HOW TO MAKE A BOX PLOT

1. SORT DATA

2. FIND MEDIAN (Point that divides the observations in half -- same number above and below)

n odd : median is the middle observation

Formula : (n+1)/2 (If n = 7, median is the (7+1)/2 = 4th largest observation)

n even : There is no middle observation. Median is between two. Put it half way. If n = 8, formula gives 9/2 = 4.5. Median is half way from the 4th observation to the 5th--the average of the 4th and 5th observations.

3. FIND HINGES (Edges of box)

Formula : (n+1)/4 -- Round up. If n = 20, (20+1)/4 = 5.25 -- Round up to 6. Count up from smallest, down from largest. Lower hinge is 6th from bottom; upper is 6th from top.

4. FIND FENCES

(a) Find distance between hinges (the "H-spread")
(b) Multiply by 1.5.
(c) Add to upper hinge to get upper fence, subtract from lower hinge to get lower fence.

5. DRAW

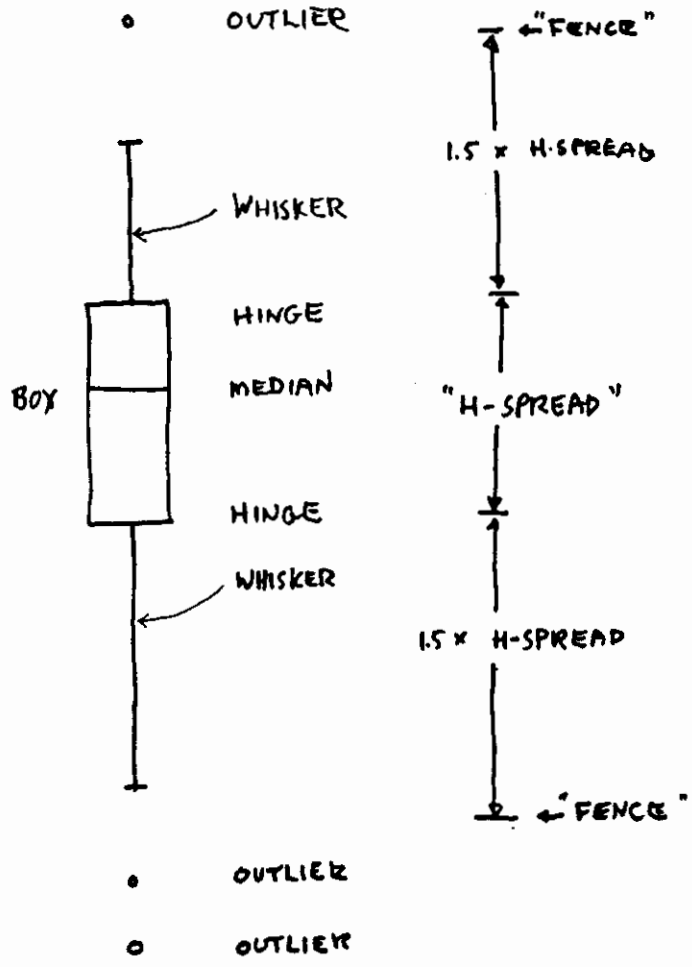
(a) Draw box with median line across it.
(b) Find fences. Go back towards box until you reach an observation.
(c) Draw "whisker" from that observation to box[2].
(d) Add outliers.

6. IDENTIFY OUTLIERS, EXTREMES, POINTS OF SPECIAL INTEREST

7. LOOK, THINK, ...ASK WHY

---

[2]The whisker ends at the observation, not at the fence.



Outliers

Outliers are points that fall so far from the rest that they deserve special attention.

Why are they so extreme ?
Are they correct ?
Are they special in some way?

Maybe they should be excluded from generalizations about this data set.

Maybe they should be treated differently from the rest, or set aside and given a separate analysis.

