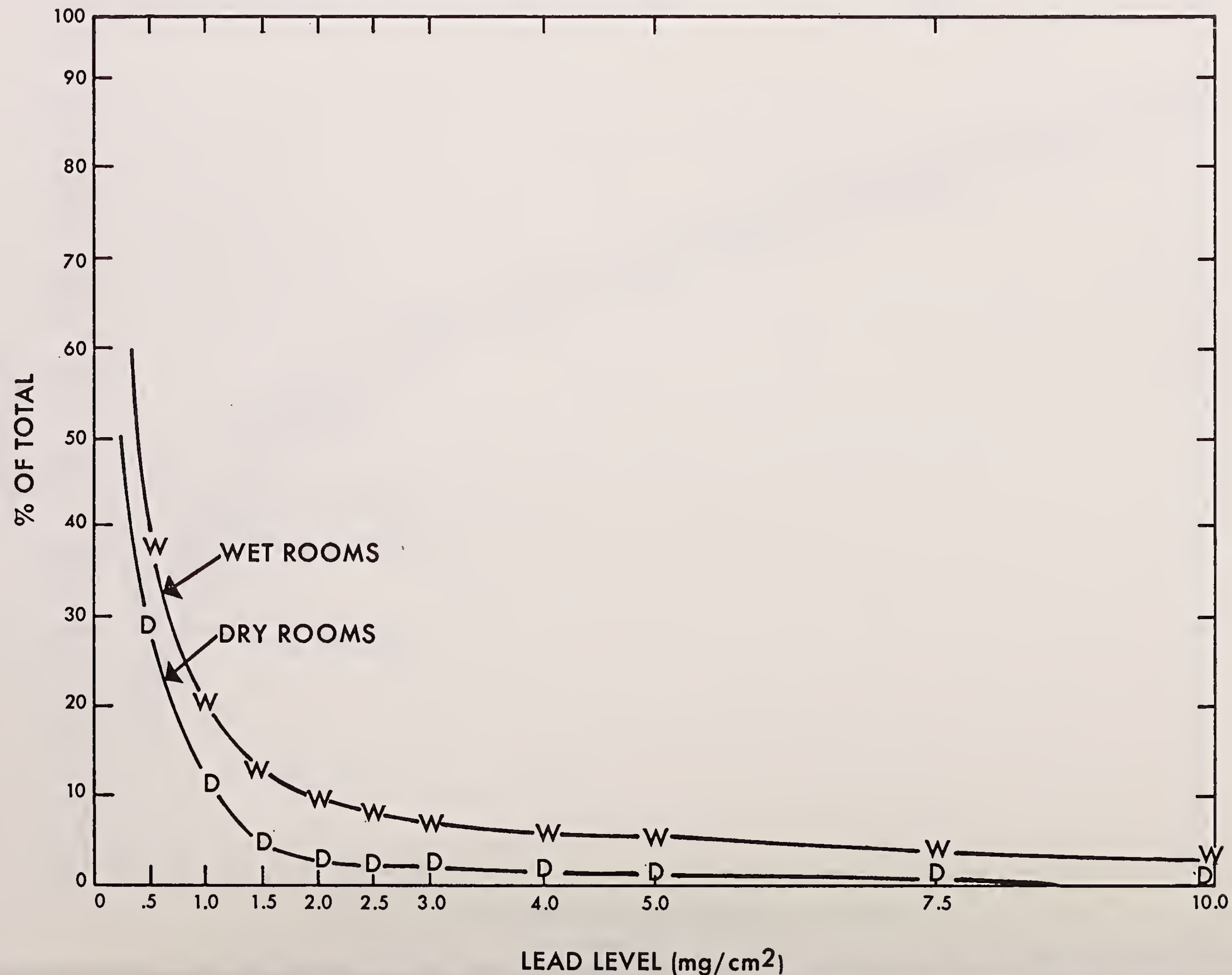


Figure 9D

Cumulative Distributions for Doors in 1940-1959 Units, by Room Type

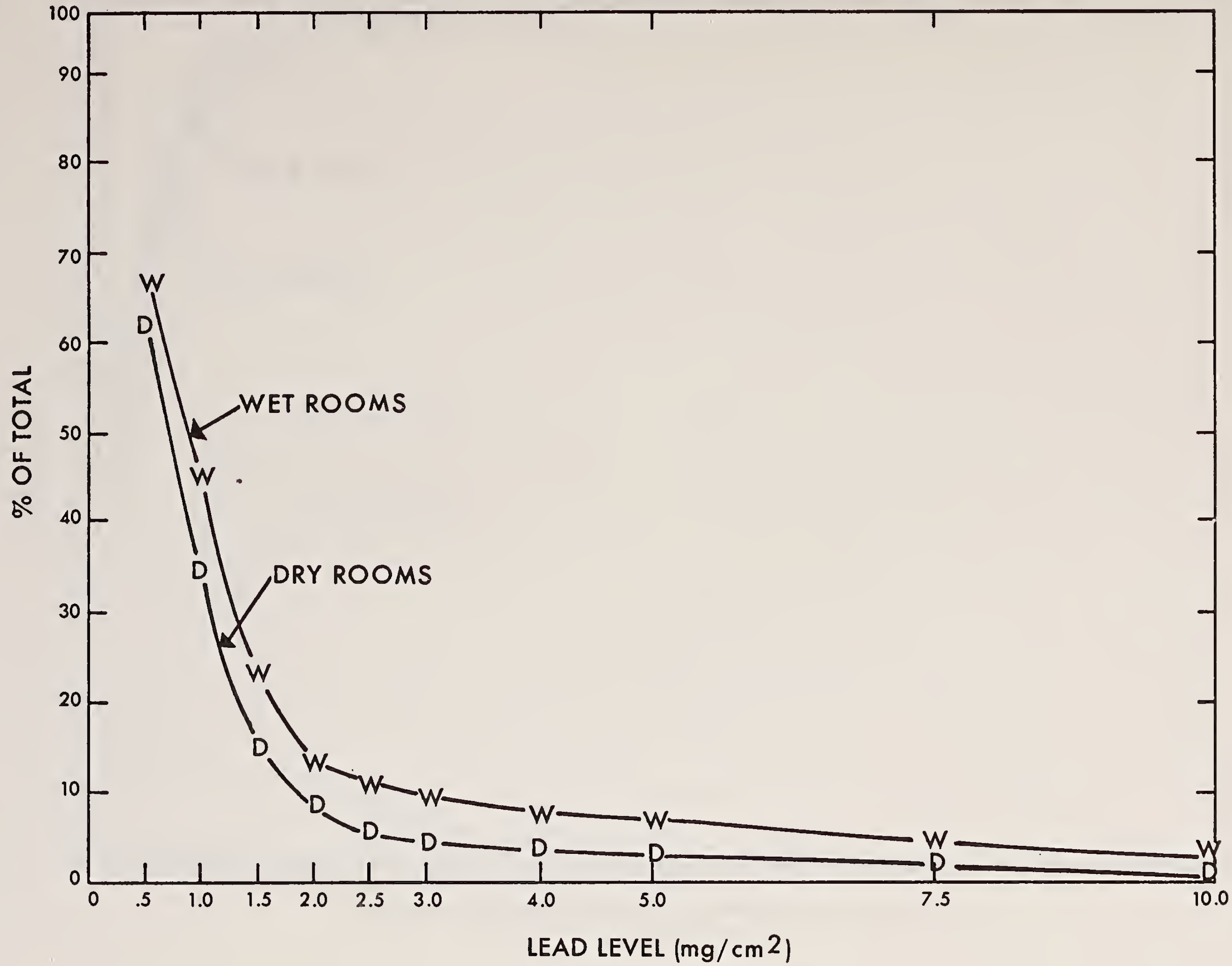


Figure 9E

Cumulative Distributions for Walls in 1960-1975 Units, by Room Type

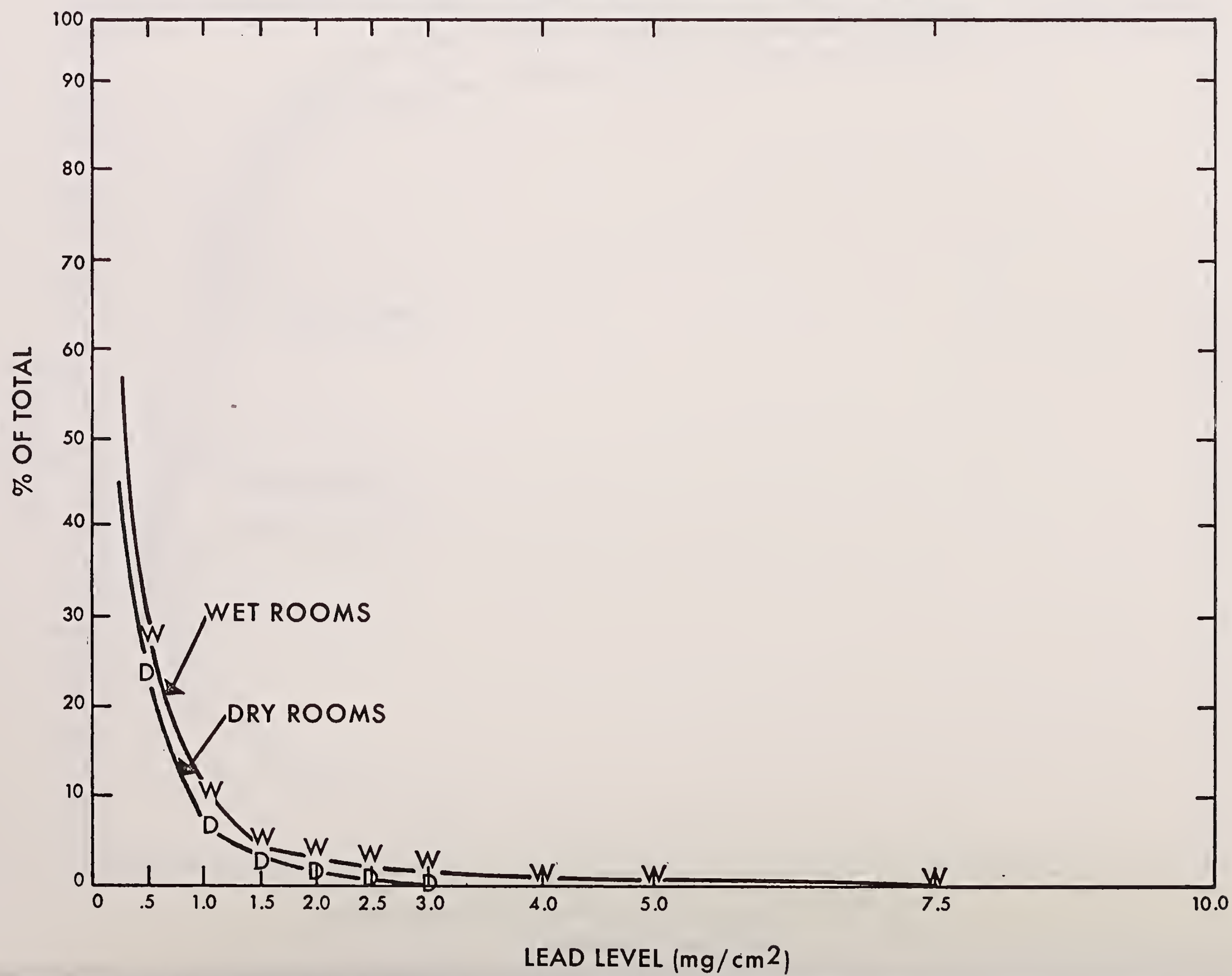


Figure 9F

Cumulative Distributions for Doors in 1960-1975 Units, by Room Type

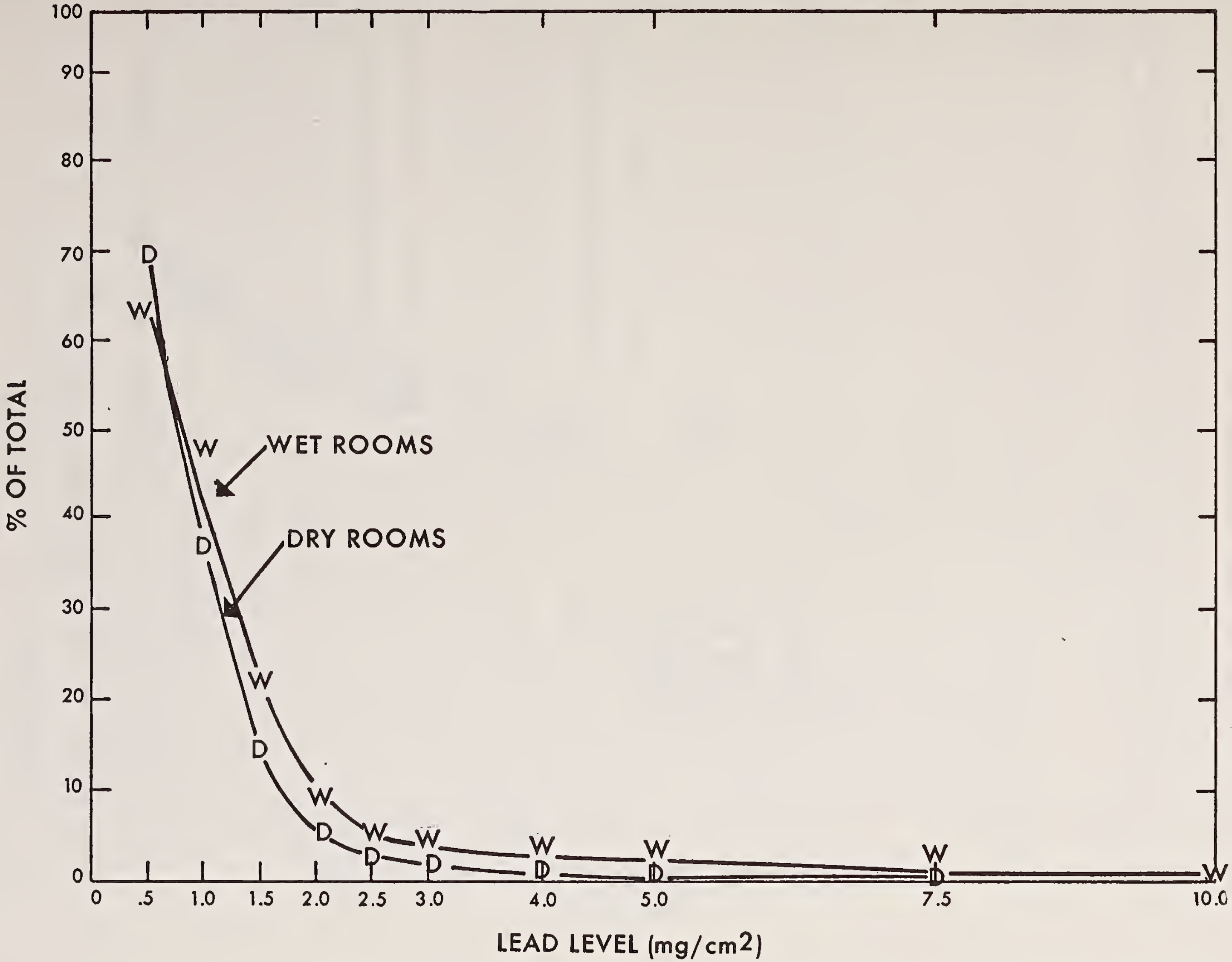
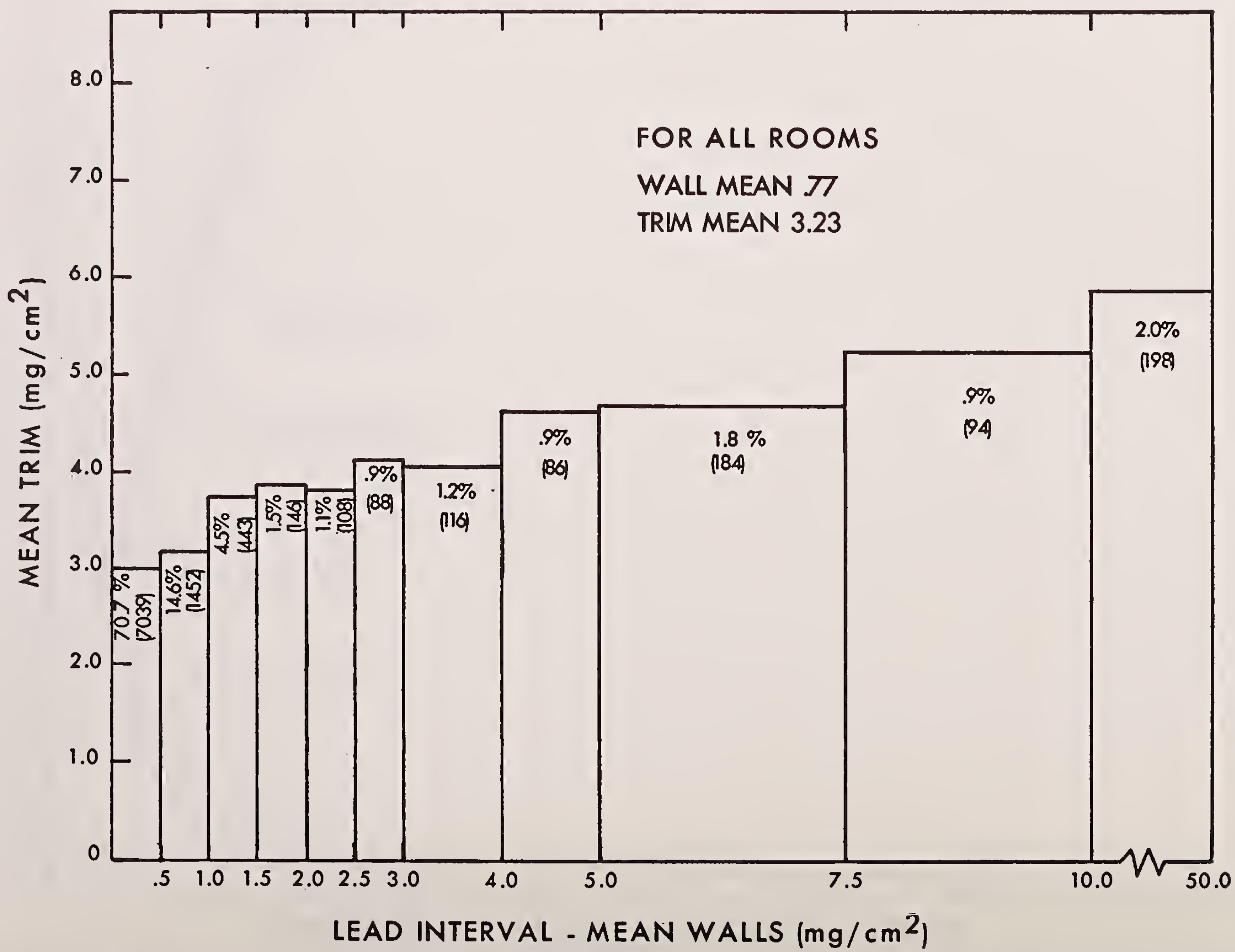


Figure 10A

Mean Trim vs. Lead Interval for Walls in Dry Rooms, Pre 1940 Units



have a mean of 3.0 mg/cm^2. Note that the number of rooms having a mean lead level within the specified interval is shown in parentheses immediately beneath the percentage of rooms within the interval. This figure indicates that at mean wall lead levels less than 5.0 mg/cm^2 (including 95.2% of all rooms), the mean trim lead levels can be substantially greater than the mean wall levels. A similar relationship is observed in Figure 10B, which pertains to pre 1940 wet rooms.

