Reference Intervals: Practical Approaches

Gary L. Horowitz, MD

Director of Clinical Chemistry
Beth Israel Deaconess Medical Center

Associate Professor of Pathology
Harvard Medical School

9 November 2012

Leiden, The Netherlands

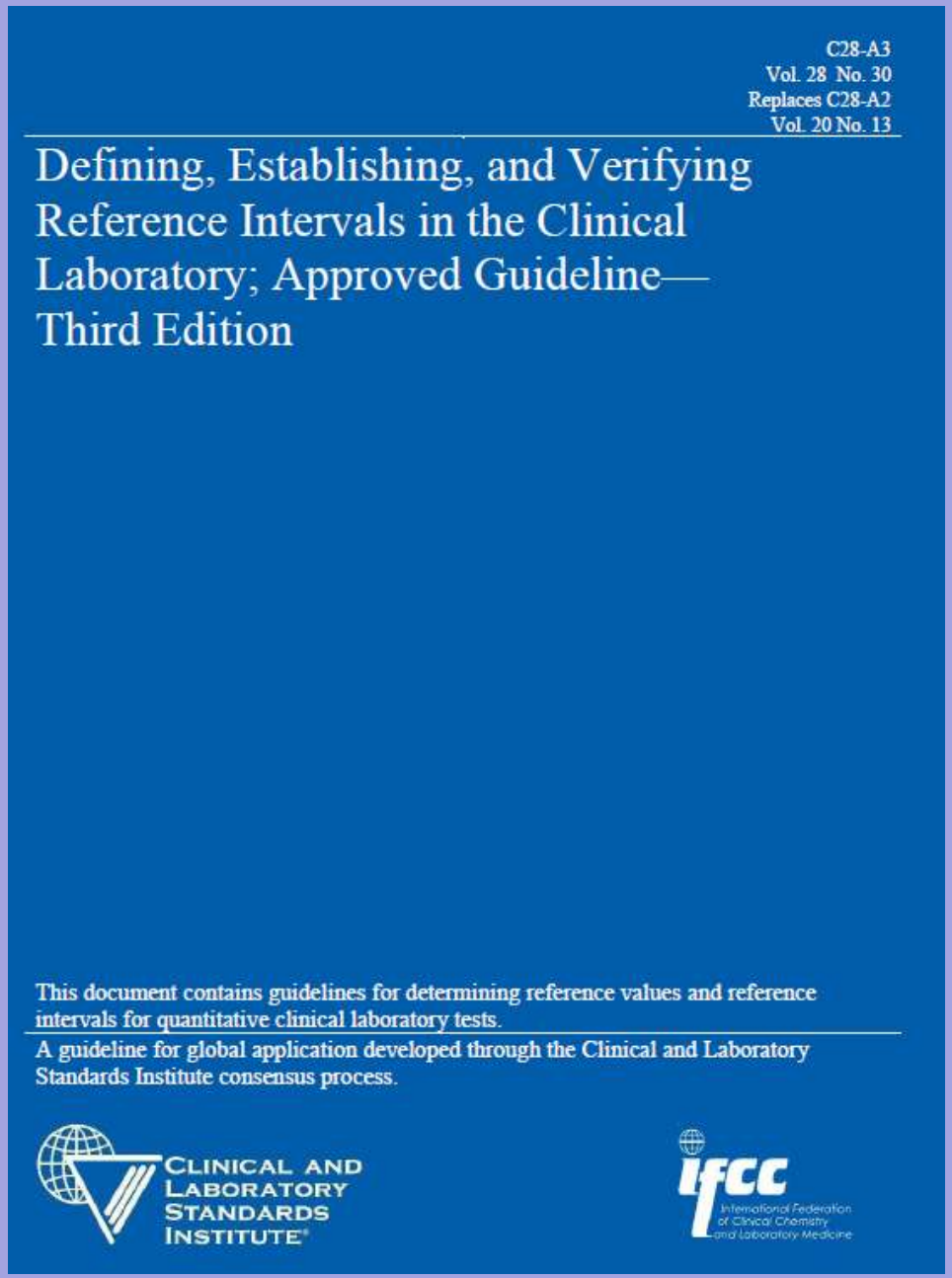
8th Annual ECAT Foundation Meeting

Disclosures

- no financial disclosures

- relevant professional committee positions
  - **Chair, CAP Chemistry Resource Committee**
  - **Past Chair, CLSI Working Group on Reference Intervals**

- use of conventional units



www.clsi.org

Document C28-A3

Establishment of Reference Intervals

- strong endorsement of CLSI C28-A2 recommendation
- preferred method:
  - carefully collect samples from 120 reference individuals
  - use non-parametric method of analysis

Decision Limits

- To use national (or international) guidelines,
  one's method must give accurate results

- Tests where accuracy trumps peer group agreement
  - Cholesterol
  - Hemoglobin A1c
  - Neonatal Bilirubin
  - Glucose (Diabetes Diagnosis)
  - Urine Albumin ("Microalbuminuria")
  - Creatinine (estimated GFR)

- How does one make this assessment?

- Proficiency Testing
  - traditionally, one is assessed against peer groups
  - this is because of "matrix effects" of survey material
  - if the material used is "real", then one can assess accuracy

Hemoglobin A1c Data
(Based on CAP GH2-A 2006 Survey)

GH2-03

Method	No. Labs	Mean	S.D.	C.V.	Median	Low Value	High Value
	24	8.27	0.43	5.3	8.3	7.4	8.9
	23	8.03	0.46	5.8	8.0	7.1	9.3
	291	**7.88**	**0.38**	**4.8**	**7.9**	**6.8**	**9.0**
	15	7.98	0.44	5.5	8.0	6.9	8.6
	20	8.43	0.16	1.9	8.4	8.1	8.8
	253	8.68	0.25	2.9	8.7	8.0	9.4
	41	8.41	0.21	2.5	8.4	7.8	8.9
	489	8.11	0.26	3.2	8.1	7.4	8.9
	15	8.33	0.63	7.6	8.2	7.2	9.4
	22	8.81	0.49	5.5	8.9	7.9	9.7
	250	**8.74**	**0.33**	**3.7**	**8.7**	**7.8**	**9.5**
	192	8.76	0.23	2.6	8.8	8.2	9.5
	195	8.61	0.21	2.5	8.6	7.8	9.8
	25	8.16	0.30	3.7	8.2	7.6	9.0
REFERENCE METHOD *		**8.40**					

more than 50% of values less than 7.9 !

more than 50% of values over 8.7 !

Virtually all values were graded “acceptable” (peer group grading)
Do the labs know it’s not acceptable?

