Biological Variation: From Principles to Practice

Dr Callum G Fraser
Biochemical Medicine - NHS Tayside
Ninewells Hospital and Medical School
Dundee Scotland

Content and expected outcomes

After this lecture, participants should :

- *be able to list the types of biological variation,*
- *know how to generate and/or find data on random biological variation, and*
- *appreciate how to use the data in laboratory medicine.*

Content and expected outcomes

Participants should appreciate how to use data on biological variation in:

- *setting quality specifications,*
- *assessing the significance of changes in serial results from an individual,*
- *deciding the utility of conventional population based reference values, and*
- *managing aspects of quality.*

Biological variation

Variation over the span of life

Predictable rhythmical/cyclical variation

Random variation around setting points

Life-long biological variation

Some analytes change over the span of life -

- *neonatal period*
- *childhood*
- *puberty*
- *menopause*
- *old age*

This is taken care of by creation of age stratified reference values when needed.

Predictable rhythmical/cyclical rhythms

- *Daily*

  cortisol, growth hormone
- *Monthly*

  LH, FSH, progesterone
- *Seasonal*

  vitamin D - also cholesterol, HbA1c

Problems associated with rhythms/cycles

- *it is impossible to develop good reference values for every time point during the cycle*
- *knowledge of the expected values throughout the cycle is vital for clinical interpretation*
- *for some analytes, samples should be taken at relevant times for the clinical purpose*
- *absence of the expected cycle may provide clinically useful information*

Most analytes - random variation

A series of four results taken from an individual

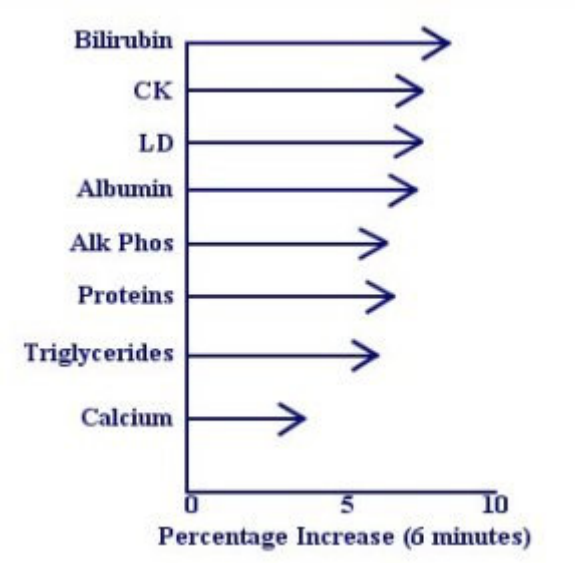
Sodium	*[mmol/L]*	*137*	*139*	*136*	*138*	*[135 -147]*
Potassium	*[mmol/L]*	*4.3*	*4.6*	*4.5*	*4.4*	*[3.5 - 5.0]*
Urea	*[mmol/l]*	*4.0*	*4.4*	*4.1*	*3.9*	*[3.3 - 6.6]*
Creatinine	*[µmol/L]*	*88*	*97*	*89*	*92*	*[64 - 120]*
Bilirubins	*[µmol/L]*	*19*	*21*	*16*	*20*	*[up to 17]*

Pre-analytical sources of variation

- *Preparation of subject*
  - *fasting*
  - *exercise*
  - *posture*

Sources of random variation

- *Pre-analytical - preparation of subject*
  sample collection and handling



Examples of increase in large molecules and quantities bound to large molecules on venous stasis

Pre-analytical sources of variation

- *Preparation of subject*
  - *fasting*
  - *exercise*
  - *posture*
- *Sample collection and handling*
  - *type of sample*
  - *anticoagulant*
  - *tourniquet*
  - *transport time*
  - *centrifugation*

Sources of random analytical variation

Analytical - imprecision

changes in bias

Sources of random analytical variation

*Imprecision - random error - usually expressed as SD or CV [SD/mean]*100*

- *intrinsic to all measurements,*
- *due to factors such as variations in temperature, volume of sample or reagent delivered, AND*
- *is method dependent.*

Sources of random analytical variation

Bias - systematic error [due to calibration and other constant sources of error]

constant bias does not affect results over time [they may all be low or all be high due to bias] but changes in bias - most importantly on re-calibration of methods - is a source of variation in serial results