Another factor which has been considered in analyzing the present data on lead levels is the condition of the surface. Recall that a PEELING surface indicates one on which the paint film is peeling, flaking, blistering or powdering. A BAD surface is one whose substrate contains imperfections (apart from hairline cracks, nail holes, etc.). Since a condition code was assigned to a room (and not to individual surfaces within a room), it is appropriate to assess the effect of condition on lead levels using aggregate measures for rooms. In particular, Table 5 displays the percentage of rooms whose highest lead level exceeds certain fixed lead levels, classified by whether the surface condition is PEELING or BAD or both (P/B), or whether it is neither PEELING nor BAD (i.e., GOOD). Thus, 62.9% of all pre 1940 dry rooms with walls in GOOD condition have their highest wall reading in the room exceeding 0.5 mg/cm^2. For virtually all combinations of surface type, room type and age category, rooms with a PEELING/BAD condition code show greater lead levels than rooms with a GOOD condition code. Furthermore, Table 6 indicates that for each of the indicated surface types about 25% of all dwelling units can be classified as having at least one room in PEELING or BAD condition (P/B). Approximately

Figure 10B
Mean Trim vs. Lead Interval for Walls in Wet Rooms, Pre 1940 Units

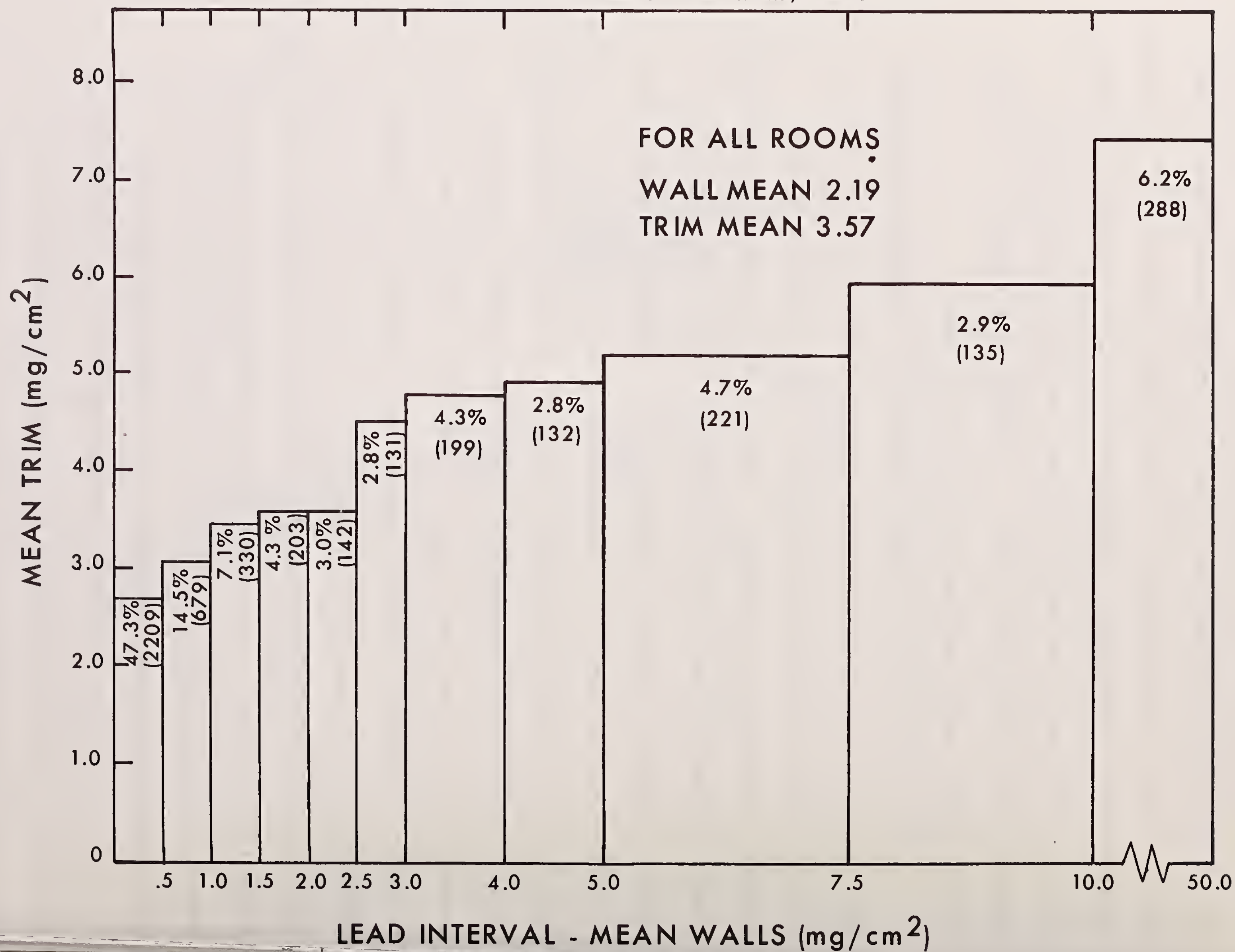


TABLE 5

Percentage of Rooms with Highest Reading
Exceeding Fixed Lead Levels

		WALLS				WINDOWS			
		DRY		WET		DRY		WET	
Age	Level (mg/cm^2)	P/B*	Good	P/B	Good	P/B	Good	P/B	Good
Pre 1940	0.5	68.2	62.9	86.2	77.7	83.9	68.8	85.7	75.0
	1.0	38.8	31.8	67.3	54.1	70.4	52.0	74.2	59.2
	2.0	19.6	12.0	46.8	34.9	55.7	36.0	57.2	42.4
	5.0	12.0	6.7	33.0	19.3	36.6	21.6	35.3	24.4
	7.5	8.3	4.6	25.1	14.0	22.4	13.6	24.9	15.1
1940-1959	0.5	46.9	56.6	79.5	66.4	66.2	43.6	62.5	44.4
	1.0	21.9	23.5	46.2	36.4	41.2	22.3	46.9	25.7
	2.0	15.6	4.8	33.3	14.9	22.1	9.0	28.1	11.4
	5.0	9.4	2.0	23.1	7.6	11.8	3.7	25.0	5.9
	7.5	6.3	1.2	12.8	5.5	5.9	2.1	15.6	4.0
1960-1975	0.5	57.1	48.8	60.0	53.8	36.4	37.3	75.0	40.9
	1.0	14.3	15.8	40.0	22.7	27.3	17.7	75.0	22.7
	2.0	0.0	2.0	40.0	5.6	18.2	4.4	75.0	10.9
	5.0	0.0	0.3	20.0	1.0	0.0	1.6	50.0	4.5
	7.5	0.0	0.0	0.0	0.8	0.0	0.4	0.0	1.8

*P/B = PEELING or BAD or BOTH

TABLE 6

Percentage of DU's (Rooms) Having Surfaces in Peeling or Bad Condition

	WALLS		WINDOWS		DOORS	
	DU*	ROOMS	DU	ROOMS	DU	ROOMS
Pre 1940	32.2	13.8	29.2	15.3	27.4	14.8
1940 - 1959	10.0	3.1	12.8	6.5	9.4	4.2
1960 - 1975	2.5	1.1	8.1	4.0	6.1	2.7
All	26.5	11.2	26.3	13.9	23.0	12.3

*The abbreviation "DU" is used here and subsequently to indicate "dwelling unit".

12-13% of the rooms in all dwelling units have surfaces considered in PEELING or BAD condition. As this table also shows, these percentages are highest for pre 1940 units, intermediate for 1940-1959 units and lowest for 1960-1975 units.

In all the analyses presented so far, attention has focused on the interior surfaces of the dwelling unit. Lead level readings were also taken at specified locations on the exterior of the unit, and these data are now presented. Tables 7A-7B permit a comparison of lead levels detected on exterior surfaces with those found on interior surfaces. In these tables, each dwelling unit has been characterized by a mean (or high) reading for all surfaces and classified according to age category, surface type, room type and location (interior/exterior). As Tables 7A and 7B indicate, exterior surfaces show considerably higher lead levels than interior surfaces, regardless of age category, surface type or room type. Moreover, this difference between exterior and interior lead levels is greater for walls than for trim. In addition, lead levels found on exterior surfaces are seen to vary according to the age of housing, with the pre 1940 units displaying greater levels (by surface type) than 1940-1959 units, which in turn show greater levels than 1960-1975 units.

Figures 11A-11D compare some of the cumulative distributions from the present Pittsburgh study with those obtained from the Washington, D.C. pre-test [7]. Since the Washington data (also based on a simple random sample) represent only about 100 dwelling units, it is not clear how well such data characterize the housing population of that city. However, in most instances the distributions for Pittsburgh and Washington

TABLE 7A

Average Value for DU Mean Reading (mg/cm^2)

		INTERIOR			EXTERIOR
AGE	SURFACE	DRY	WET	ALL	ALL
Pre 1940					
	Walls	.89	2.46	1.33	10.22
	Trim	3.55	3.77	3.58	6.07
	All	1.97	3.05	2.26	6.45
1940 - 1959					
	Walls	.38	.77	.49	4.75
	Trim	1.14	1.36	1.22	2.51
	All	.57	.96	.68	2.54
1960 - 1975					
	Walls	.21	.27	.21	1.37
	Trim	.76	.89	.79	2.14
	All	.31	.47	.34	2.11
All					
	Walls	.75	2.03	1.11	9.72
	Trim	3.01	3.24	3.04	5.33
	All	1.62	2.54	1.87	5.63

TABLE 7B

Average Value for DU High Reading (mg/cm^2)

AGE	SURFACE	INTERIOR DRY	INTERIOR WET	INTERIOR ALL	EXTERIOR ALL
Pre 1940					
	Walls	3.22	4.90	5.67	11.04
	Trim	8.35	7.29	9.96	11.32
	All	8.80	8.32	11.11	12.65
1940 - 1959					
	Walls	1.46	1.75	2.09	5.41
	Trim	2.48	2.51	3.13	3.85
	All	2.61	2.69	3.43	4.12
1960 - 1975					
	Walls	.83	.78	1.04	1.37
	Trim	1.46	1.63	1.85	2.79
	All	1.46	1.57	1.88	2.81
All					
	Walls	2.76	4.09	4.76	10.50
	Trim	7.03	6.24	8.40	9.74
	All	7.26	6.96	9.21	10.86

Figure 11A

Cumulative Distributions (Pittsburgh and Washington) for Interior Walls in Pre 1940 Units