Male Life Expectancy : North America
(1993 Demographic Yearbook)

country	m_life
Anguilla	-
Antigua and Barbuda	-
Aruba	68.30
Bahamas	68.32
Barbados	67.15
Belize	69.95
Bermuda	70.23
British Virgin I.	-
Canada	73.02
Cayman Islands	-
Costa Rica	72.89
Cuba	72.74
Dominica	-
Dominican Republic	67.63
El Salvador	50.74
Greenland	60.40
Grenada	-
Guadeloupe	66.40
Guatemala	55.11
Haiti	54.95
Honduras	65.43
Jamaica	71.41
Martinique	67.00
Mexico	62.10
Montserrat	-
Netherlands Antilles	71.13
Nicaragua	64.80
Panama	69.78
Puerto Rico	69.60
Saint Kitts & Nevis	65.87
Saint Lucie	68.00
St.Pierre & Miquelon	-
St.Vincent & G'dines	-
Trinidad and Tobago	66.88
Turks and Caicos I.	-
U.S. Virgin Islands	-
United States	72.00

mlif_na1

Male Life Expectancy : North America
(1993 Demographic Yearbook)

	country	m_life
1.	Canada	73.02
2.	Costa Rica	72.89
3.	Cuba	72.74
4.	United States	72.00
5.	Jamaica	71.41
6.	Netherlands Antilles	71.13
7.	Bermuda	70.23
8.	Belize	69.95
9.	Panama	69.78
10.	Puerto Rico	69.60
11.	Bahamas	68.32
12.	Aruba	68.30
13.	Saint Lucie	68.00
14.	Dominican Republic	67.63
15.	Barbados	67.15
16.	Martinique	67.00
17.	Trinidad and Tobago	66.88
18.	Guadeloupe	66.40
19.	Saint Kitts & Nevis	65.87
20.	Honduras	65.43
21.	Nicaragua	64.80
22.	Mexico	62.10
23.	Greenland	60.40
24.	Guatemala	55.11
25.	Haiti	54.95
26.	El Salvador	50.74
27.	U.S. Virgin Islands	-
28.	Grenada	-
29.	St.Vincent & G'dines	-
30.	Dominica	-
31.	Turks and Caicos I.	-
32.	British Virgin I.	-
33.	Cayman Islands	-
34.	Antigua and Barbuda	-
35.	Anguilla	-
36.	St.Pierre & Miquelon	-
37.	Montserrat	-

mlif_na2

Male Life Expectancy : Africa
(1993 Demographic Yearbook)

	country	m_life
1.	Algeria	65.75
2.	Angola	44.90
3.	Benin	45.92
4.	Botswana	52.32
5.	Burkino Faso	45.84
6.	Burundi	48.42
7.	Cameroon	54.50
8.	Cape Verde	63.53
9.	Cent.African Rep.	46.87
10.	Chad	45.93
11.	Comoros	55.50
12.	Congo	48.91
13.	Cote d'Ivoire	49.69
14.	Djibouti	46.72
15.	Egypt	62.86
16.	Eq.Guinea	44.86
17.	Eritrea	48.85
18.	Ethiopia	45.93
19.	Gabon	51.86
20.	Gambia	43.41
21.	Ghana	54.22
22.	Guinea	44.00
23.	Guinea-Bissau	41.92
24.	Kenya	54.18
25.	Lesotho	58.00
26.	Liberia	45.80
27.	Libyan Arab J.	61.58
28.	Madagascar	55.00
29.	Malawi	43.51
30.	Mali	55.24
31.	Mauritania	49.90
32.	Mauritius	65.57
33.	Morocco	61.58
34.	Mozambique	44.88
35.	Namibia	57.50
36.	Niger	44.90
37.	Nigeria	48.81
38.	Reunion	69.38
39.	Rwanda	45.10
40.	Sao Tome-et-Principe	.
41.	Senegal	48.30
42.	Seychelles	65.26
43.	Sierra Leone	37.47
44.	Somalia	45.41
45.	South Africa	60.01
46.	St.Helena	.
47.	Sudan	51.58
48.	Swaziland	42.90
49.	Tanzania	47.00
50.	Togo	53.23
51.	Tunisia	66.85
52.	Uganda	43.57
53.	Zaire	50.40
54.	Zambia	50.70
55.	Zimbabwe	52.39