Steps in the Traditional Method

1) Determine biological variables & analytical interferences
2) Determine selection/exclusion/partitioning criteria
3) Obtain written consent and completed questionnaire
4) Categorize reference individuals
5) Exclude individuals as determined a priori
6) Insure an adequate number of reference individuals
7) Prepare reference individuals for sample collection
8) Collect samples
9) Analyze samples
10) Inspect frequency distribution of data
11) Identify data errors and outliers
12) Determine reference intervals (and confidence limits)

Selection of Reference Individuals

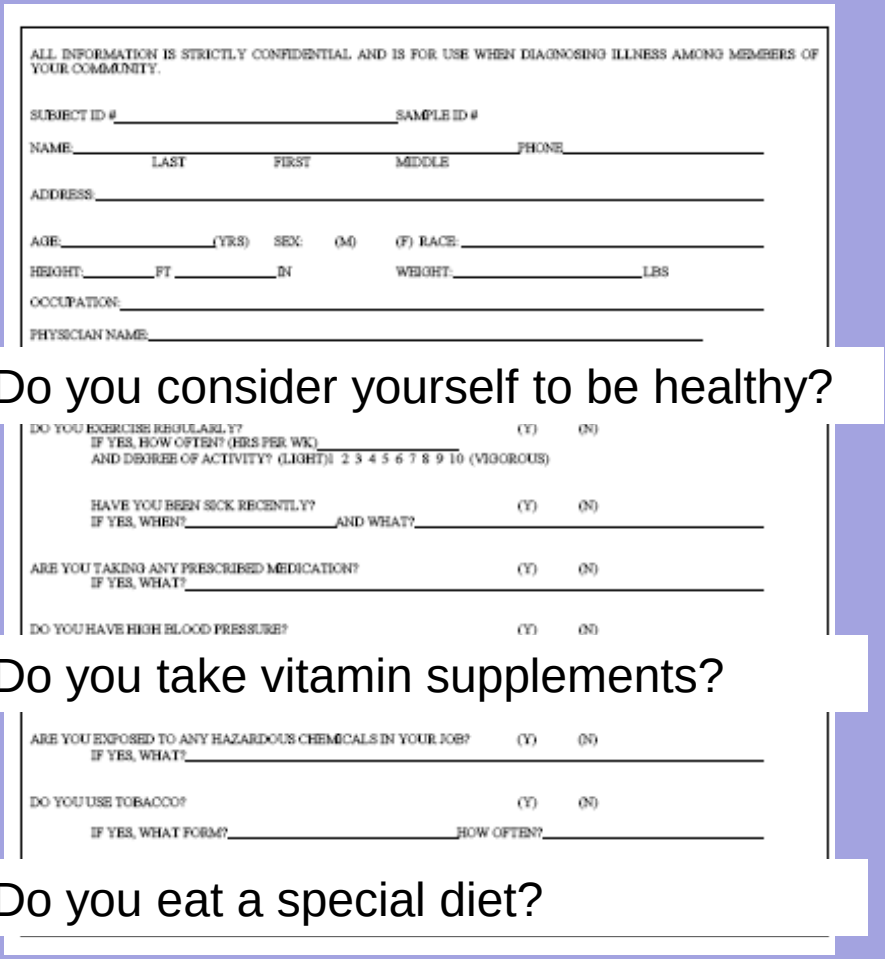
- Exclusion/Partitioning
  - Informed Consent
  - Coding for Privacy

Alcohol Consumption	Illness, recent
Blood donor	Lactation
Blood pressure, abnormal	Obesity
Drug abuse	Occupation
Drugs, prescription	Oral contraceptives
Drugs, over the counter	Pregnancy
Environment	Surgery, recent
Fasting or nonfasting	Tobacco use
Genetic Factors	Transfusion, recent
Hospitalization, current or recent	Vitamin abuse

Age	Posture when sampled
Blood group	Race
Circadian variation	Sex
Diet	Stage of menstrual cycle
Ethnic Background	Stage of pregnancy
Exercise	Time of day when sampled
Fasting or nonfasting	Tobacco use
Geographic location	

Selection of Reference Individuals

- Exclusion/Partitioning
  - Informed Consent
  - Coding
- Questionnaire



ALL INFORMATION IS STRICTLY CONFIDENTIAL AND IS FOR USE WHEN DIAGNOSING ILLNESS AMONG MEMBERS OF YOUR COMMUNITY.

SUBJECT ID # _____ SAMPLE ID # _____

NAME: _____ PHONE: _____
LAST FIRST MIDDLE

ADDRESS: _____

AGE: _____ (YRS) SEX: (M) (F) RACE: _____

HEIGHT: _____ FT _____ IN WEIGHT: _____ LBS

OCCUPATION: _____

PHYSICIAN NAME: _____

Do you consider yourself to be healthy?

DO YOU EXERCISE REGULARLY? (Y) (N)
IF YES, HOW OFTEN? (HRS PER WK) _____
AND DEGREE OF ACTIVITY? (LIGHT) 1 2 3 4 5 6 7 8 9 10 (VIGOROUS)

HAVE YOU BEEN SICK RECENTLY? (Y) (N)
IF YES, WHEN? _____ AND WHAT? _____

ARE YOU TAKING ANY PRESCRIBED MEDICATION? (Y) (N)
IF YES, WHAT? _____

DO YOU HAVE HIGH BLOOD PRESSURE? (Y) (N)

Do you take vitamin supplements?

ARE YOU EXPOSED TO ANY HAZARDOUS CHEMICALS IN YOUR JOB? (Y) (N)
IF YES, WHAT? _____

DO YOU USE TOBACCO? (Y) (N)

IF YES, WHAT FORM? _____ HOW OFTEN? _____

Do you eat a special diet?

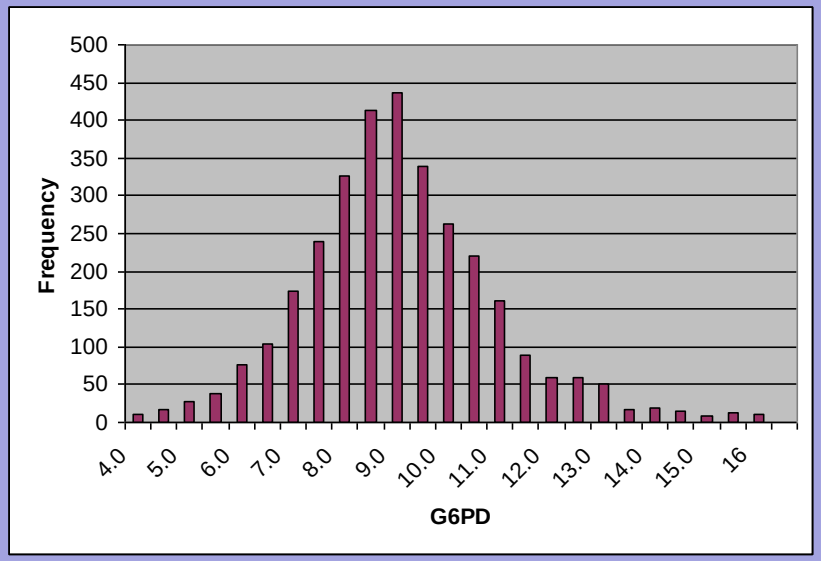
Selection of Reference Individuals

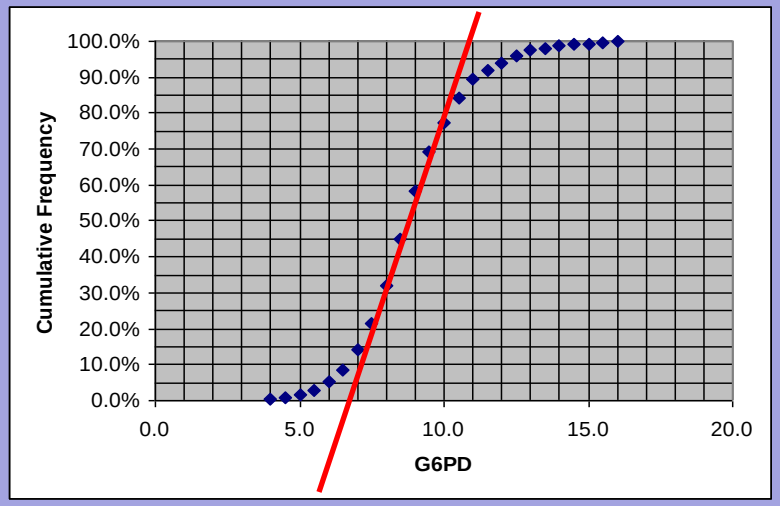
- Exclusion/Partitioning
  - Informed Consent
  - Coding
- Questionnaire
- Sampling Methods:
  - Direct
    - A priori
    - A posteriori
  - ~~Indirect~~