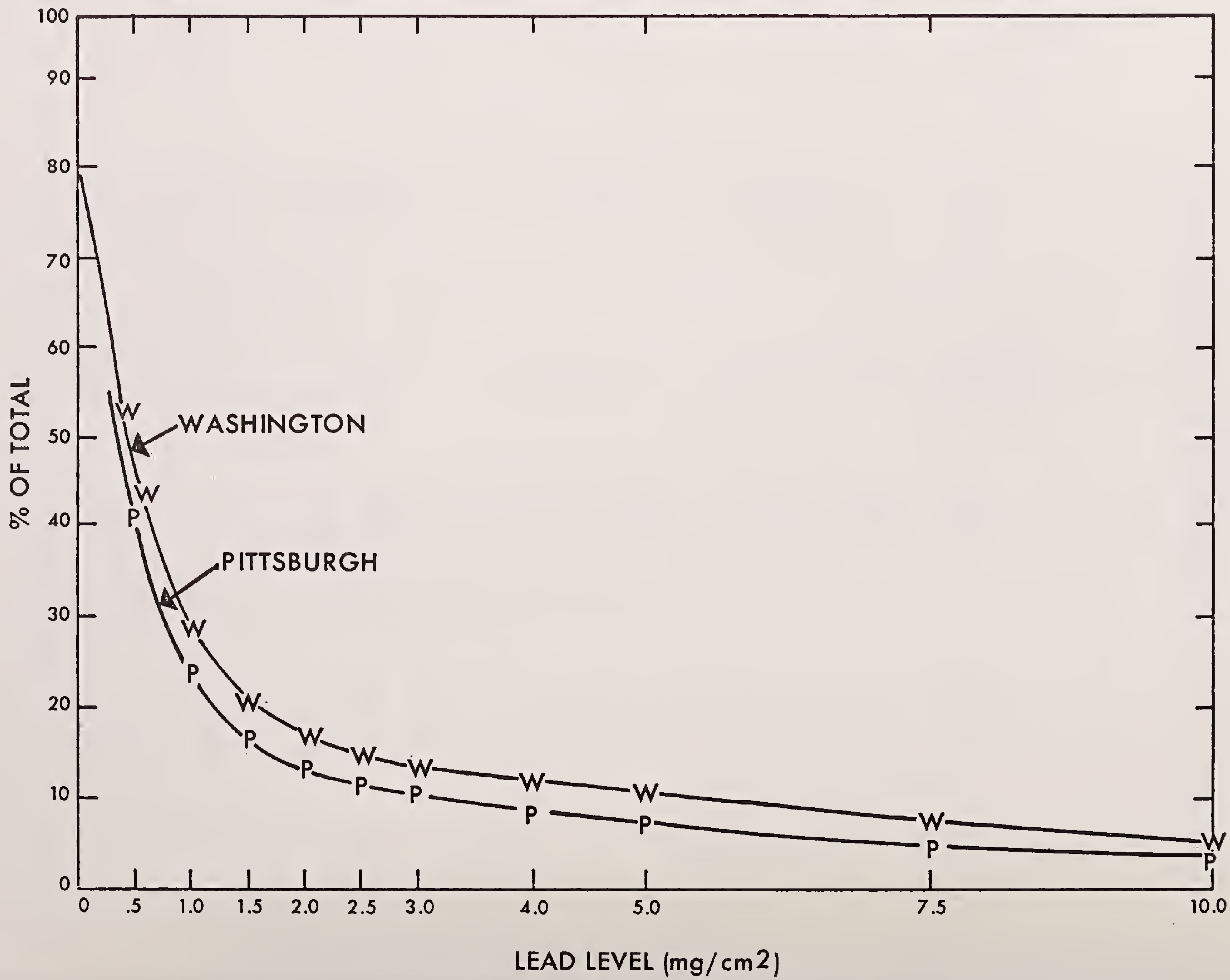


Figure 11B

Cumulative Distributions (Pittsburgh and Washington) for Interior Walls in 1940-1959 Units

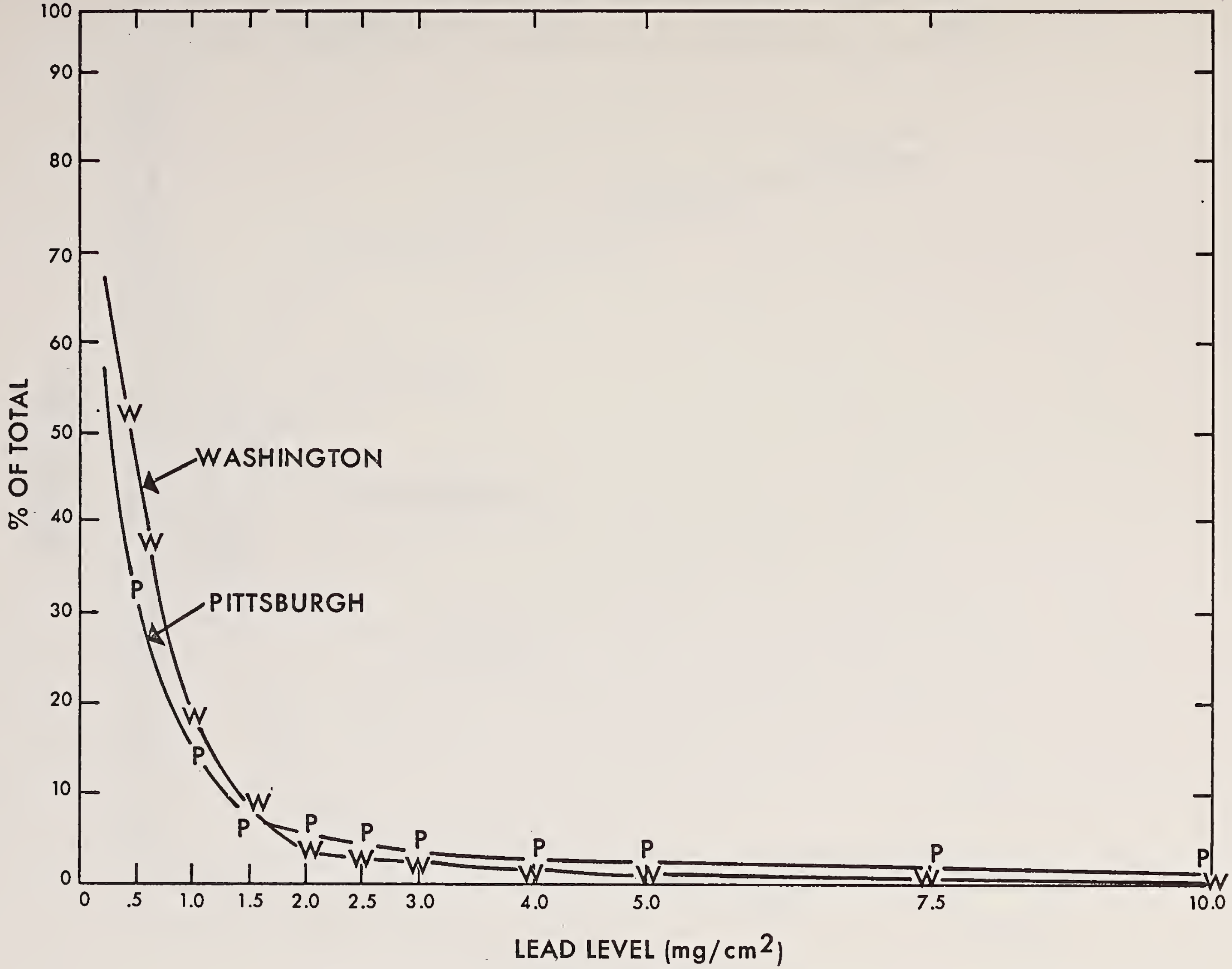


Figure 11C

Cumulative Distributions (Pittsburgh and Washington) for Interior Doors in Pre 1940 Units