A quality control material for sodium - replicate analyses

Series 1	Series 2	
140	140	Series 1 done with one
141	142	calibrator - mean = 140 mmol/L
139	141	SD = 1.15 mmol/L
140	141	
142	139	Series 2 - after re-calibration
138	143	mean = 141 mmol/L
139	140	SD = 1.15 mmol/L
141	141	
140	142	Overall SD = 1.24 mmol/L
140	141	

Sources of random variation

- *Pre-analytical - preparation of subject*
  sample collection/handling
- *Analytical - imprecision*
  changes in bias
- *Biological - within-subject*
  biological variation

Random test result variation

A series of four results taken from each of a cohort of four individuals - sodium [mmol/L]

Individual 1	*137*	*139*	*136*	*138*	*[135 -147]*
Individual 2	*144*	*146*	*145*	*144*	*[135 -147]*
Individual 3	*141*	*143*	*142*	*140*	*[135 -147]*
Individual 4	*136*	*138*	*137*	*135*	*[135 -147]*

homeostatic setting points do vary amongst individuals - between-subject biological variation

Variation of derived indices

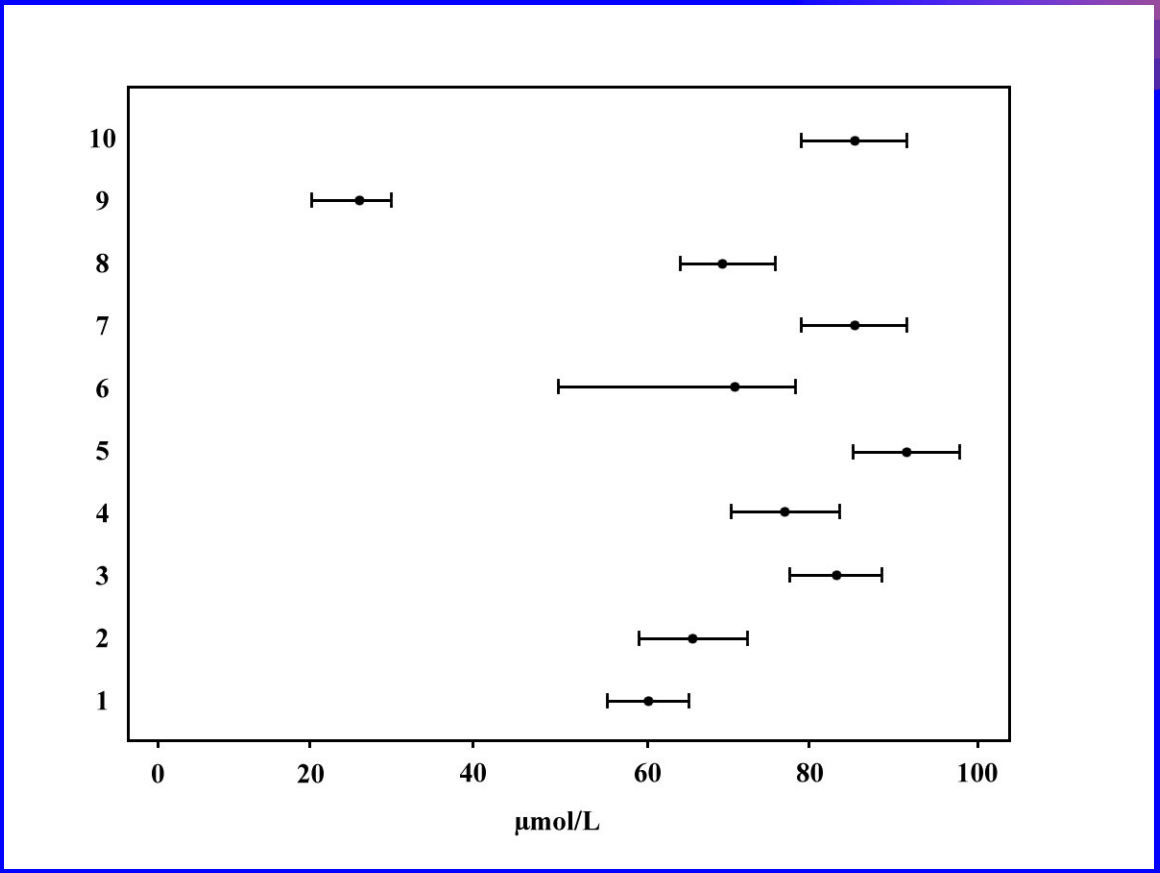
Date	Creat	Urea	Alb	Age	eGFR
11 March 1992	92	4.4	45	47	87.4
21 April 1995	97	4.3	46	50	82.9
23 January 1996	86	3.7	45	51	94.9
28 August 1998	94	4.7	44	53	82.3
12 July 1999	88	4.3	44	54	88.9
13 January 2003	89	4.4	42	58	85.2