Involves applying statistical methods to values in a laboratory database ***without*** selection of reference individuals

Working group
strongly prefers direct over indirect sampling
but recognizes potential utility of indirect sampling in some situations (e.g., pediatrics)

Hoffmann Technique
(Beth Israel Deaconess Medical Center Data)





assumes Gaussian (normal) distribution of reference individual data!

Hoffman RG. Statistics in the practice of medicine. JAMA 1963; 185:864-873.

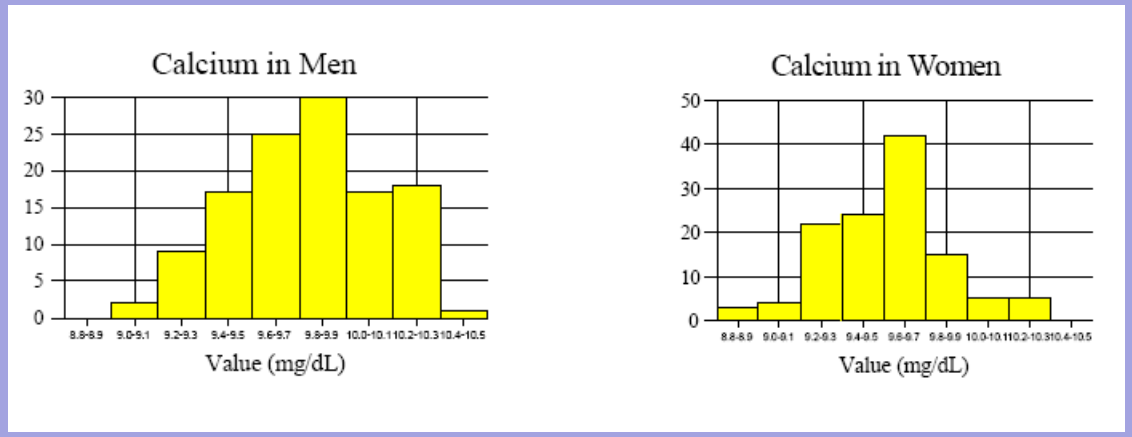
Subject Preparation and Other Pre-Analytical Considerations

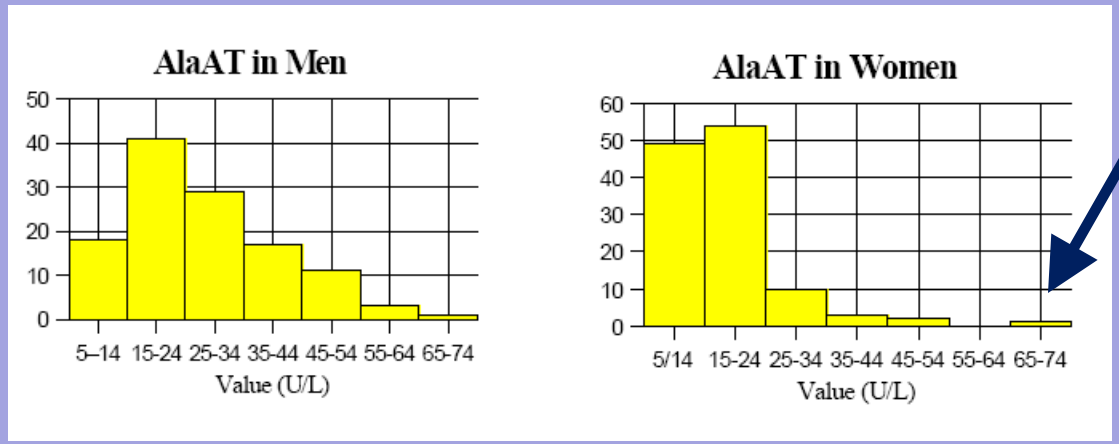
Subject Preparation	Specimen Collection	Specimen Handling
• Prior diet • Fasting vs. nonfasting • Abstinence from pharmacologic agents • Drug regimen • Sampling time in relation to biological rhythms • Physical activity • Rest period before collection • Stress	• Environmental conditions during collection • Time • Body posture • Specimen type • Collection site • Site preparation • Blood flow • Equipment • Technique	• Transport • Clotting • Separation of serum/plasma • Storage • Preparation for analysis

Analysis of Data

- **Determine number of subjects needed**
  - non-parametric: n=120 per (potential) partition
- **Check for outliers**
  - review frequency distribution visually!
  - transform if needed, then Reed/Dixon or Tukey
- **Check for partitioning**
  - normal deviate test, z
- **Determine RI**
  - non-parametric method
  - "robust" methods
- **Establish confidence limits on RI**

Frequency Distributions





Dixon/Reed rule

(extreme – next)
range of values

(65-47) / 60 = 0.30

< 0.33

➔ not an outlier

Rank Order Calcium Reference Values (for non-parametric analysis)

rank order	female n=120	male n=120
1	8.8	9.1
2	8.9	9.1
3	8.9	9.2
4	9.0	9.3
5	9.1	9.3
6	9.1	9.3
7	9.1	9.3
8	9.2	9.3
9	9.2	9.3
10	9.2	9.3
↓	↓	↓
111	10.0	10.2
112	10.0	10.3
113	10.0	10.3
114	10.1	10.3
115	10.1	10.3
116	10.2	10.3
117	10.2	10.3
118	10.2	10.3
119	10.3	10.4
120	10.3	10.6

percentile	rank order n=120	rank order n=240
2.5th	3	6
97.5th	118	235

rank order	combined n=240
1	8.8
2	8.9
3	8.9
4	9.0
5	9.1
6	9.1
7	9.1
8	9.1
9	9.1
10	9.2
↓	↓
231	10.3
232	10.3
233	10.3
234	10.3
235	10.3
236	10.3
237	10.3
238	10.3
239	10.4
240	10.6

Partitioning Calculation

- Use standard normal deviate test
  - for n>60, data need not be normally distributed
  - however, if highly skewed, transformation indicated
- Using calcium data,
  - **men:** **$x = 9.80$, $s_1 = 2.9$**
  - **women:** **$x = 9.57$, $s_2 = 3.1$**
- Suggests difference is statistically significant
- But is difference clinically significant?