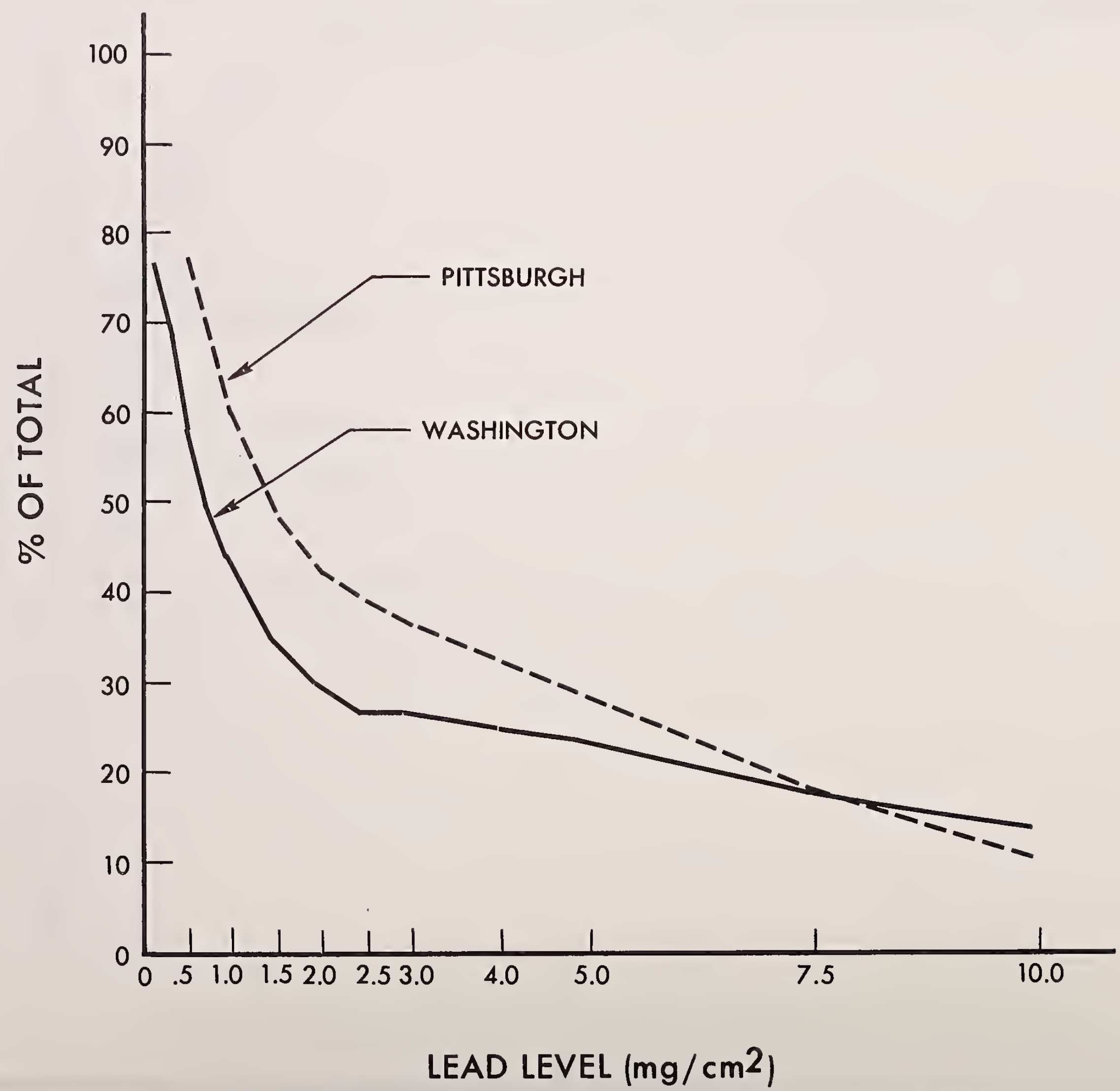
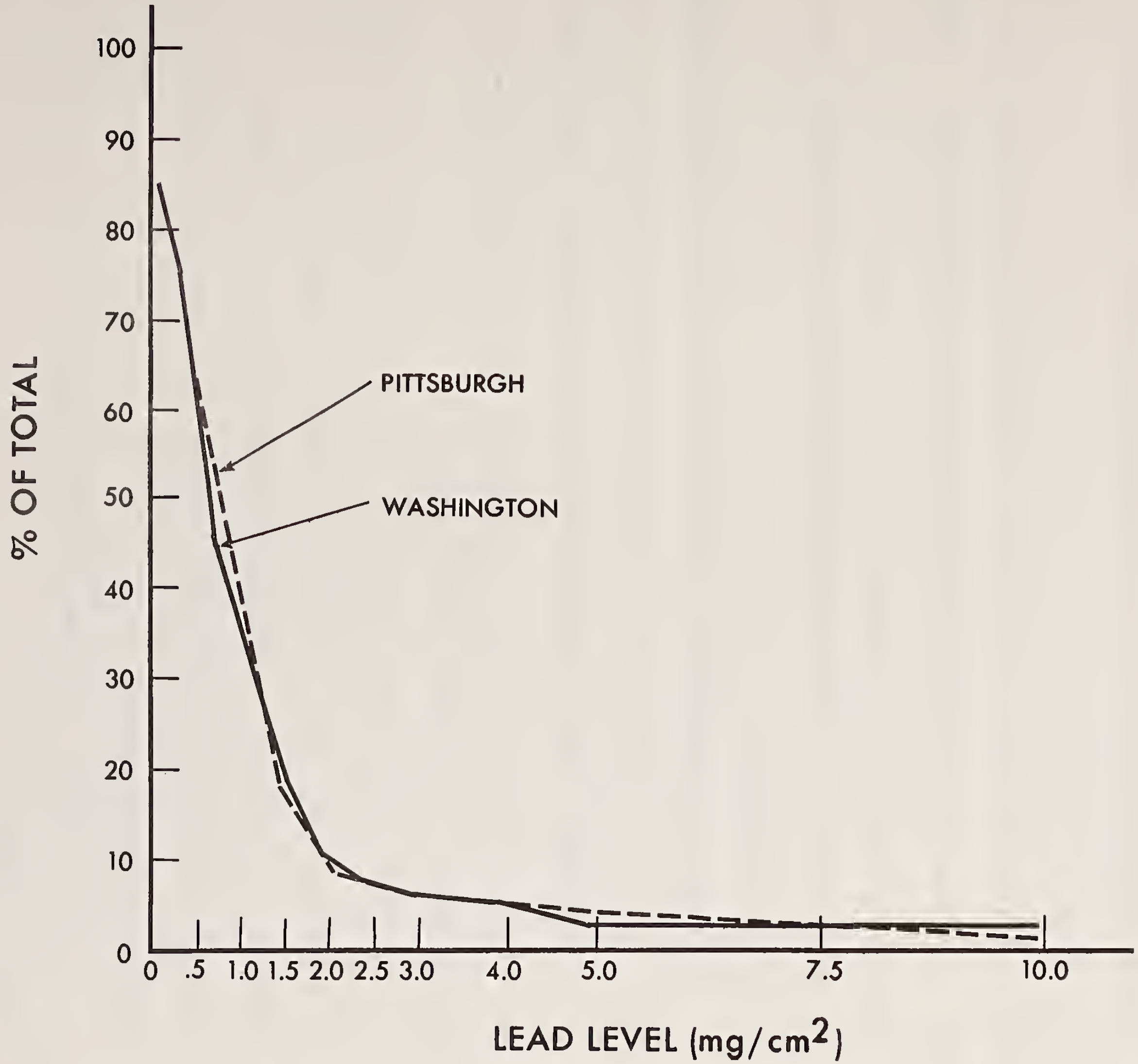


Figure 11D

Cumulative Distributions (Pittsburgh and Washington) for Interior Doors in 1940-1959 Units



follow quite similar shapes. Figures 11A-11B suggest that there is very little difference in the wall lead levels for the two cities; the most pronounced difference occurs for doors in pre 1940 units, as shown in Figure 11C.

4.4 Summary of Results

This section has presented various analyses of the Pittsburgh study data, performed in order to understand better the relation of certain housing characteristics to measured lead levels in the dwelling unit. Further analyses, in addition to those already described here, have also been conducted using data from the Pittsburgh study. A table of contents to all these analyses -- those presented here and in a companion report [8], and those which have not been reported -- can be found in Appendix C.

The major conclusions to be drawn from the analyses given here are now briefly summarized:

1. OCCUPANCY CLASS - No substantial differences in lead levels are observed among the various occupancy classes, except that public housing tends to show lower lead levels.

2. AGE - Measurable lead levels (e.g., in excess of 2.0 mg/cm^2) are present in all dwelling unit age categories, and even in a significant number of newer units. However, lead levels show a pronounced variation by the age of the unit, with the highest lead levels observed in older dwelling units.

3. ROOM TYPE - Wet rooms exhibit considerably greater lead levels than do dry rooms.

4. SURFACE TYPE - Lead levels found on trim (e.g., doors, windows and baseboards) are markedly greater than those found on walls.

5. CONDITION - Lead levels detected in rooms having a bad substrate or a peeling paint film are somewhat greater than those levels detected in rooms having a good substrate and a tightly adhered paint film.

6. LOCATION - Lead levels for exterior surfaces are considerably greater than those found on interior surfaces of a dwelling unit; these exterior lead levels vary consistently by the age category of the unit.

7. COMPARISON WITH WASHINGTON - The levels of lead found on walls are quite similar for both Pittsburgh and Washington, D.C.

In addition, the analyses conducted here have been based on consideration of several summary statistics, rather than on any one single statistical measure. This practice appears warranted for the present data set, inasmuch as pronounced departures from normality are observed in the statistical distributions of lead levels.

5. REFERENCES

[1] Gilsinn, J., "Estimates of the Nature and Extent of Lead Paint Poisoning in the United States," Technical Note 746, National Bureau of Standards, December 1972.

[2] Laurer, G.R., et al., "The Distribution of Lead Paint in New York City Tenement Buildings," American Journal of Public Health, Vol. 63, No. 2, pp. 163-168, 1973.

[3] Jamison, C.S., "Lead Paint Danger to Pre-Schoolers in Houston, Texas," thesis (Master of Public Health), University of Texas, School of Public Health at Houston, June 1971.

[4] McAllister, J.J. and H.H. Tilson, "Multnomah County Oregon Lead Poisoning Project Report," June 1974.

[5] "Childhood Lead Poisoning Control Project for New Mexico," Two Year Summary Report, New Mexico Environmental Improvement Agency, Santa Fe, New Mexico, 1976.

[6] "Childhood Lead Poisoning: A Summary Report of a Survey for Undue Lead Absorption and Lead-Based Paint Hazard in 27 Cities," DHEW Publication No. (HSM) 73-10002, September 1972.

[7] Hall, W., Ayers, T. and D. Doxey, "Survey Plans and Data Collection and Analysis Methodologies: Results of a Pre-Survey for the Magnitude and Extent of the Lead Based Paint Hazard in Housing," NBSIR 74-426, National Bureau of Standards, January 1974.

[8] Urban, W., "Statistical Analysis of Blood Lead Levels of Children Surveyed in Pittsburgh, Pennsylvania: Analytical Methodology and Summary Results," NBSIR 76-1024, National Bureau of Standards, April 1976.

[9] Hall, W. and L. Slovic, "Survey Manual for Estimating the Incidence of Lead Paint in Housing," Technical Note 921, National Bureau of Standards, September 1976.

[10] "Polk Directory for Pittsburgh, Pennsylvania," R.L. Polk and Company, 1974.

[11] "Lead-Based Paint Poisoning," Hearing before the Subcommittee on Health of the Committee on Labor and Public Welfare, United States Senate, 91st Congress, Second Session, November 23, 1970.

[12] New York State Sanitary Code, Part 67 (Lead Poisoning Control), Section 67.3-(b), added 1970.

[13] Godette, M., Post, M. and R. Higgons, "Preparation of Standard Lead Bearing Samples," Report 10631, National Bureau of Standards, November 1971.

[14] Clark, E., Waksman, D. and D. Hattis, "Identification of Potential Building Material Sources of Lead Poisoning," Report 10632, National Bureau of Standards, November 1971.

[15] Rasberry, S., "Investigation of Portable X-ray Fluorescence Analyzers for Determining Lead on Painted Surfaces," Applied Spectroscopy, Vol. 27, No. 2, pp. 102-108, 1973.

[16] Spurgeon, J., "Response Characteristics of a Portable X-ray Fluorescence Lead Detector: Detection of Lead in Paint," NBSIR 73-231, National Bureau of Standards, June 1973.

[17] Kaplan, E.H., et al., "Portable X-ray Fluorescence Instruments for the Analysis of Lead in Paints," Public Health Reports, Vol. 90, pp. 223-230, May-June 1975.