Derived indices have random variation over time!

Generation of numerical estimates of components of biological variation

- *select a small number of reference individuals*
- *apply exclusion criteria [just as for reference values]*
- *take samples at intervals minimizing pre-analytical sources of variation [subject preparation and sample collection and handling]*
- *store as to ensure stability*
- *analyze in random duplicate in one batch*
- *look for outliers [complex]*
- *use nested ANOVA to determine CV_A, CV_I, CV_G*

Mean and absolute ranges for serum creatinine in 10 subjects



To generate or not to generate?

It is dogma that all laboratories generate their own reference values

- analogously, do all laboratories have to generate their own data on components of biological variation?

The answer is ……..absolutely NOT!

Within-subject biological variation of sodium and urea as CV

No	Time	Sex	Na	Urea	Country
11	2 weeks	M	0.7	12.3	Denmark
10	4 weeks	M	0.9	14.3	USA
10	8 weeks	M	0.6	9.5	Germany
14	8 weeks	F	0.5	11.3	Germany
9	12 weeks	M	1.4	13.6	USA
11	15 weeks	M	0.6	15.7	Denmark
37	22 weeks	M	0.5	11.1	England
15	40 weeks	M&F	0.7	13.9	Scotland

Biological variation in young and elderly as CV

Quantity	Young	Elderly
Sodium	0.7	0.9
Urea	13.8	10.3
Calcium	2.1	1.6
Albumin	2.2	2.6
Cholesterol	4.9	5.8
Glucose	4.8	4.7

Within-subject variation in urine as CV

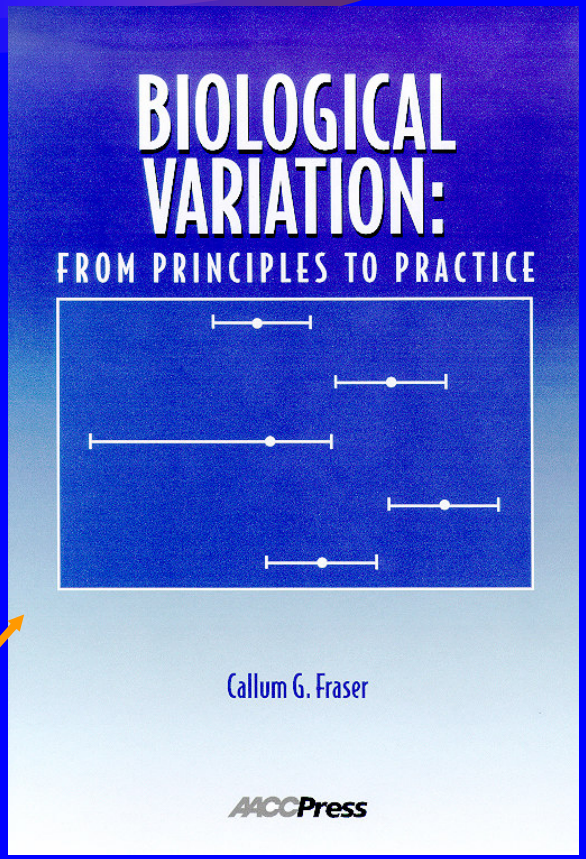
Analyte	Australia	Scotland	Spain
Sodium	28.0	26.5	28.7
Calcium	25.1	26.2	27.5
Creatinine	11.2	11.0	15.0
Phosphate	16.6	16.9	20.6

Data on biological variation

Over the years, many compilations, most recent -

Ricos C, et al. Current databases on biologic variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59:491-500