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\left[\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)\right]^{1/2}}$$

$z = 5.94$

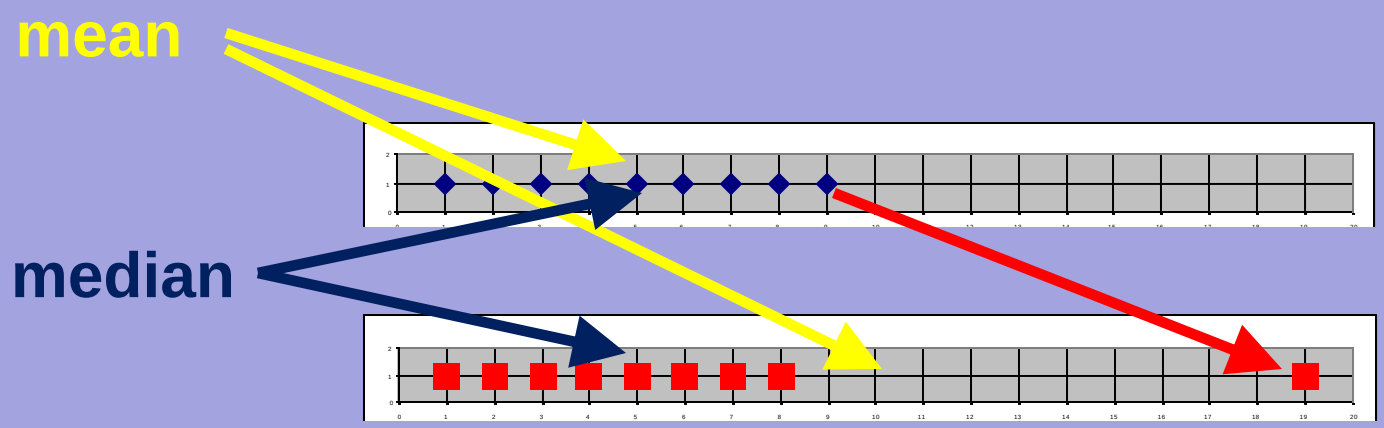
$threshold_{120} = 3$

What If I Don't Have 120 Values?

- **Transform data into a Gaussian distribution**
  - not so easy to do
  - but, for completeness, here's what's involved
  - statistical tests to prove that transformed data is Gaussian
  - then central 95% is: $(\bar{x} - 1.96*SD)$ to $(\bar{x} + 1.96*SD)$
  - 90% confidence limits (later): $2.81*SD / \sqrt{n}$

- **Robust Techniques**

What is Robust, Anyway?



- Think of median vs mean
  - one extreme value can change the mean
  - but it may have no effect on median!

- Similarly, robust techniques take a distribution
  - Make initial robust estimates of "location" and "spread"
  - Give more weight to values towards "center"
  - Calculate, iteratively, new values for "location" and "spread"

More Details on Robust Iterations

x_i		weight iteration 1	weight iteration 2	weight iteration 3	weight iteration 4	weight iteration 5	weight iteration 6
8.9		0.000	0.000	0.000	0.000	0.000	0.000
9.2		0.219	0.180	0.167	0.167	0.162	0.161
9.4		0.752	0.713	0.700	0.700	0.694	0.693
9.4		**0.752**	**0.713**	**0.700**	**0.700**	**0.694**	**0.693**
9.5		0.935	0.912	0.904	0.904	0.900	0.900
9.5		0.935	0.912	0.904	0.904	0.900	0.900
9.5		0.935	0.912	0.904	0.904	0.900	0.900
9.6		1.000	0.998	0.997	0.997	0.996	0.996
9.6		1.000	0.998	0.997	0.997	0.996	0.996
9.6		1.000	0.998	0.997	0.997	0.996	0.996
9.6		1.000	0.998	0.997	0.997	0.996	0.996
9.7		0.935	0.954	0.960	0.960	0.962	0.962
9.7		0.935	0.954	0.960	0.960	0.962	0.962
9.7		0.935	0.954	0.960	0.960	0.962	0.962
9.7		0.935	0.954	0.960	0.960	0.962	0.962
9.7		0.935	0.954	0.960	0.960	0.962	0.962
9.8		0.752	0.788	0.800	0.800	0.805	0.806
9.9		0.491	0.537	0.552	0.552	0.558	0.559
9.9		0.491	0.537	0.552	0.552	0.558	0.559
10.2		0.000	0.000	0.000	0.000	0.000	0.000
T_{bi} =	9.6	9.616	9.622	9.624	9.624	9.624	9.624

Reference Interval Robust Technique Calcium, Women, n=20

Randomly Selected Data Sets of n=20
From Original Date Set of 120 Females

	Sample 1	Sample 2	Sample 3
1	9.5	9.9	9.7
2	9.9	9.5	9.8
3	9.5	10.1	9.5
4	9.7	9.6	9.7
5	9.6	9.9	9.7
6	9.9	10.2	9.8
7	9.7	9.2	9.4
8	10.2	9.7	9.3
9	9.2	9.5	9.6
10	9.4	9.7	9.7
11	9.7	9.3	9.7
12	9.7	9.5	9.7
13	8.9	9.5	9.7
14	9.8	10.0	9.3
15	9.5	9.2	9.6
16	9.6	9.4	10.3
17	9.6	9.6	9.1
18	9.7	10.0	9.7
19	9.4	10.2	9.5
20	9.6	9.7	9.2
Reference Interval	**9.0 - 10.2**	**9.1 - 10.4**	**9.1 - 10.1**

Reference Intervals:

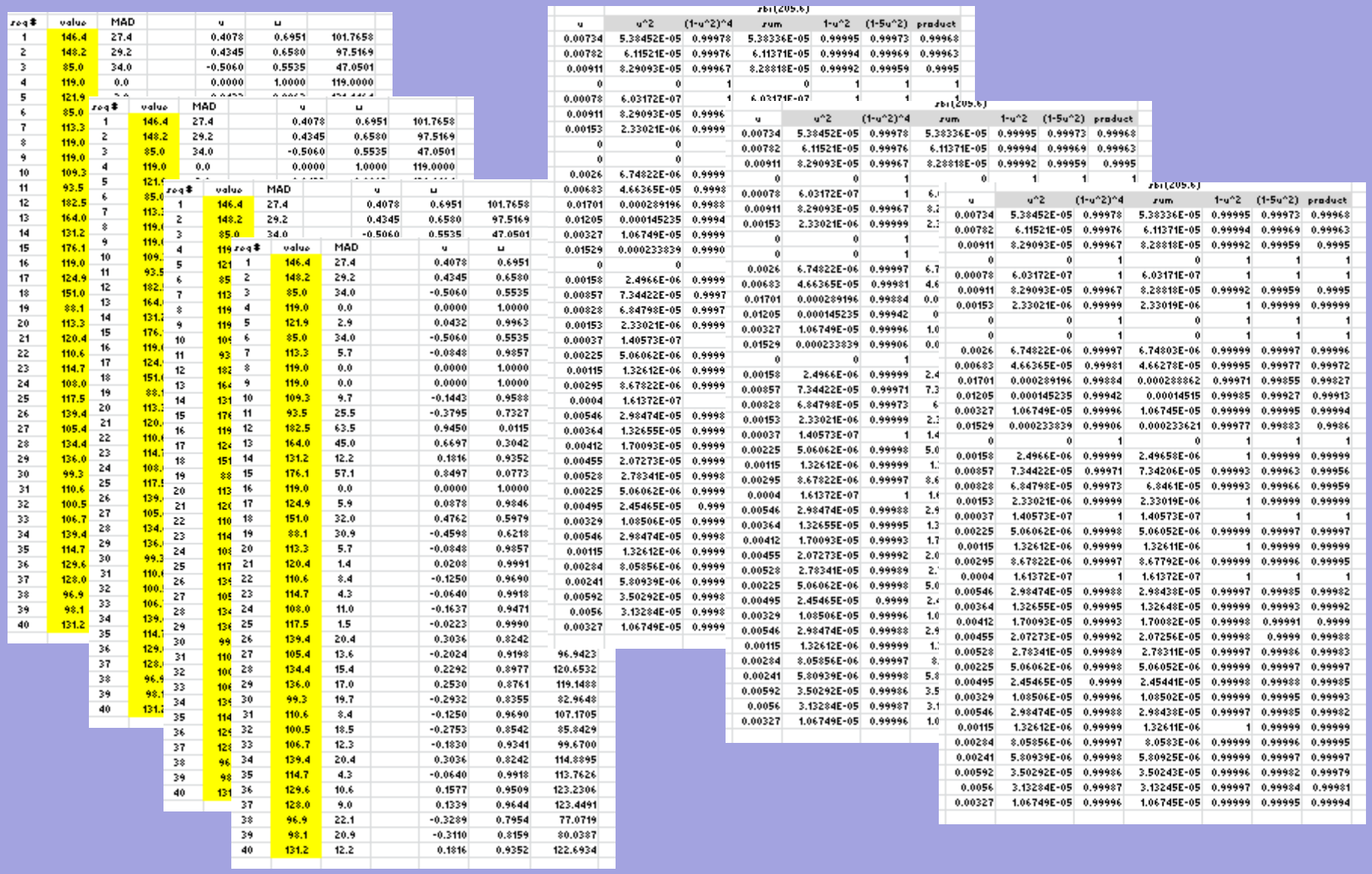
Non-Parametric (n=120) vs Robust (n=20)

	Calcium	Calcium	ALT	ALT
	non-parametric	**robust**	**non-parametric**	**robust**
women	8.9-10.2	9.0-10.2	6-46	7-39
men	9.2-10.3	9.0-10.5	10-55	9-57

Recent Application of Robust Technique

- **BIDMC implementing new system for coagulation factor assays**
  - more than 2 in 20 outside proposed limits
  - succeeded in recruiting ~40 reference individuals
  - robust technique to the rescue!
- **one of my colleagues had heard this talk and asked me to help with analysis**
  - Excel vs StatisPro

Calculations in Excel



req#	value	MAD		u	u	
1	146.4	27.4		0.4078	0.6951	101.7658
2	148.2	29.2		0.4345	0.6580	97.5169
3	85.0	34.0		-0.5060	0.5535	47.0501
4	119.0	0.0		0.0000	1.0000	119.0000
5	121.9	2.9		0.0432	0.9963	121.4432
6	85.0	34.0		-0.5060	0.5535	47.0501
7	113.3	5.7		-0.0848	0.9857	111.6857
8	119.0	0.0		0.0000	1.0000	119.0000
9	119.0	0.0		0.0000	1.0000	119.0000
10	109.3	9.7		-0.1443	0.9588	104.7988
11	93.5	25.5		-0.3795	0.7327	68.5127
12	182.5	63.5		0.9450	0.0115	2.1008
13	164.0	45.0		0.6697	0.3042	49.8888
14	131.2	12.2		0.1816	0.9352	122.6934
15	176.1	57.1		0.8497	0.0773	13.6128
16	119.0	0.0		0.0000	1.0000	119.0000
17	124.9	5.9		0.0878	0.9846	122.9954
18	151.0	32.0		0.4762	0.5979	90.2729
19	88.1	30.9		-0.4598	0.6218	54.7866
20	113.3	5.7		-0.0848	0.9857	111.6857
21	120.4	1.4		0.0208	0.9991	120.3291
22	110.6	8.4		-0.1250	0.9690	107.1705
23	114.7	4.3		-0.0640	0.9918	113.7626
24	108.0	11.0		-0.1637	0.9471	102.2868
25	117.5	1.5		-0.0223	0.9990	117.3490
26	139.4	20.4		0.3036	0.8242	114.8895
27	105.4	13.6		-0.2024	0.9198	96.9423
28	134.4	15.4		0.2292	0.8977	120.6532
29	136.0	17.0		0.2530	0.8761	119.1488
30	99.3	19.7		-0.2932	0.8355	82.9648
31	110.6	8.4		-0.1250	0.9690	107.1705
32	100.5	18.5		-0.2753	0.8542	85.8429
33	106.7	12.3		-0.1830	0.9341	99.6700
34	139.4	20.4		0.3036	0.8242	114.8895
35	114.7	4.3		-0.0640	0.9918	113.7626
36	129.6	10.6		0.1577	0.9509	123.2306
37	128.0	9.0		0.1339	0.9644	123.4491
38	96.9	22.1		-0.3289	0.7954	77.0719
39	98.1	20.9		-0.3110	0.8159	80.0387
40	131.2	12.2		0.1816	0.9352	122.6934

zbi(205.6)

u	u^2	(1-u^2)^4	sum	1-u^2	(1-5u^2)	product
0.00734	5.38452E-05	0.99978	5.38336E-05	0.99995	0.99973	0.99968
0.00782	6.11521E-05	0.99976	6.11371E-05	0.99994	0.99969	0.99963
0.00911	8.29093E-05	0.99967	8.28818E-05	0.99992	0.99959	0.9995
0	0	1	0	1	1	1
0.00078	6.03172E-07	1	6.03171E-07	1	1	1
0.00911	8.29093E-05	0.9996				
0.00153	2.33021E-06	0.9999				
0	0					
0	0					
0.0026	6.74822E-06	0.9999				
0.00683	4.66365E-05	0.9998				
0.01701	0.000289196	0.9988				
0.01205	0.000145235	0.9994				
0.00327	1.06749E-05	0.9999				
0.01529	0.000233839	0.9990				
0	0					
0.00158	2.4966E-06	0.9999				
0.00857	7.34422E-05	0.9997				
0.00828	6.84798E-05	0.9997				
0.00153	2.33021E-06	0.9999				
0.00037	1.40573E-07					
0.00225	5.06062E-06	0.9999				
0.00115	1.32612E-06	0.9999				
0.00295	8.67822E-06	0.9999				
0.0004	1.61372E-07					
0.00546	2.98474E-05	0.9998				
0.00364	1.32655E-05	0.9999				
0.00412	1.70093E-05	0.9999				
0.00455	2.07273E-05	0.9999				
0.00528	2.78341E-05	0.9998				
0.00225	5.06062E-06	0.9999				
0.00495	2.45465E-05	0.999				
0.00329	1.08506E-05	0.9999				
0.00546	2.98474E-05	0.9998				
0.00115	1.32612E-06	0.9999				
0.00284	8.05856E-06	0.9999				
0.00241	5.80939E-06	0.9999				
0.00592	3.50292E-05	0.9998				
0.0056	3.13284E-05	0.9998				
0.00327	1.06749E-05	0.9999				