[18] Massachusetts Lead Poisoning Regulations, Department of Public Health, Sections 4.05 and 5.1, 1975.

[19] New Jersey State Sanitary Code, Chapter 13 (Lead Poisoning), Section 2.3, July 1972.

[20] Census of Population and Housing: 1970, Census Tracts, Final Report PHC(1)-162, Pittsburgh, Pa. SMSA, U.S. Bureau of the Census, May 1972.

[21] Johnson, N.L. and S. Kotz, Continuous Univariate Distributions, Vols. 1 and 2, Houghton Mifflin Co., Boston, 1970.

[22] Feller, W., An Introduction to Probability Theory and Its Applications, Vol. I, John Wiley and Sons, New York, 1968.

[23] Filliben, J.J., "Techniques for Tail Length Analysis," Proceedings of the Eighteenth Conference on the Design of Experiments in Army Research Development and Testing, Part 2, ARO Report 73-2, pp. 425-450, 1973.

APPENDIX A

INSTRUCTIONS FOR COMPLETING THE HOUSING SURVEY FORM

The Data Collection Form (NBS-744), hereafter called the DCF, is the sole reporting instrument for the housing portion of the survey. Hence there are several different types of data appearing on the DCF. Some of the data elements will have been entered prior to the inspector's receipt of the DCF; others will be transcribed from the XRF log book; the remainder will be collected on-site for the individual dwelling unit.

In the description of the data elements, the term "field" is used to denote a sequence of characters or digits which are treated as an entity. The designation "DU" is used as an abbreviation for "dwelling unit."

The DCF consists of 19 lines, each pre-numbered in the field CODE. The field SERIAL NUMBER forms the unique identification for the DCF. Generally, if the entity defined by a line of the DCF does not exist (a one story DU, for example, would have no stairway), that line is left completely blank when the data are collected. The position of the relevant (or non-blank) characters within a field is immaterial; e.g., ΔΔ1 is considered identical to Δ1Δ or 1ΔΔ, but the characters must be successive (a field such as 1Δ2 is meaningless whereas Δ12 and 12Δ are meaningful and equivalent)*. Decimal points are ignored in entering XRF readings; a decimal point is automatically generated during

*Here, "Δ" is used to indicate a blank character in a field.

subsequent processing. No particular convention must be observed to avoid ambiguity of l's and I's or O's and o's.

The right hand side (RHS) of the DCF contains a comprehensive list of the codes required in the body of the form.

When the inspection is completed, lines 1, 2, 3, and one or more of the lines from 4 to 19 will have been filled out. If the DU did not exist (resulting from an error in the city directory), only line 1 will be filled out. If the DU exists but could not be inspected, lines 1 and 3 will be filled out.

Section I - IDENTIFICATION

Line 01 - This line is not the responsibility of the inspector. All the fields here except "V" will be filled out before the inspector receives the form. The field "V" will be filled out, if necessary, by the supervisor after the form is returned.

SERIAL NUMBER - The arbitrarily assigned unique identifying number for the dwelling unit. It must be numeric.

TRACT - The census tract in which the DU lies.

BLOCK - The census block in which the DU lies. (This field was not used in the Pittsburgh housing survey.)

ZIP CODE - The postal ZIP CODE of the DU mailing address.

V - The visitation code which is entered after the DCF is returned. If an inspection has been made, this field is left blank; otherwise this field describes the reason the inspection could not be made. The codes used are listed in I - Visitation on the RHS. The text accompanying the codes is self-explanatory; note, however, that codes 3 and 6 are different.

STREET NAME and NUMBER - Self-explanatory. This field is used only by the inspector for locating the DU. This information has been dropped in the subsequent construction of the data base, in keeping with privacy requirements.

Line 02 - These data are collected daily and copied onto the DCF from the XRF log book. This information relates to the day of inspection rather than the specific DU.

XRF SERIAL - The three-character numeric serial number of the particular XRF instrument used.

TEST BLOCK - The XRF reading obtained from the painted panel (medium lead content) test block included with each instrument. This reading is taken after the instrument has been calibrated.

ZERO READING - The reading obtained from the non-leaded wood block included with each instrument. This reading is taken during calibration of the instrument.

ZERO VERNIER - The zero vernier (dial) setting corresponding to the zero reading.

CALIBRATE READING - The reading obtained from the lead-foil-on-wood block included with the instrument. This reading is taken during calibration of the instrument.

CALIBRATE VERNIER - The vernier setting corresponding to the calibrate reading.

PEAK READING - The reading obtained from the lead-foil-on-wood block included with each instrument. This reading is taken during calibration of the instrument.

PEAK VERNIER - The vernier setting corresponding to the peak reading.

INSPECTORS - The initials of both members of the inspecting team.

DATE/MONTH - Conventional numeric code (1-12) for month of year.

DATE/DAY - Conventional numeric code (1-31) for day of month.

Line 03 - This line is completed on-site. The fields YEAR through CHILD require the active participation of the householder.

SEQ - Indicates the sequence in which the DU was inspected that day (first, second, third, etc.). Do not include DU's which were visited but not inspected.

TEST BLOCK- The average of five on-site XRF readings taken on the (medium lead) painted panel test block included with the instrument.

TYPE - Coded according to II--Type of Construction--on the RHS. If the type is mixed, enter the code for the predominant type.

XS - The material of the exterior according to III--Outside Surface of Building--on the RHS. If the surface is mixed, enter the predominant type code.

OC - Coded according to IV--Occupancy--on the RHS. For multi-unit buildings a mailbox count is a convenient way of ascertaining the number of units.

YEAR - Coded according to V--Year Built--on the RHS. This will be obtained from the householder if possible; otherwise it is to be estimated by the inspector.

OWNER - Coded according to VI--Owner/Renter--on the RHS.

MORT - Coded according to VII--Mortgage--on the RHS.

PUBLIC - Coded according to VIII--Public Housing--on the RHS. Public housing denotes those units which are operated by some agency of government whether federal, state, or local.

SUB - The subsidized housing code from IX--Subsidized--on the RHS. Subsidized housing denotes units for which at least a portion of the rent is paid by some federal, state, or local government agency.

CHILD - The number of children, age 6 and under, resident in the DU.

Section II - INTERIOR

This section of the DCF includes lines 4 through 17. Each of these lines has the same format; the pre-printed CODE field identifies the room type.

No attempt is to be made to move furniture, pictures, etc. or to stand on anything in order to obtain readings.

Walls and Ceilings

COND - The material and condition of the walls, ceilings, etc. This information consists of a three-character field in which one character is a material code according to XI (a) of the RHS (V is used to indicate any surface of synthetic or plastic material such as laminated plastic, etc.); one character is a base or substrate condition code according to XI (b) of the RHS and the third character is a surface condition code according to XI (c) of the RHS. The order in which these characters appear is immaterial. The code P2Q, for example, is equivalent to PQ2, 2PQ, or QP2.

These codes are assigned to the room in general, rather than to each specific surface. If the surfaces are mixed, the code which most nearly characterizes the room at a height of four feet or less should be entered.

Base condition codes are described using an estimate of the magnitude of work involved if redecoration were considered:

1. If no work is required.
2. If minor work is required (small cracks, but no deep or large holes).
3. If major work is required (large cracks, deep or large holes, structural defects).

Since COND reflects the general condition of the room, it is best to complete this field after the room inspection has been completed.

WALL 1, WALL 2, WALL 3, WALL 4, and CEILING - Each of these fields contains an XRF reading taken on the indicated surface. The walls are numbered clockwise beginning with the wall to the left of the entry. The wall reading should be taken at any point less than four feet from the floor if such a point is accessible. The ceiling reading is to be taken only if it is possible to do so by standing on the floor or stairs; do not stand on furniture or counter tops, etc. If the wall is inaccessible enter an "X" in the field; if the surface does not exist (as in a three-walled room, for example) leave the reading field blank. If the ceiling is inaccessible, the CEILING field may be left blank instead of entering an "X"; subsequent editing procedures assume that a ceiling exists for each room, and thus a blank and an "X" are equivalent for this particular field.

Trim

COND - Similar to the COND field for walls and ceilings, this field contains a general condition code for all trim within the room (i.e., windows, doors and baseboards).

WINDOW NUMBER - Enter the number of windows in the room ("0" if none).

WINDOW READING - Enter the XRF reading obtained from a window frame or sill; if possible this reading should be taken from a vertical surface within four feet of the floor.

DOOR NUMBER - Enter the number of doors in the room ("0" if none).

DOOR READING - Enter the XRF reading from any point of a door's surface which is within four feet of the floor. If the door is of variable thickness, take this reading from the thick part of the door to avoid the problem of detecting lead on the back surface of the door.

BASEBOARD - Enter the XRF reading from an accessible part of the baseboard from any convenient wall of the room. If the baseboard is inaccessible enter an "X"; if there is no baseboard leave this field blank.

Other

FLOOR - Enter an XRF reading only if the floor is a painted one.

RADIATOR, CABINET, FIREPLACE - Enter XRF readings for each of these which exist; leave the field blank for each non-existent entity.

Line 17 - COND - is a numeric field containing the number of moves the family has made in the preceding six years.

Section III - EXTERIOR

The exterior section differs from the interior section in that there is a condition code associated with each XRF reading taken. Criteria for each condition code are, however, the same as for the interior section. Readings are to be made only if the designated areas are painted. As with the interior readings, the XRF readings should be taken from a point at a height of four feet or less.

WALL - An outside wall which is typical of the predominant exterior surface of the dwelling unit.

PORCH - The floor of the porch.

DOOR - An exterior door. If the door is of variable thickness, this reading is taken from the thick portion.

WINDOW - A window frame or sill.

RAILING - The porch or stair railing.

FENCE - The fence.

GARAGE - The garage wall.

EXCLUDED ROOMS - Enter here the number of rooms of the indicated types which were not inspected.

APPENDIX B

DATA BASE SPECIFICATIONS

There are three files currently maintained which can be made available* to interested researchers. The first, and by far the largest of these, MAST16, contains all housing and child data which were collected. The second file, BLF014, contains all child data and the housing data for those units where child data are present. A third file, BLMI14, contains data for housing units in which there was at least one child for which no child data were collected.