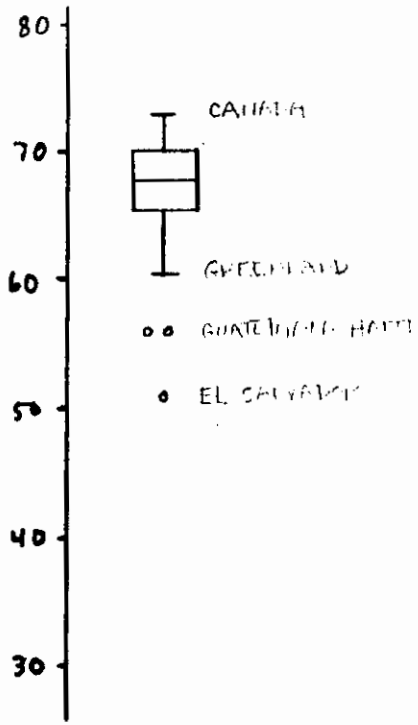
mlif_af1

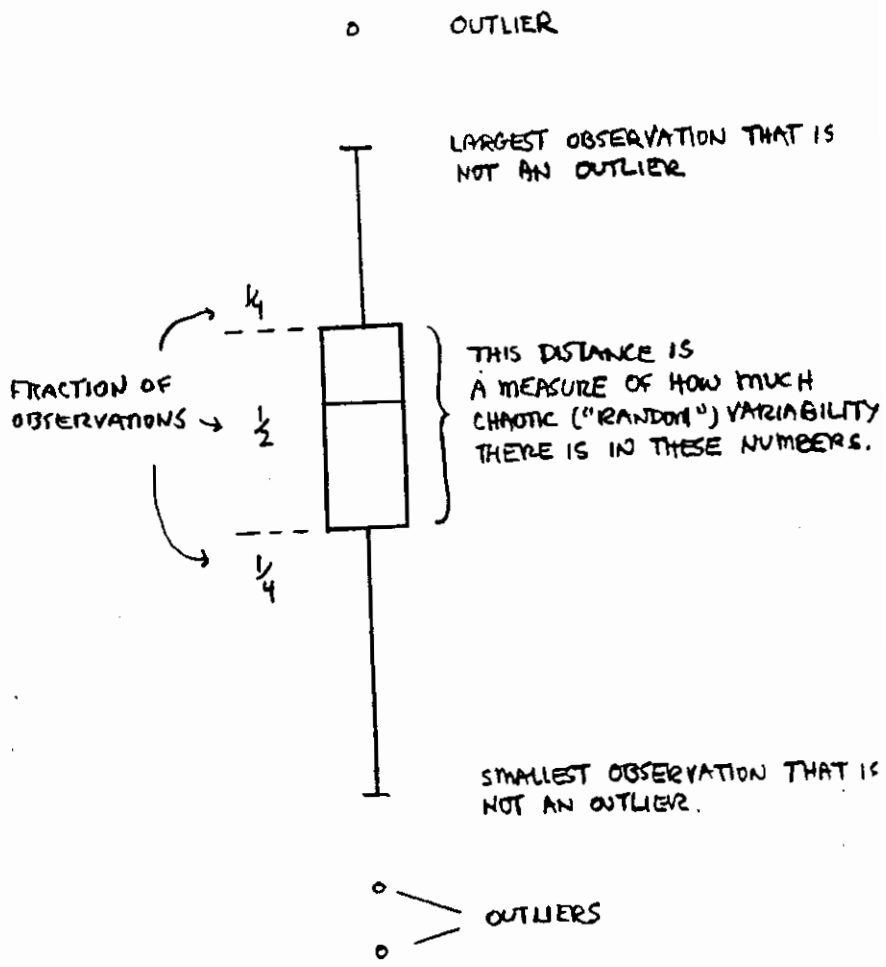
Male Life Expectancy : Africa
(1993 Demographic Yearbook)