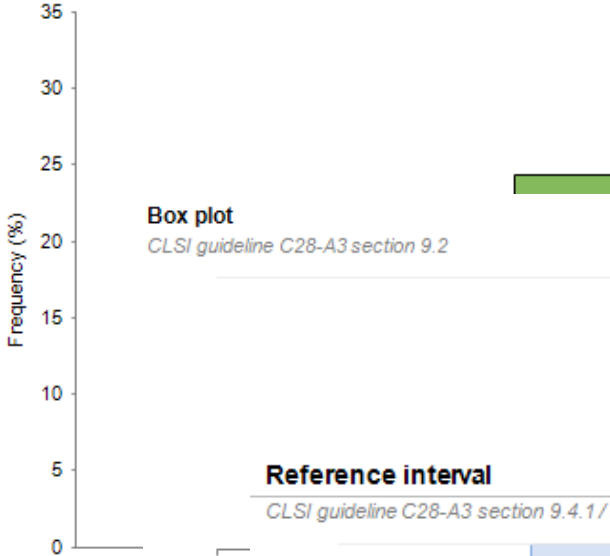
zbi(205.6)

u	u^2	(1-u^2)^4	sum	1-u^2	(1-5u^2)	product
0.00734	5.38452E-05	0.99978	5.38336E-05	0.99995	0.99973	0.99968
0.00782	6.11521E-05	0.99976	6.11371E-05	0.99994	0.99969	0.99963
0.00911	8.29093E-05	0.99967	8.28818E-05	0.99992	0.99959	0.9995
0	0	1	0	1	1	1
0.00078	6.03172E-07	1	6.03171E-07	1	1	1
0.00911	8.29093E-05	0.99967	8.28818E-05	0.99992	0.99959	0.9995
0.00153	2.33021E-06	0.99999	2.33019E-06	1	0.99999	0.99999
0	0	1	0	1	1	1
0	0	1	0	1	1	1
0.0026	6.74822E-06	0.99997	6.74803E-06	0.99999	0.99997	0.99996
0.00683	4.66365E-05	0.99981	4.66278E-05	0.99995	0.99977	0.99972
0.01701	0.000289196	0.99884	0.000288862	0.99971	0.99855	0.99827
0.01205	0.000145235	0.99942	0.00014515	0.99985	0.99927	0.99913
0.00327	1.06749E-05	0.99996	1.06745E-05	0.99999	0.99995	0.99994
0.01529	0.000233839	0.99906	0.000233621	0.99977	0.99883	0.9986
0	0	1	0	1	1	1
0.00158	2.4966E-06	0.99999	2.49658E-06	1	0.99999	0.99999
0.00857	7.34422E-05	0.99971	7.34206E-05	0.99993	0.99963	0.99956
0.00828	6.84798E-05	0.99973	6.8461E-05	0.99993	0.99966	0.99959
0.00153	2.33021E-06	0.99999	2.33019E-06	1	0.99999	0.99999
0.00037	1.40573E-07	1	1.40573E-07	1	1	1
0.00225	5.06062E-06	0.99998	5.06052E-06	0.99999	0.99997	0.99997
0.00115	1.32612E-06	0.99999	1.32611E-06	1	0.99999	0.99999
0.00295	8.67822E-06	0.99997	8.67792E-06	0.99999	0.99996	0.99995
0.0004	1.61372E-07	1	1.61372E-07	1	1	1
0.00546	2.98474E-05	0.99988	2.98438E-05	0.99997	0.99985	0.99982
0.00364	1.32655E-05	0.99995	1.32648E-05	0.99999	0.99993	0.99992
0.00412	1.70093E-05	0.99993	1.70082E-05	0.99998	0.99991	0.9999
0.00455	2.07273E-05	0.99992	2.07256E-05	0.99998	0.9999	0.99988
0.00528	2.78341E-05	0.99989	2.78311E-05	0.99997	0.99986	0.99983
0.00225	5.06062E-06	0.99998	5.06052E-06	0.99999	0.99997	0.99997
0.00495	2.45465E-05	0.9999	2.45441E-05	0.99998	0.99988	0.99985
0.00329	1.08506E-05	0.99996	1.08502E-05	0.99999	0.99995	0.99993
0.00546	2.98474E-05	0.99988	2.98438E-05	0.99997	0.99985	0.99982
0.00115	1.32612E-06	0.99999	1.32611E-06	1	0.99999	0.99999
0.00284	8.05856E-06	0.99997	8.0583E-06	0.99999	0.99996	0.99995
0.00241	5.80939E-06	0.99998	5.80925E-06	0.99999	0.99997	0.99997
0.00592	3.50292E-05	0.99986	3.50243E-05	0.99996	0.99982	0.99979
0.0056	3.13284E-05	0.99987	3.13245E-05	0.99997	0.99984	0.99981
0.00327	1.06749E-05	0.99996	1.06745E-05	0.99999	0.99995	0.99994

StatisPro v2.0 (from CLSI)

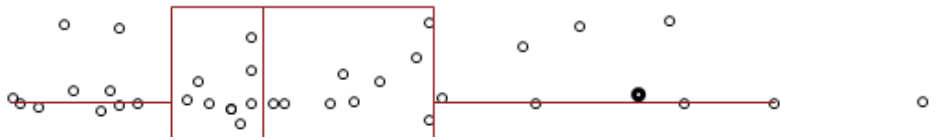
Frequency histogram

CLSI guideline C28-A3 section 9.2



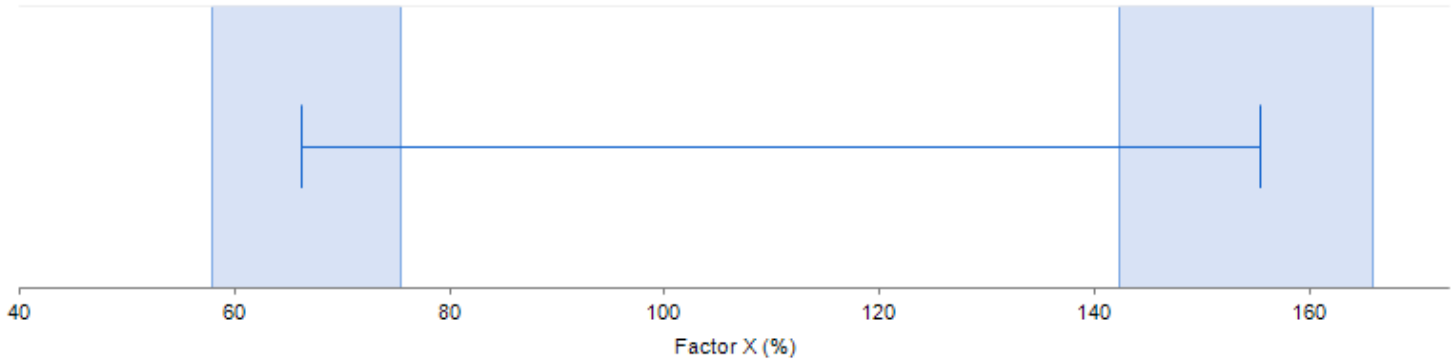
Box plot

CLSI guideline C28-A3 section 9.2



Reference interval

CLSI guideline C28-A3 section 9.4.1 / 9.5.1



	Computation method	95% reference interval	Lower reference limit 90% CI	Upper reference limit 90% CI
	Robust symmetric	66.28 to 155.50	57.81 to 75.37	142.26 to 165.99