The files have almost identical specifications; BLF014 and BLMI14 are actually subfiles which were extracted from MAST16. All files are completely "processible"; that is, all records are of proper length and format, and all data fields are of the indicated form and have values within the specified range.

File Format and Structure

The files have no labels and each is terminated by a file mark preceded by a sentinel record; the sentinel is a record with "999999" in the first field. The file elements are basically images of the lines present on the data collection form (DCF) with the serial number from line 1 replacing the leading blanks from lines 2, 3, 16, 17 and the indicated room names from the other lines. In addition a "pointer" is used to indicate the successor line type, and an image of the psycho-socio-environmental (PSE) form is included where available for each child.

---

*Data tapes for these files can be obtained from the Lead-Based Paint Program, Housing & Urban Development, Washington, D.C. 20410.

These lines are sorted in ascending order* on a two-part key consisting of the serial number and the line type; the line type indicates the room type for the housing data. The formatted PSE line for each child is assigned a line number from 20 - 29: 20 for the first child tested, 21 for the second child tested, up to 27 for the eighth child tested. These child line numbers were assigned as the data were received and are not in order by age or sex. The comments by the blood collection team (a portion of the PSE form without fixed formats) are included as lines 30 - 39 in free form text.

In each of the files, a sequence of lines having the same serial number and including all lines present through line 19 defines a dwelling unit. Line numbers which are 20 or greater indicate that child data have been included.

Line Formats

The description of the line contents and formats are keyed to the definitions and formats appearing on the DCF. For each of the files there are different acceptable combinations of lines.

---

*Some arithmetic is involved to transform the serial numbers as they exist in the file into true numeric sort. The serial numbers were assigned such that the first four digits of the serial number correspond to the last four digits of the census tract in which the unit is located and the last two digits were index numbers. More than the 99 units thus allowed for were inspected in some census tracts. The indices for units 100-199 in these tracts were formed by adding 5 to the leading position of the serial number and using the last two digits of the index in the index portion of the serial number. This causes no ambiguity in identification since the only permissible characters in the first position of the serial number are 0-4. To illustrate the method for sorting the serial numbers into order, let each character of the serial number be denoted by the letters a through f; then a-b-c-d-e-f is a typical input serial number. If a>4, set g=a-5 and h=1; otherwise, set g=a and h=0. Then the data can be placed in ascending order by sorting on the key g-b-c-d-h-e-f.

For MAST16, acceptable sequences are:

(1) Line 1 only (if the dwelling unit does not exist).

(2) Lines 1 and 3 only (if the dwelling unit exists, was observable but the inspection was not performed).

(3) Lines 1, 2, 3 and at least one line from among lines 4 - 19. A line for each PSE form submitted is included; these are numbered sequentially beginning with 20 for the first child. If at least one line from among 20 - 29 appears, there may be one or more comment lines beginning at line 30 and terminating (after a sequence of consecutive line numbers) at any line through line 39. The totality of these lines represent dwelling units on which inspections were performed.

For BLF014, the acceptable sequences are the same as for (3) of MAST16, except that exactly one line from among lines 20 - 29 appears. In BLF014, there are duplicate dwelling units; the housing data are duplicated for each child after the first. The criterion for generation into this subfile is that a file entry was extracted from MAST16 for each child about whom PSE data were collected.

BLMI14 has the same acceptable sequences as (3) for MAST16. The criterion for generation into this subfile is that an entry was extracted from MAST16 for each dwelling in which the number of children represented by PSE data was less than the number of children indicated on the DCF.

The format of each of the possible line types is identical for each of the three files. Each field is right-adjusted. The format and definitions of variables are given below on a line-by-line basis keyed to the DCF. In what follows, RHS is used to denote the right hand side of the DCF, where definitions for the codes appear.

LINE 1 - Identification Line - FORTRAN Format (I6, I2, I6, 3X, I5, I1, 2I6, 4A6, 2I2).

Field 1 - DU serial number (I6).

Field 2 - Line type number (I2), set equal to 1.

Field 3 - Census tract (I6); either greater than 30100 and less than 33299, or greater than 54810 and less than 54814.

Field 4 - Zip code (I5); if not reported, this field contains 0.

Field 5 - Visit code (I1) as defined in RHS-I. This has the value 0 if the inspection was performed.

Field 6 - Census tract (I6).

Field 7 - DU serial number within tract (I6).

Field 8 - Δ*Δ*Δ* (A6).

Field 9 - Δ*Δ*Δ* (A6).

Field 10 - Δ*Δ*Δ* (A6).

Field 11 - Δ*Δ*Δ* (A6).

Field 12 - Batch number (I2) - This field was used as part of the quality control procedures in the creation and updating of the file. It is the index of the last cycle in which the line was changed. If no changes have been made, it is the index of the cycle in which the line was inserted into the file. Batch number appears in every line of each file.

Field 13 - Successor index (I2) - This is simply the index of the next line type which appears in the file. For the last line of the file this index is 99. This field appears in every line of the file.

LINE 2 - XRF Calibration and Identification Line - FORTRAN Format (I6, I2, 8I3, 2A3, 2I3, 15X, 2I2).

Field 1 - DU serial number (I6).

Field 2 - Line type number (I2), set equal to 2.

Field 3 - XRF serial number (I3).

Field 4 - Test block reading (I3) - The XRF reading for the medium lead test block, taken during the daily calibration process.

Field 5 - Zero reading (I3) - The XRF reading from the non-lead test block, taken during the daily calibration process.

Field 6 - Zero vernier (I3) - The setting of the zero vernier scale after calibration.

Field 7 - Calibration reading (I3) - The XRF reading from the lead block taken during the daily calibration process.

Field 8 - Calibrate vernier (I3) - The setting of the calibrate vernier scale.

Field 9 - Peak reading (I3) - The XRF reading for the lead block taken during the daily calibration process.

Field 10 - Peak vernier (I3) - The setting of the peak vernier scale.

Field 11 - Initials of the first inspector (A3).

Field 12 - Initials of the second inspector (A3).

Field 13 - Month code (I3): i.e., 1 - 12.

Field 14 - Day of month (I3): i.e., 1 - 31.

Field 15 - Batch number (I2).

Field 16 - Successor index (I2).

LINE 3 - Questionnaire Line - FORTRAN Format (I6, I2, 6I3, 4A3, I3, 18X, 2I2).

Field 1 - DU serial number (I6).

Field 2 - Line type number (I2), set equal to 3.

Field 3 - Sequence (I3) - The index of this inspection within the day.

Field 4 - Test block (I3) - The XRF reading from the medium lead test block, taken at the time of inspection.

Field 5 - Type (I3) - The type of construction, coded according to RHS-II. An unknown type is denoted by 4.

Field 6 - Exterior surface (I3) - The material of the exterior surface, coded according to RHS-III. Unknown is denoted by 8.

Field 7 - Occupancy class (I3) - Occupancy class of the unit, coded according to RHS-IV. Unknown is denoted by 8.

Field 8 - Year built (I3) - Year the structure was built, coded according to RHS-V. Unknown is denoted by 4.

Field 9 - Owner/Renter (A3) - The code for owner/renter category according to RHS-VI. Unknown is denoted by U.

Field 10 - Mortgage type (A3) - Mortgage type, coded according to RHS-VII. Unknown is denoted by U.

Field 11 - Public housing (A3) - Coded according to RHS-VIII.

Field 12 - Subsidized (A3) - Always N in this survey.

Field 13 - Children resident (I3) - This is a single digit specifying the number of children aged six or less, except that 9 denotes an unknown number of children.

Field 14 - Batch number (I2).

Field 15 - Successor index (I2).

LINES 4 - 16 - Room Lines - FORTRAN Format (I6, I2, A3, 5I3, A3, 9I3, 3X, 2I2).

Each line present denotes a room of the specified type. Only those rooms for which one or more XRF readings were taken appear in the file. All numeric fields representing XRF readings are given in tenths of mg/cm^2; a numeric value greater than 990 indicates that no reading was taken.

Field 1 - DU serial number (I6).

Field 2 - Line (room) type (I2); a room type = 16 indicates an unknown room type.

Field 3 - Wall condition code (A3) - This refers to the representative condition of the walls and ceilings in the room.

Character 1 - The wall material according to RHS-XI(a).

Character 2 - The wall substrate condition according to RHS-XI(b).

Character 3 - The wall surface condition according to RHS-XI(c).

Fields 4-8 - XRF readings (5I3) - XRF readings for each of the four walls and the ceiling.

Field 9 - Trim condition code (A3) - Exactly as in Field 3 except that it describes the condition of the trim surfaces within the room.

Field 10 - Number of windows (I3) - The number of windows within the room.

Field 11 - Window reading (I3) - The XRF reading taken from a window within the room.

Fields 12 & 13 - Number of doors/door reading (2I3) - Exactly as given in Fields 10 and 11, except these fields refer to the doors of the room.

Fields 14 - 18 - XRF readings (5I3) - XRF readings from a baseboard, floor, radiator, cabinet, and fireplace within the room.

Field 19 - Batch number (I2).

Field 20 - Successor index (I2).

<u>LINE 17</u> - Move Line - FORTRAN Format (I6, I2, I3, 48X, 2I2).

Field 1 - DU serial number (I6).

Field 2 - Line type (I2), set equal to 17.

Field 3 - Moves (I3) - The number of moves the family has made within the preceding six years.

Field 4 - Batch number (I2).

Field 5 - Successor index (I2).

<u>LINES 18 & 19</u> - Exterior Lines - FORTRAN Format (I6, I2, 7(A3,I3), 3I3, 2I2).

Field 1 - DU serial number (I6).

Field 2 - Line type number (I2).

Field 3 - Wall condition (A3) - Like Field 3 of the interior section, but here refers to the exterior wall.

Field 4 - Wall reading (I3) - XRF reading for the wall.

Fields 5 & 6 - (A3, I3) - Like Fields 3 and 4, for the porch floor.

Fields 7 & 8 - (A3, I3) - Like Fields 3 and 4, for exterior door.

Fields 9 & 10 - (A3, I3) - Like Fields 3 and 4, for exterior window sill or frame.

Fields 11 & 12 - (A3, I3) - Like Fields 3 and 4, for porch railings.

Fields 13 & 14 - (A3, I3) - Like Fields 3 and 4, for fence.