	country	m_life
1.	Sierra Leone	37.47
2.	Guinea-Bissau	41.92
3.	Swaziland	42.90
4.	Gambia	43.41
5.	Malawi	43.51
6.	Uganda	43.57
7.	Guinea	44.00
8.	Eq.Guinea	44.86
9.	Mozambique	44.88
10.	Angola	44.90
11.	Niger	44.90
12.	Rwanda	45.10
13.	Somalia	45.41
14.	Liberia	45.80
15.	Burkino Faso	45.84
16.	Benin	45.92
17.	Ethiopia	45.93
18.	Chad	45.93
19.	Djibouti	46.72
20.	Cent.African Rep.	46.87
21.	Tanzania	47.00
22.	Senegal	48.30
23.	Burundi	48.42
24.	Nigeria	48.81
25.	Eritrea	48.85
26.	Congo	48.91
27.	Cote d'Ivoire	49.69
28.	Mauritania	49.90
29.	Zaire	50.40
30.	Zambia	50.70
31.	Sudan	51.58
32.	Gabon	51.86
33.	Botswana	52.32
34.	Zimbabwe	52.39
35.	Togo	53.23
36.	Kenya	54.18
37.	Ghana	54.22
38.	Cameroon	54.50
39.	Madagascar	55.00
40.	Mali	55.24
41.	Comoros	55.50
42.	Namibia	57.50
43.	Lesotho	58.00
44.	South Africa	60.01
45.	Morocco	61.58
46	Libyan Arab J.	61.58
47.	Egypt	62.86
48.	Cape Verde	63.53
49.	Seychelles	65.26
50.	Mauritius	65.57
51.	Algeria	65.75
52	Tunisia	66.85
53.	Reunion	69.38
54.	Sao Tome-et-Principe	.
55.	St.Helena	.

mlif_af2

Box Plot
Male Life Expectancy : North America
(Demographic Yearbook, 1993)