Importance of Confidence Limits

- Provide a **quantitative measure**
  of the **variability** of the **estimate of RI**

- This **variability narrows**
  as the **sample size increases**

- To get improved precision of the RI,
  one may choose to obtain larger n

Confidence Limits
Non-Parametric Technique

90% confidence intervals
vary with sample size

for n=120

2.5% ile: 1st - 7th point
97.5% ile: 114th - 120th point

for calcium in women,

lower limit 8.9 (8.8 - 9.1)
upper limit 10.2 (10.1 - 10.3)

rank order	female n=120
1	8.8
2	8.9
3	8.9
4	9.0
5	9.1
6	9.1
7	9.1
8	9.2
9	9.2
10	9.2
111	10.0
112	10.0
113	10.0
114	10.1
115	10.1
116	10.2
117	10.2
118	10.2
119	10.3
120	10.3

Confidence Limits: Robust Techniques

- no formula
- rather, use "bootstrapping"
  - sample, with replacement, many times
  - each time, calculate lower reference interval
    - take upper & lower 5% of these determinations as 90% confidence intervals
  - repeat for upper reference limit

n	Lower Reference Limit (Calcium, Women)	Upper Reference Limit (Calcium, Women)
20	8.4 - 10.2	9.9 - 10.4
40	8.7 - 10.0	10.0 - 10.4
80	8.9 - 9.1	10.0 - 10.2

vs. non-parametric, n = 120

8.8 - 9.1 10.1 - 10.3

Transference

- If a laboratory has a current, well-established RI,

  it **may not need to establish** a new RI

  (which requires new samples from reference individuals)

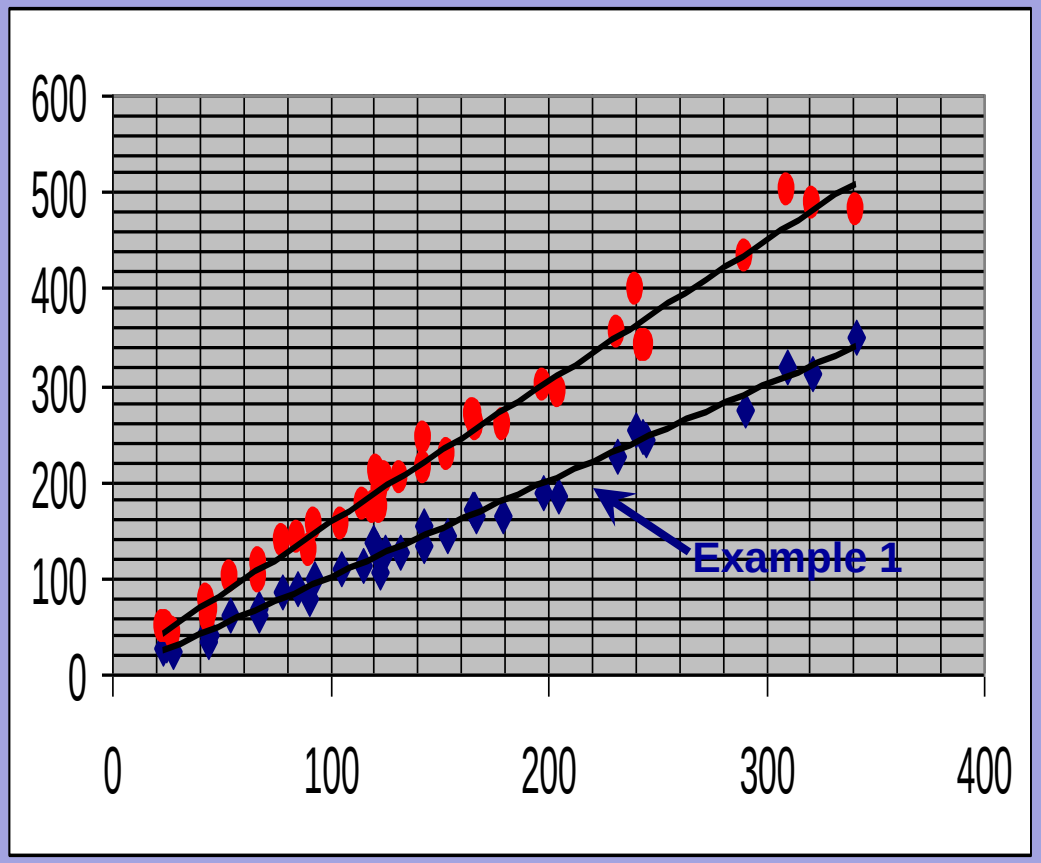
- Rather,

  it **may be able to transfer** the current RI

  (using samples already in lab from typical patients)

  no samples from reference individuals needed!*

Transference Examples
(adapted from CLSI EP9-A2)



$y = 1.466x - 8.960$
$r^2 = 0.985$

Current RI 50 – 150

very high correlation
negligible intercept

Transferred RI
1.466 (50) – 8.960 = 78
1.466 (150) – 8.960 = 224

i.e., 78 - 224

Validation

- **Every laboratory** is capable of validation

- Working Group **strongly endorses C28-A2 method:**
  - pay strict attention to pre-analytic and analytic variables
  - collect samples from 20 reference individuals
  - if no more than 2 (of 20) is outside proposed RI, the proposed RI can be used
  - probability of false rejection is 5-7%
    - i.e., rejecting proposed RI when, in fact, it is valid
    - based on binomial distribution

- C28-A3 adds information on other, sophisticated tests

2004 Normal "Study" (Vetted BIDMC Volunteers)

- protocol:
  - find several labs using same methods
  - each collects a relatively small number of "normals" (n~20)
  - pool data to get number needed (>120) to establish reference interval

- BIDMC data set:
  - screened, healthy volunteers
  - age: range 27-63, mean 41
  - predominantly white, ~2/3 female

2004 Normal "Study" (Vetted BIDMC Volunteers)

	Cholesterol <200		
n	20		
mean	190.65		
"SD"	37.79		
"mean-2SD"	115 (min=138)		
"mean+2SD"	266 (max=253)		
below/above Roche RR	0 8		

2004 Normal "Study" (Vetted BIDMC Volunteers)

	Cholesterol <200	Calcium 8.4-10.2	
n	20	20	
mean	190.65	9.535	
"SD"	37.79	0.274	
"mean-2SD"	115 (min=138)	8.99 (min=9.0)	
"mean+2SD"	266 (max=253)	10.08 (max=10.1)	
below/above Roche RR	0 8	0 0	

2004 Normal "Study" (Vetted BIDMC Volunteers)

	Cholesterol <200	Calcium 8.4-10.2	TSH 0.3-4.2
n	20	20	20
mean	190.65	9.535	2.51
"SD"	37.79	0.274	1.481
"mean-2SD"	115 (min=138)	8.99 (min=9.0)	-0.4 (min=0.55)
"mean+2SD"	266 (max=253)	10.08 (max=10.1)	5.1 (max=5.9)
below/above Roche RR	0 8	0 0	0 2

Multicenter Trials

- **Why should each laboratory establish its own RI?**
  - differences ***in methods***
    - given traceability, this may no longer be necessary
  - differences ***in populations***
    - alleged, but not frequently documented

- **Requirements to insure success of multicenter trials:**
  - a priori selection, insuring adequate numbers of subjects
  - pre-analytic phase requirements
  - ***traceability, inclusion of commutable reference materials***
  - QC program

CAP Reference Range Service
(based on CAP RRS-B 2006 Survey)

		n	min	mean	max	2.5%ile		insert
Calcium	BIDMC	20	8.60	9.35	10.			
	all platform	335					10.29 (10.18 - 10.40)	8.4 - 10.2
TS				2.169	4.370			
		267	0.090	2.032	5.360	0.59 (0.440 - 0.750)	4.38 (4.110 - 4.650)	0.27 - 4.2

CAP dropped this survey in 2011 for lack of participants!