2006 update at www.westgard.com/guest32.htm



Available at www.aacc.org

Clinical Chemistry. 2006;52:650-656

Estimation and Application of Biological Variation of Urinary Delta -Aminolevulinic Acid and Porphobilinogen in Healthy Individuals and in Patients with Acute Intermittent Porphyria

Aasne K. Aarsand[1,a], Per Hyltoft Petersen[2] and Sverre Sandberg[1,2]

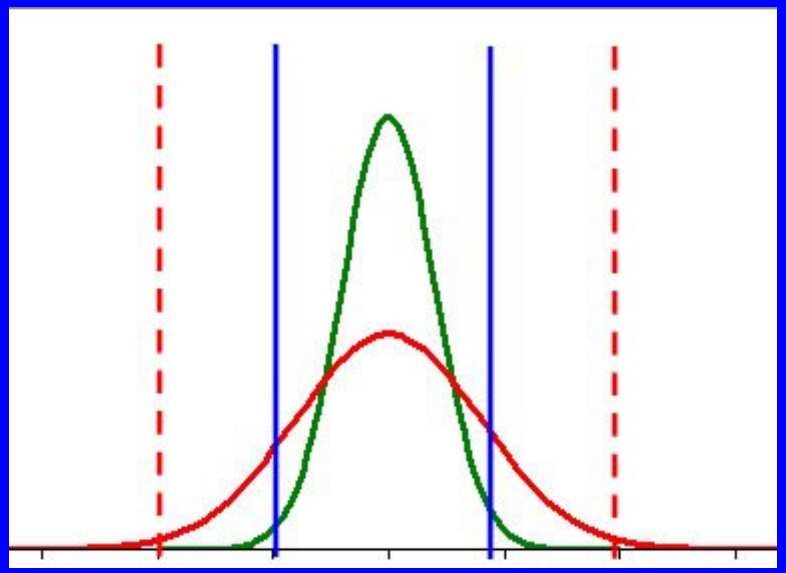
Application of data on biological variation

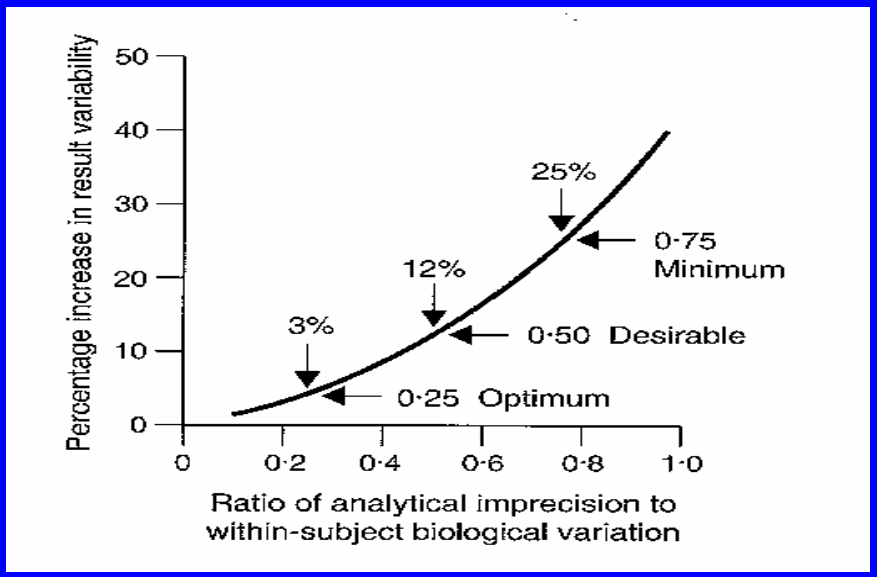
- *setting quality specifications*
- *assessing the significance of changes in serial results from an individual*
- *deciding the utility of conventional population based reference values*
- *other uses in quality management*

Setting quality specifications

$$Imprecision < x \bullet CV_I$$

For most situations, x should be 0.5 [1/2].