Fields 15 & 16 - (A3, I3) - Like Fields 3 and 4, for a painted surface of the garage. (Line 18 only.)

Fields 17 - 19 - (3I3) - The number of bedrooms, bathrooms, and rooms other than bedrooms or bathrooms which were not inspected. (Line 18 only.)

Field 20 - Batch number (I2); this will be found in Field 15 for line 19.

Field 21 - Successor index (I2); this will be found in Field 16 for line 19.

<u>LINES 20 - 29</u> - Formatted PSE Line - One line appears for each child from whom blood was collected. FORTRAN Format (I6, I2, A2, 7I2, 2A1, 2I2, I1, A1, I5, A1, I1, I2, I5, 2I2, I5, I2, 2I1, 14X/5I1, 2I2, 2A5, 2I1, 2I2).

Field 1 - Serial number (I6).

Field 2 - Child number + 19 (I2); thus child 1 is denoted by

20, child 2 is denoted by 21 etc.

Field 3 - Inspector's identification (A2).

Field 4 - Month of inspection (I2).

Field 5 - Day of month (I2).

Field 6 - Blood lead level (I2), initial screening micro.

Field 7 - Age of child in months (I2); "99" indicates 99 months or older.

Field 8 - Rank of child in family (I2).

Field 9 - Height of child (inches) (I2).

Field 10 - Weight of child (pounds) (I2).

Field 11 - Sex (A1); (M=Male; F=Female).

Field 12 - Race (A1); (W=White; B=Black; O=Oriental).

Field 13 - Months child has lived in present dwelling unit (I2).

Field 14 - Months child has lived in present census tract (I2).

Field 15 - Percent of waking hours child spends in his unit (I1):

1. 0 - 25%
2. 25% - 50%
3. 50% - 75%
4. 75% - 100%

Field 16 - More than 25% of waking hours spent at friends or relatives? (A1). Y=Yes; N=No.

Field 17 - Friend's or relative's census tract (I5) if Field 16 indicates Yes.

Field 18 - Does child attend preschool? (A1). Y=Yes; N=No.

Field 19 - Preschool days/week (I1).

Field 20 - Preschool hours/day (I2); in tenths of hours.

Field 21 - Preschool census tract (I5).

Fields 22 & 23 - Favorite play areas (2I2):

1. Grass
2. Dirt
3. Pavement

Field 24 - Play area census tract (I5).

Field 25 - Indoor play area (I2); room number used on DCF.

Field 26 - Pica evidence (I1):

1. No
2. Possible
3. Confirmed
4. Other

Field 27 - Number of siblings six or under (I1).

Field 28 - Number of siblings 7 to 11 (I1).

Field 29 - Number of siblings 12 or over (I1).

Field 30 - Family income (I1):

1. 0 - 3,999
2. 4,000 - 6,999
3. 7,000 - 10,999
4. 11,000 or more
5. Cares not to answer
6. None of the above

Field 31 - Work situation (I1):

0. Neither works
1. Mother works
2. Father works
3. Both work

Field 32 - Child care (I1):

0. Not answered

1. Older sibling

2. Adult non-relative

3. Adult relative

4. Day care

Field 33 - Years of education of mother (I2).

Field 34 - Years of education of father (I2).

Field 35 - Mother's origin (A5); a five-character code defined as follows:

Character 1:

0 = Pennsylvania, but not Allegheny County

1 = Allegheny County

Δ = Not Pennsylvania

Characters 2 - 3:

ΔΔ = Non-U.S.

otherwise = Two-letter state postal code

Characters 4 - 5:

ΔΔ = If state code in 2 - 3

otherwise = Two-letter country of origin abbreviation

Field 36 - Father's origin code (A5); as in Field 35.

Field 37 - Number of parents resident (I1).

Field 38 - Number of adults resident (I1).

Field 39 - Batch number (I2).

Field 40 - Successor index (I2).

LINES 30 - 39 - Blood Collector's Comments - FORTRAN Format (I6, I2, 8A6, A3, 14X/3A6, A3, 2I2).

Field 1 - Serial number (I6).

Field 2 - Line number (I2).

Fields 3 - 15 - Free form text (8A6, A3, 3A6, A3).

Field 16 - Batch number (I2).

Field 17 - Successor index (I2).

APPENDIX C

TABLE OF CONTENTS TO DATA ANALYSES PERFORMED

The data analyses discussed in this report represent selected results from a considerably wider class of analyses that were in fact performed using the Pittsburgh survey data. The object of this appendix is to provide a brief summary of the types of tabulations and statistical analyses that have already been conducted. Such analyses are grouped together under the following headings:

I. HOUSING - investigations of housing-related variables and measured lead levels.

II. CHILD - preliminary investigations of the blood data and certain PSE information.

III. HOUSING/CHILD - investigations of child data (especially blood lead levels) in relation to housing variables.

IV. SAMPLING - investigations of sampled dwelling unit characteristics, and comparison of dwelling units containing tested/non-tested children.

Certain analyses have been presented in this housing analysis report (Parts I, II) or in a companion blood analysis report [8]. Such analyses are indicated below by the designations (H) and (B), respectively. Other analyses have been performed using the Pittsburgh data but not reported, and these are indicated below by (U), for "unpublished."

I. HOUSING ANALYSES

(H) A. Multidimensional tabulations of highest lead reading, by

- occupancy class
- room type

- lead level
- surface type
- DU age category
- surface condition
- level of aggregation (dwelling, room, surface)

(H) B. Multidimensional tabulations of different room-based statistics, giving their mean and standard deviation, by

- occupancy class
- room type
- lead interval
- surface type
- DU age category
- surface condition
- statistic (mean, standard deviation, median, high)

(U) C. Wall lead levels and census tract variables, plotted in relation to one another. Census tract variables include:

- fraction of population under the age of 7
- education level
- income
- fraction of households below poverty level
- value of dwelling
- housing age
- fraction of non-rent income

(H) D. Statistical analyses of lead level distributions by

- DU age category
- surface type

- surface condition
- statistic (mean, high DU reading)

and using probability plots for the following distributions:

- normal
- lognormal
- Weibull
- extreme value
- uniform
- Cauchy
- Tukey's λ

II. CHILD ANALYSES

(B) A. Summary statistics for blood lead levels, by

- age of child
- sex of child

(U) B. Distributions and probability plots for blood lead levels, by

- age of child
- sex of child

(U) C. Correlation matrices for combinations of

- age of child
- sex of child
- blood lead level

III. HOUSING/CHILD ANALYSES

(U) A. Blood lead levels plotted against census tract variables

(B) B. Summary statistics for blood lead levels, by

- DU age category
- age of child
- sex of child

(U) C. Graphical plots and correlation matrices for combinations of

- blood lead levels
- DU age category
- age of child
- sex of child

(U) D. Probability plots for blood lead levels, by

- DU age category
- age of child
- sex of child

(U) E. Correlation matrices for combinations of

- blood lead levels
- room type
- surface type
- surface condition
- surface lead level statistics

  fraction of surfaces exceeding fixed levels

  mean

  median

  high

  standard deviation

(U) F. Blood lead level distributions (frequency/cumulative) by

- DU age category
- age of child

- sex of child
- fraction of surfaces exceeding certain fixed levels

(H) G. Multidimensional tabulations of different room-based statistics, giving their mean and standard deviation, by

- occupancy class
- room type
- blood lead interval
- surface type
- DU age category
- surface condition
- statistic (mean, standard deviation, median, high)

(B) H. Distributions and summary statistics for blood lead levels, by

- fraction of surfaces exceeding fixed levels
- surface condition

(B) I. Summary statistics for blood lead levels, by

- DU age category
- maximum surface lead level

IV. SAMPLING ANALYSES

(H) A. Comparison of sample characteristics with 1970 census data, by

- census tract counts
- dwelling unit counts

(B) B. Comparison of dwelling units containing children tested, children not tested, and all sampled dwelling units, by

- occupancy class
- surface type

- room type
- level of aggregation (room, surface)
- statistic (mean, standard deviation)

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBSIR 77-1250	2. Gov't Accession No.	3. Recipient's Accession No.

4. TITLE AND SUBTITLE

ANALYSIS OF HOUSING DATA COLLECTED IN A LEAD-BASED PAINT SURVEY IN PITTSBURGH, PENNSYLVANIA: PART I

5. Publication Date March 1977

6. Performing Organization Code

7. AUTHOR(S) Douglas R. Shier and William G. Hall

8. Performing Organ. Report No.

9. PERFORMING ORGANIZATION NAME AND ADDRESS

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

10. Project/Task/Work Unit No.

11. Contract/Grant No.

12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)*

Office of Policy Development and Research
Department of Housing and Urban Development
Washington, D.C. 20410

13. Type of Report & Period Covered Final

14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

This report is a companion document to a previous report (NBSIR 76-1024) on blood lead levels of children tested during a lead-based paint survey in Pittsburgh, Pennsylvania. The emphasis in this report is on the methodology used and types of housing-related information collected by the survey. Through the use of portable x-ray fluorescence lead detectors, measurements were taken from a variety of surfaces within rooms of the dwelling unit as well as at locations exterior to the unit. Analyses of these x-ray fluorescence measurements established that older housing units exhibit considerably greater lead levels than newer housing units. In addition, wet rooms (kitchen and bathrooms) have higher levels than other (dry) rooms, rooms with a poor surface/substrate condition have higher levels than rooms with a good surface/substrate condition, and trim surfaces (e.g., doors, windows, baseboards) have higher levels than walls. Also, exterior surfaces show higher readings than functionally similar interior surfaces. While the present report concentrates on the housing aspect of the survey, subsequent processing of the Pittsburgh data is under way to determine possible relationships among blood lead levels, socioeconomic variables and housing-related characteristics.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

Children; data analysis; housing; lead paint; lead poisoning; surveys; x-ray fluorescence.

18. AVAILABILITY [X] Unlimited

[ ] For Official Distribution. Do Not Release to NTIS

[ ] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13

[X] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151

19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
UNCLASSIFIED	100
20. SECURITY CLASS (THIS PAGE)	**22. Price**
UNCLASSIFIED	$5.00

USCOMM-DC 29042-P74

☆ U. S. GOVERNMENT PRINTING OFFICE : 1977--240-848/192