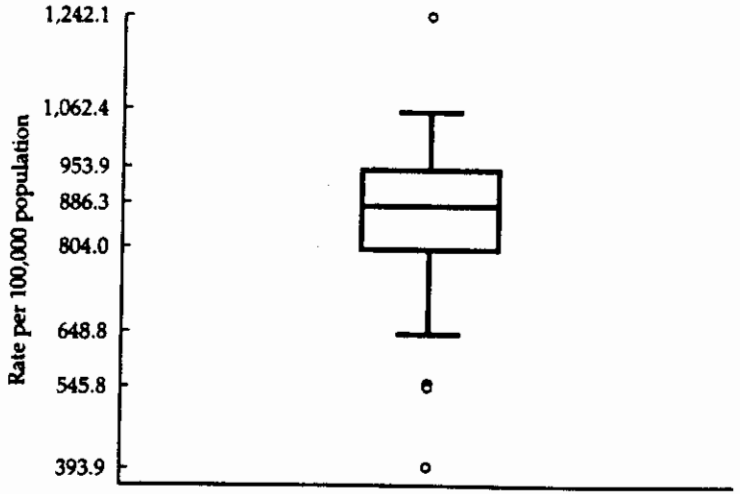
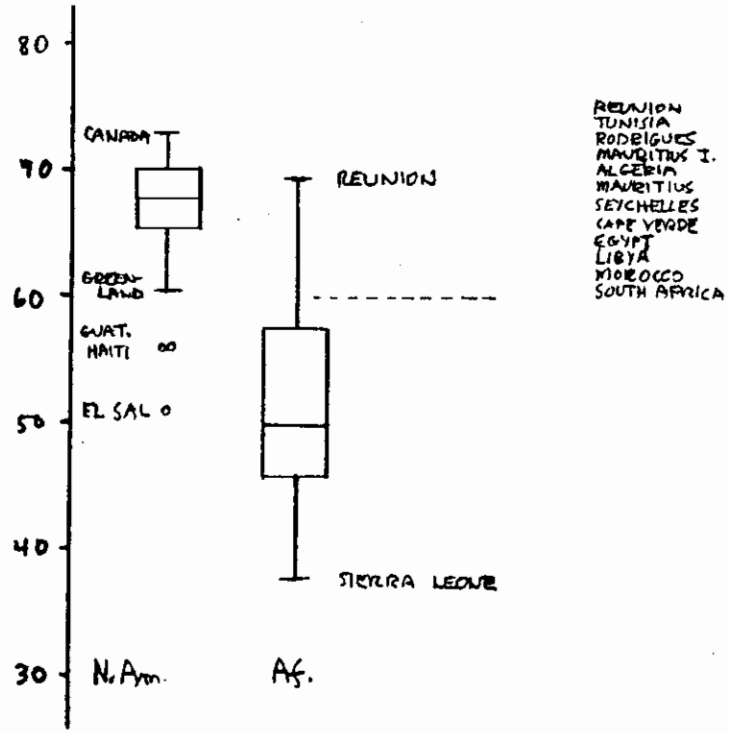
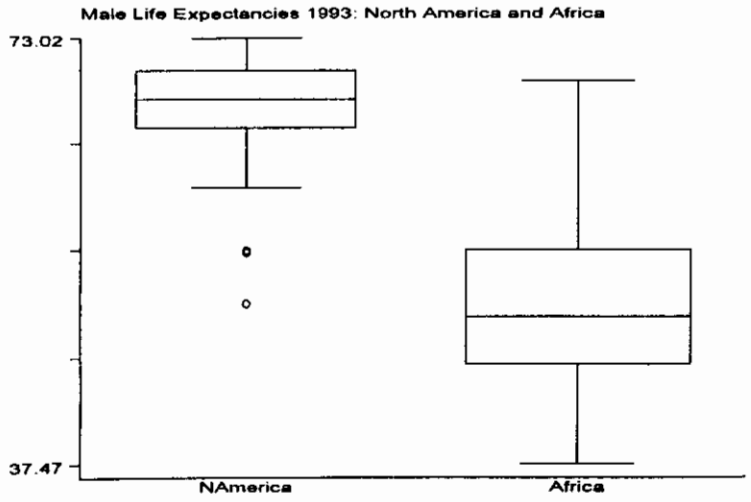
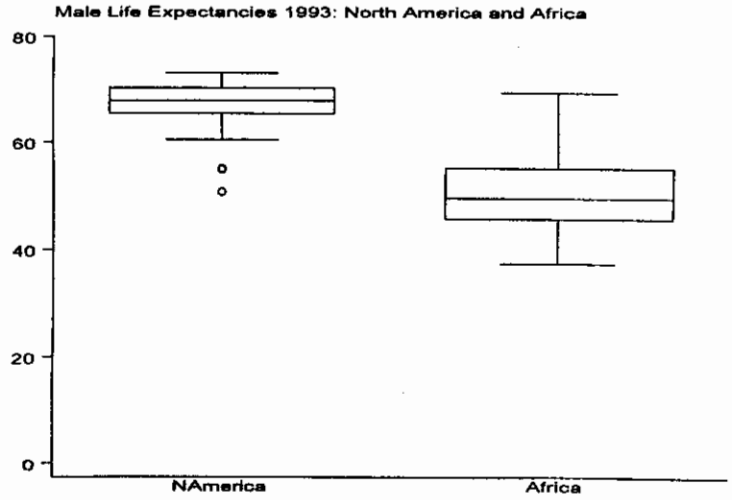


Figure 2.8. Box plot: Crude death rates for the United States, 1988

Box Plots
Male Life Expectancy : North America and Africa
(Demographic Yearbook, 1993)