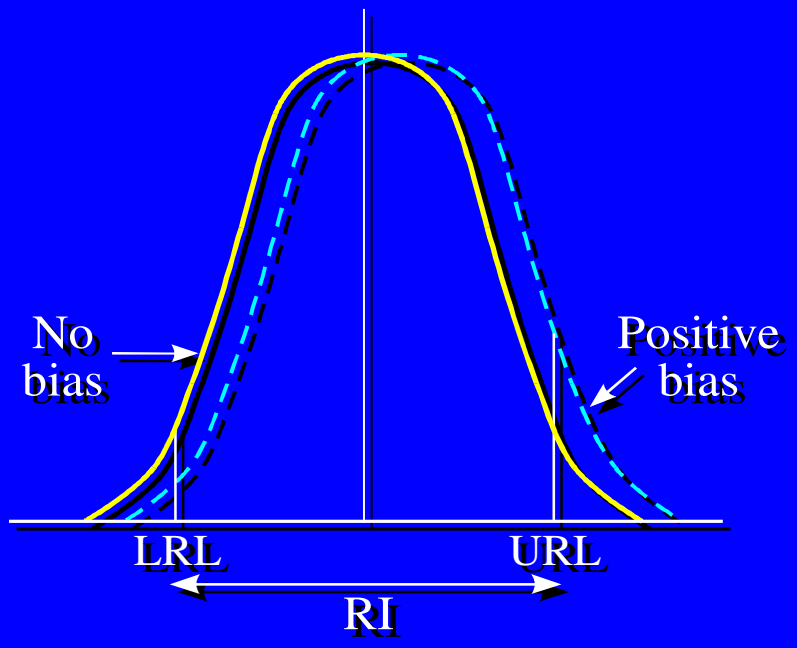




Setting quality specifications

$$Bias < y \bullet [CV_I^2 + CV_G^2]^{1/2}$$

For most situations, y should be 0.25 [1/4].



Setting quality specifications

- *Total error* $< Z \bullet x \bullet CV_I + y \bullet [CV_I^2 + CV_G^2]^{½}$

- *Other specifications based on biology include:*

  reference methods

  fixed limits for EQAS

  allowable difference between 2 methods.

Application of data

- *setting quality goals*
- ***assessing the significance of changes in serial results from an individual***
- *deciding the utility of conventional population based reference values*
- *other uses in quality management*

Reference change values

In order to decide whether a change is due to the patient improving or deteriorating, the "critical difference" or "reference change value" that would be expected due to inherent sources of variation must be exceeded

RCV depend on probability [Z], analytical [CV_A] and within-subject biological [CV_I] variation, if pre-analytical variation is minimized -