Note that 2 values from BIDMC became outliers in overall analysis (6.62, 617)

A Superbly Done Reference Interval Study
(with very practical implications)

Prevention and Rehabilitation

Distribution of creatine kinase in the general population: Implications for statin therapy

Lizzy M. Brewster, MD,[a] Gideon Mairuhu, MD,[c] August Sturk, PhD,[b] and Gert A. van Montfrans, MD, PhD[a]
Amsterdam and Nieuwegein, The Netherlands

Background Eligible subjects with mildly elevated serum creatine kinase (CK) activity are often excluded before randomization in statin trials, but patients may potentially be misclassified as having hyperCKemia when inappropriate reference limits are used. Little information is usually given regarding how reference limit data were established, although evidence suggests that the variation of CK activity in the general population is wider than reflected in reference intervals in current use.

Methods We determined reference intervals for ... according to NCCLS guidelines

1444 individuals, after 3 days of rest

...of white European (n = 503), South Asian

Results The calculated upper reference limits (97.5th percentile) for nonblack and black women and men were 2 to 5 times higher than recommended by the assay manufacturer. Respectively 13% of the white Europeans, 23% of South Asians, and 49% of the black people had serum CK activities above the manufacturer-provided limits.

Conclusion The variation in CK activity within the population is wider than previously suggested in smaller, nonrandom samples, and relatively high values occur frequently in all subgroups studied after rest. Therefore, we infer that upward adjustment of the upper reference limit is necessary for all population subgroups studied. The use of appropriately established reference intervals may improve the use of statins and particularly benefit the control of dyslipidemia in those with relatively high baseline CK activity.

Am Heart J. 2007;154:655-61

Distribution of the data

We visually inspected the distribution of the data and the values in the tails of the distribution to identify data errors and outliers. In addition, the Dixon range statistic was used. When this method is applied to the upper end of the distribution, the largest value may be an outlier if the difference between the 2 largest values is greater than one third of the difference between the maximum and minimum values of the distribution. If several probable outliers are present, the one-third rule is applied to the least extreme outlier as if it were the only outlier.[14]

Calculated reference limits

After deletion of errors and outliers, we nonparametrically assessed the 2.5th and 97.5th percentiles as the reference limits for CK[14] and compared these with the reference interval recommended by the manufacturer.

Partition

We decided that partition into subclasses of sex and self-defined ancestry would be clinically relevant, and assessed whether partition would reduce the variation in the data by inspecting the data and by statistical partitioning testing, assessing at least 120 subjects in each subgroup.[14] Several statistical criteria have been proposed to establish separate reference intervals for different population groups.[14,16,17] We followed the guidelines of the CLSI and applied the standard normal deviate test to the original data when we assessed 2 subclasses. In this test, z statistics are calculated for 2 groups of at least 120 people and compared with the critical value: $z^* = 3(n_1 + n_2/240)^{1/2}$.[14] Separate reference intervals for each subclass are appropriate when the calculated z value exceeds z^*, or if the larger SD exceeds 1.5 SD_1.[14] For the assessment of sex-ancestry groups involving more than 2 subclasses, analysis of variance with a Tukey posttest was performed on the data after a logarithmic transformation to base 10. Any significant outcome in the Tukey posttest was retested with the more stringent z^* criterion as described above.

"visually inspected the data"

"Dixon range statistic was used"

"non-parametrically assessed the 2.5th and 97.5th percentiles"

"decided to partition into subclasses"

"assessing at least 120 subjects in each subgroup"

Brewster LM et al. Am Heart J. 2007;154:655-61.

Non-Parametric Reference Intervals

should be ~2.5%

Gender	Ethnicity	n	2.5th %ile	97.5th %ile	% > ULN from manufacturer
Women	White	252	29	201	8%
Women	SouthAsian	147	37	313	16%
Women	Black	387	48	414	42%
Men	White	251	47	322	17%
Men	SouthAsian	123	47	641	32%
Men	Black	183	71	801	62%

Brewster LM et al. Am Heart J. 2007;154:655-61.

Upshot of This Data

- hyperCKemia overdiagnosed
- statins may be discontinued based on incorrect reference intervals
- labs who did not verify their reference intervals share responsibility for this problem (and that includes most of us)

Summary

- **Decision Limits vs Reference Intervals (RI)**
- **Reference Individuals:**
  - selection/partitioning/preparation
- **Data Analysis**
  - examine distribution/eliminate outliers
  - ***to establish RI***
    - n=120, non-parametric preferred
    - <120: transform to Gaussian or use robust method
  - ***to verify RI established elsewhere***
    - n=20, valid if no more than 2 outside proposed RI

Thank you for your attention!

Questions and/or Comments?

Course Objectives

Upon completion of this session, participants will be able to…..

- perform reference interval validation studies with as few as 20 samples
- identify 5 commonly performed tests where conventional reference intervals are not relevant
- list 3 resources that can be used to help with reference interval determinations in their own laboratories

Answers:

- one can verify a reference interval with
  20 reference individuals (no more than 2 outside proposed interval)

- accuracy required:
  Hemoglobin A1c
  Neonatal Bilirubin
  Cholesterol
  Glucose
  Creatinine

- reference interval resources:
  CAP Accuracy-Based Surveys
  CLSI C28-A3
  CAP Reference Range Service

Hemoglobin A1c Data
(Based on CAP GH2-A 2006 Survey)

Reference Value	Within 7% of Reference Value	Range of Peer Group Pass Rates	Overall Pass Rate
5.3	4.9 - 5.7	44.8 - 97.2	86.7
8.4	7.8 - 9.0	42.3 - 100	85.6
10.7	9.9 - 11.5	53.1 - 98.1	84.7

BIDMC Outpatient CK Data

caveats on BIDMC data:
- all outpatients (not reference individuals)
- race not known (predominantly white)

		Paper	BIDMC
women	**%>140**	**37%**	**26%**
	median	**95**	**93**
men	**%>174**	**35%**	**30%**
	median	**143**	**130**

Statins and Muscle Disease
ACC/AHA/NHLBI Clinical Advisory

- "Routine monitoring of CK is of little value in the absence of clinical signs or symptoms."

- "Therefore, all persons beginning to receive statins should be instructed to report muscle discomfort or weakness or brown urine immediately, which should then prompt a CK measurement."

- If a patient has signs or symptoms,
  - Check TSH as well as CK
  - If CK > 10 X ULN ➔ stop statin therapy immediately
  - If CK < 10 X ULN ➔ maintain statin therapy & monitor weekly

J Am Coll Cardiol 2002;40:567-572