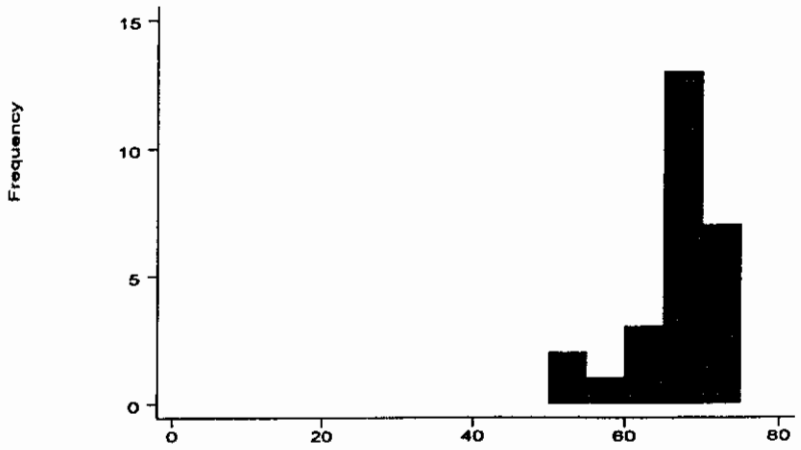



Male Life Expectancies 1993: North America and Africa
73.02
37.47
NAmerica
Africa


Male Life Expectancies 1993: North America and Africa
80
60
40
20
0
NAmerica
Africa