$$Change > RCV = 2^{1/2} * Z * [CV_A{}^2 + CV_I{}^2]^{1/2}$$

Reference change values

Each result has a dispersion dependent on -

probability [Z],analytical [CV_A] and biological [CV_I] variation -

First result - $Z * [CV_A^{\,2} + CV_I^{\,2}]^{1/2}$

Second result - $Z * [CV_A^{\,2} + CV_I^{\,2}]^{1/2}$

Total variation = [sum of squares]$^{1/2}$

$= \{Z^2 * [CV_A^{\,2} + CV_I^{\,2}] + Z^2 * [CV_A^{\,2} + CV_I^{\,2}]\}^{1/2}$

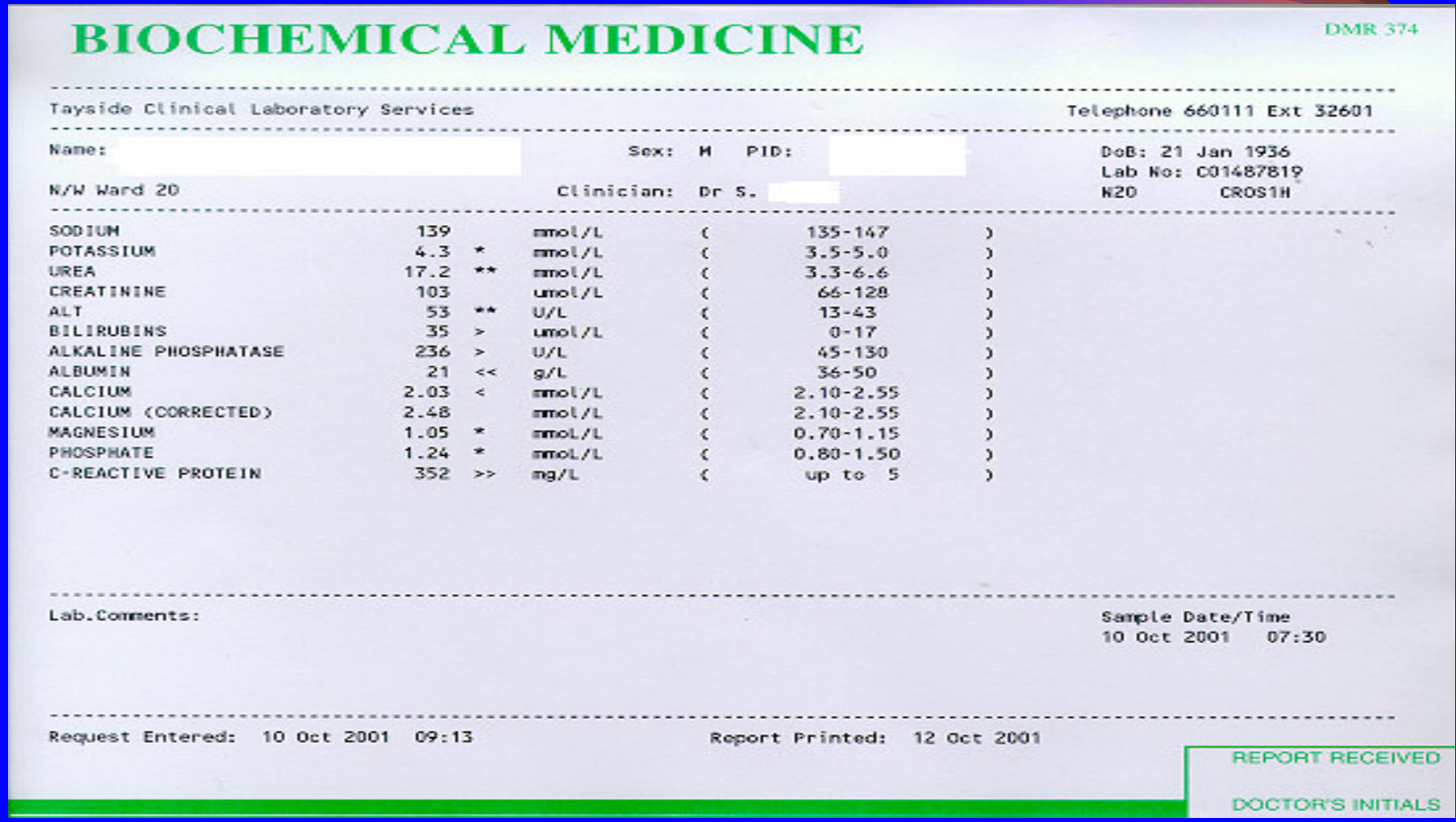
$= 2^{1/2} * Z * [CV_A^{\,2} + CV_I^{\,2}]^{1/2}$

Calculation of RCV

$$RCV = 2^{1/2} * Z * [CV_A^2 + CV_I^2]^{1/2}$$

- *use $2^{1/2}$ because we have 2 samples*
- *use 1.96 and 2.58 as the Z-scores for significant and highly significant respectively*
- *use CV_A from your own IQC program - at "clinically significant levels"*
- *use CV_I from the most recent data base available [as per Carmen Ricos and colleagues]*

An example laboratory report



BIOCHEMICAL MEDICINE

DMR 374

Tayside Clinical Laboratory Services — Telephone 660111 Ext 32601

Name: — Sex: M — PID: — DoB: 21 Jan 1936

Lab No: C01487819

N/W Ward 20 — Clinician: Dr S. — N20 — CROS1H

SODIUM	139		mmol/L	(	135-147)
POTASSIUM	4.3	*	mmol/L	(	3.5-5.0)
UREA	17.2	**	mmol/L	(	3.3-6.6)
CREATININE	103		umol/L	(	66-128)
ALT	53	**	U/L	(	13-43)
BILIRUBINS	35	>	umol/L	(	0-17)
ALKALINE PHOSPHATASE	236	>	U/L	(	45-130)
ALBUMIN	21	<<	g/L	(	36-50)
CALCIUM	2.03	<	mmol/L	(	2.10-2.55)
CALCIUM (CORRECTED)	2.48		mmol/L	(	2.10-2.55)
MAGNESIUM	1.05	*	mmol/L	(	0.70-1.15)
PHOSPHATE	1.24	*	mmol/L	(	0.80-1.50)
C-REACTIVE PROTEIN	352	>>	mg/L	(	up to 5)

Lab.Comments:

Sample Date/Time
10 Oct 2001 07:30

Request Entered: 10 Oct 2001 09:13

Report Printed: 12 Oct 2001

REPORT RECEIVED

DOCTOR'S INITIALS

Application of RCV

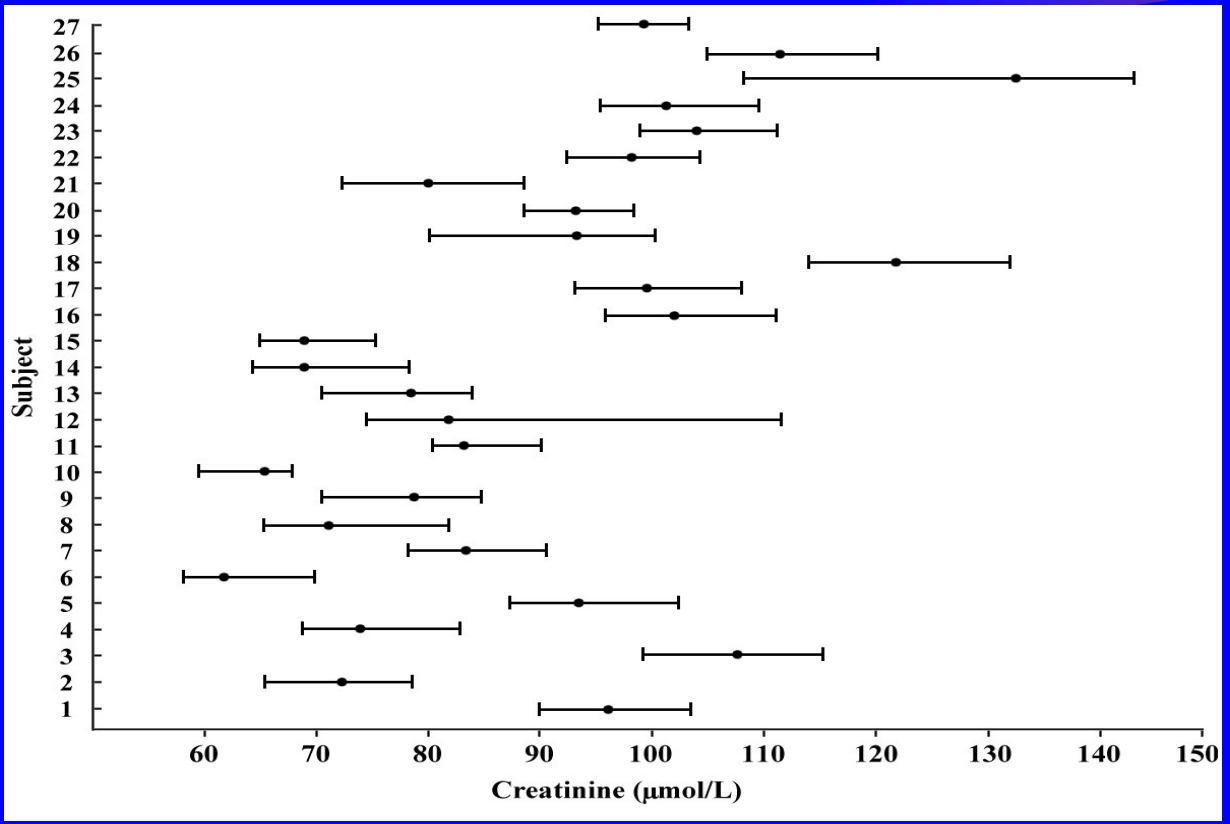
- *in reporting of results to aids clinical interpretation*
- *in auto-verification*
- *in delta checking*

This will be explored in Biological Variation 2 in more detail.