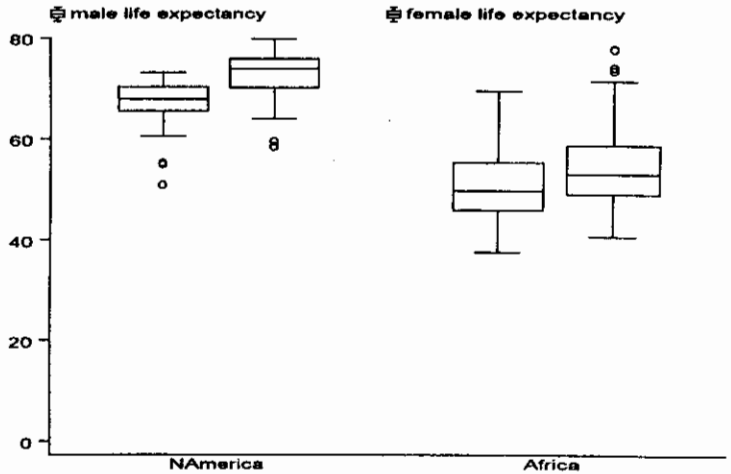
Two Histograms

Male Life Expectancy: Countries of Africa



Male Life Expectancy: Countries of North America

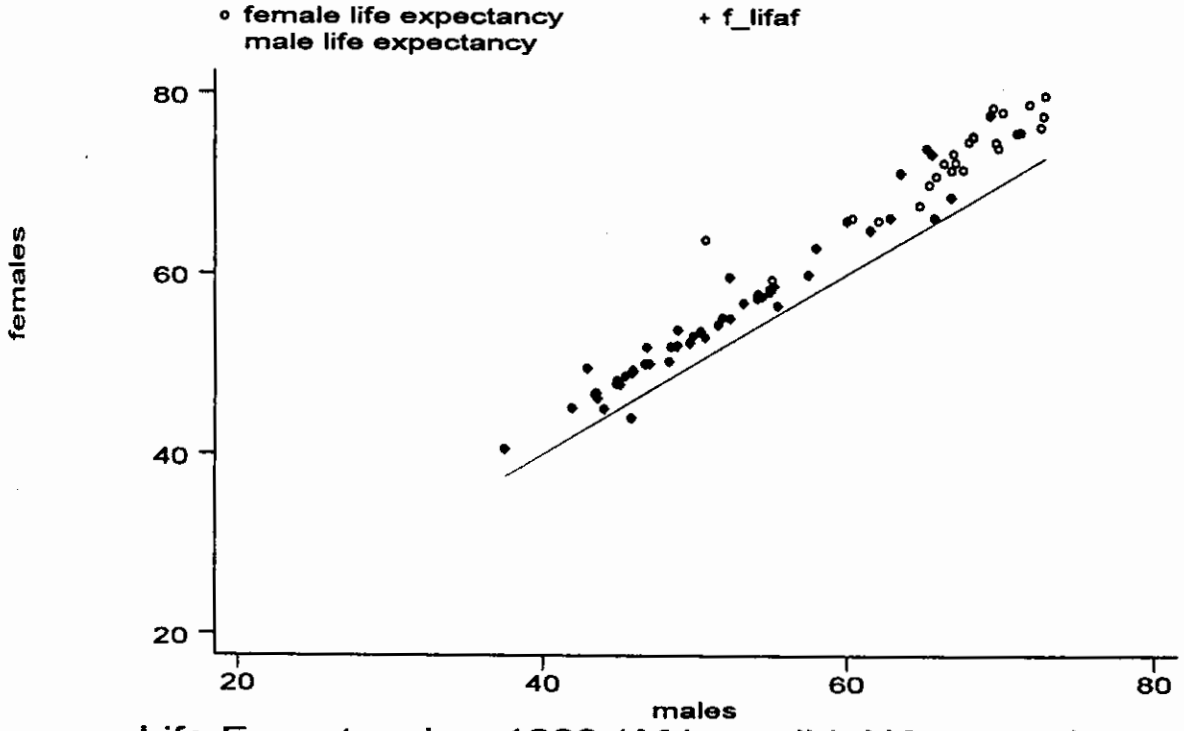
Multiple Box Plots



Life Expectancies in N.America and Africa (by Sex)

Scatter Plot

Life Expectancies in Africa and N.America : 1993



Life Expectancies: 1993 (Africa solid, NAm open)

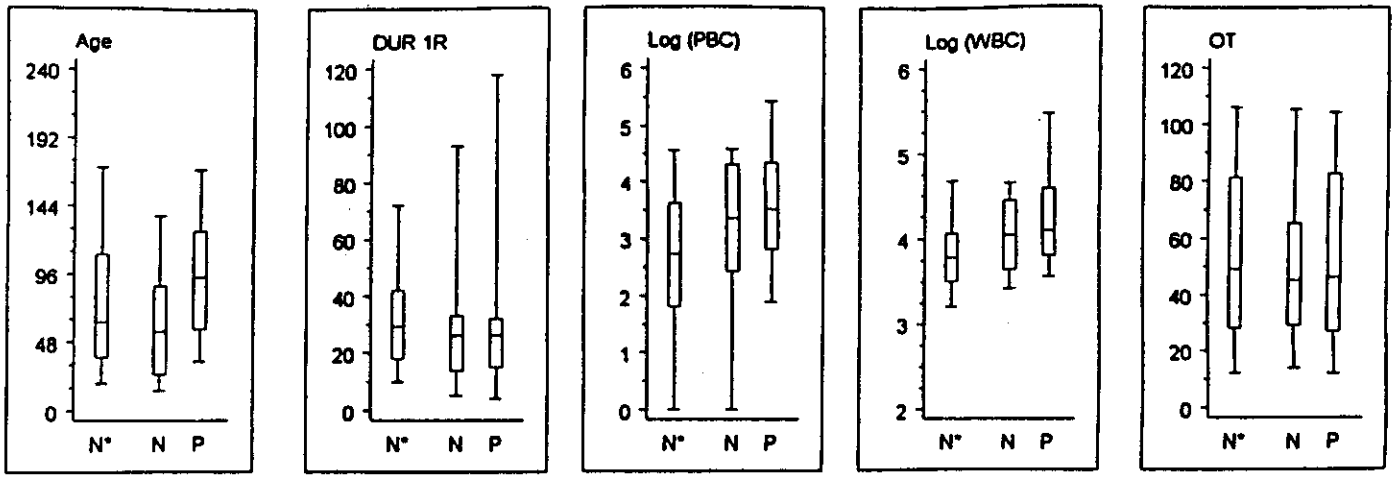


Fig 1. Boxplots of continuous variables by patient groups. Age, age at initial diagnosis (months); DUR 1R, duration of first remission (months); Log (PBC), logarithm (base 10) of (peripheral blast-cell count [1/μL] + 1); Log (WBC), logarithm (base 10) of (WBC count [1/μL] + 1); OT, observation time = (today − date of relapse diagnosis) (months); N*, Ph^1- and/or BCR-ABL-mRNA–negative, not matched, N = 247; N, Ph^1- and/or BCR-ABL-mRNA–negative, matched pairs, N = 30; P, Ph^1- and/or BCR-ABL-mRNA–positive, matched pairs, N = 30. Boxes represent first to third quartile of the distribution; cross line denotes median; upper whisker, ninety-fifth percentile; and lower whisker, fifth percentile.

Journal of Pediatrics 127:88, 1995

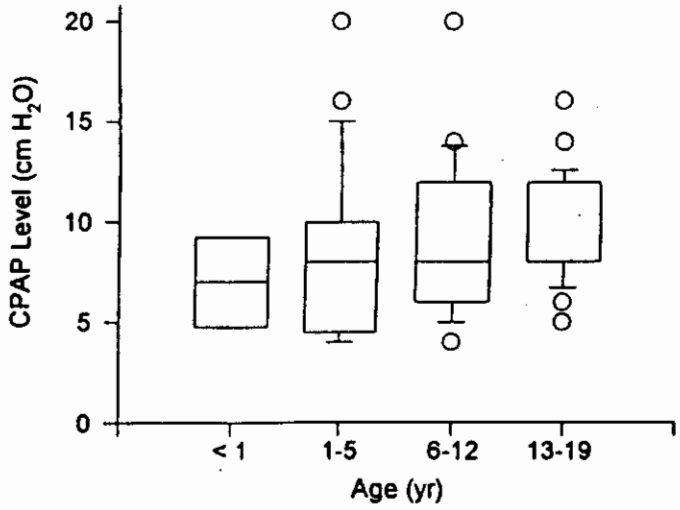


Fig. 1. Levels of *CPAP* according to age group, are shown for 70 subjects for whom levels were available. *Lower boundary* of box indicates 25th percentile of range, *upper boundary* indicates 75th percentile, and *horizontal line* within box indicates median. *Error bars* depict 10th to 90th percentiles; data points outside this range are shown as *circles*. No differences were found in pressure requirements between different age groups.

JAMA 275:1171,1996

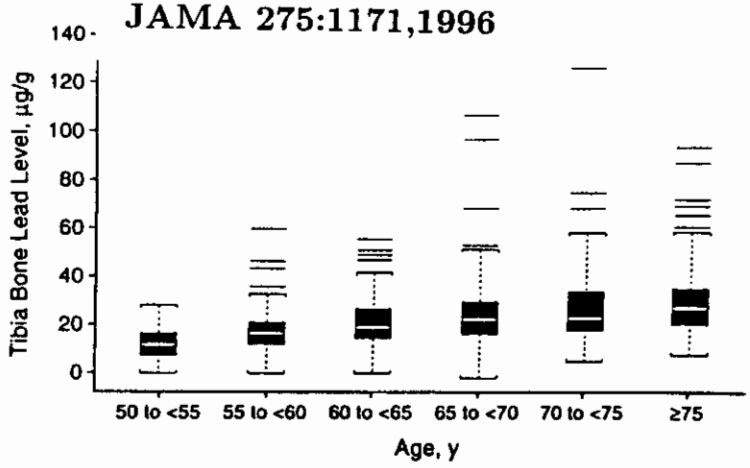


Figure 1.—Box plot of tibia lead levels measured by K x-ray fluorescence vs age among participants in the Normative Aging Study. The horizontal line in the interior of the box is located at the median of the data. The box describes the interquartile distance (IQD) between the third quartile of the data and the first quartile. The dotted lines extend a distance of 1.5 × IQD from the center to the "whiskers." The data bars that fall outside the whiskers may be considered outliers.