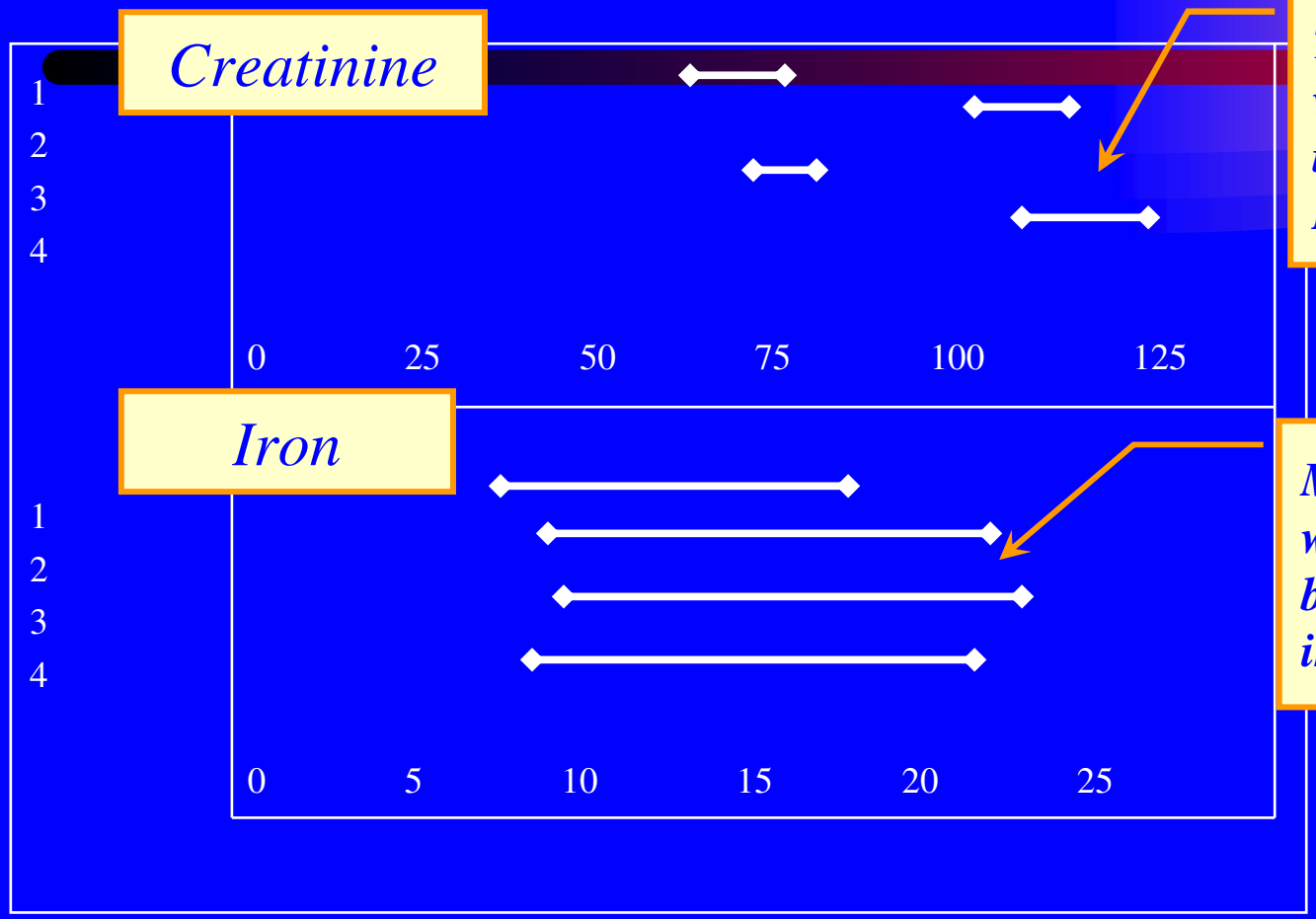
Application of data

- *setting quality goals*
- *assessing the significance of changes in serial results from an individual*
- *deciding the utility of conventional population based reference values*
- *other uses in quality management*

Individuality and reference values: creatinine



Ranges for 4 subjects for creatinine and iron



There is LITTLE variation within each individual but much BETWEEN individuals

MUCH variation within each individual but LITTLE between individuals

Index of individuality

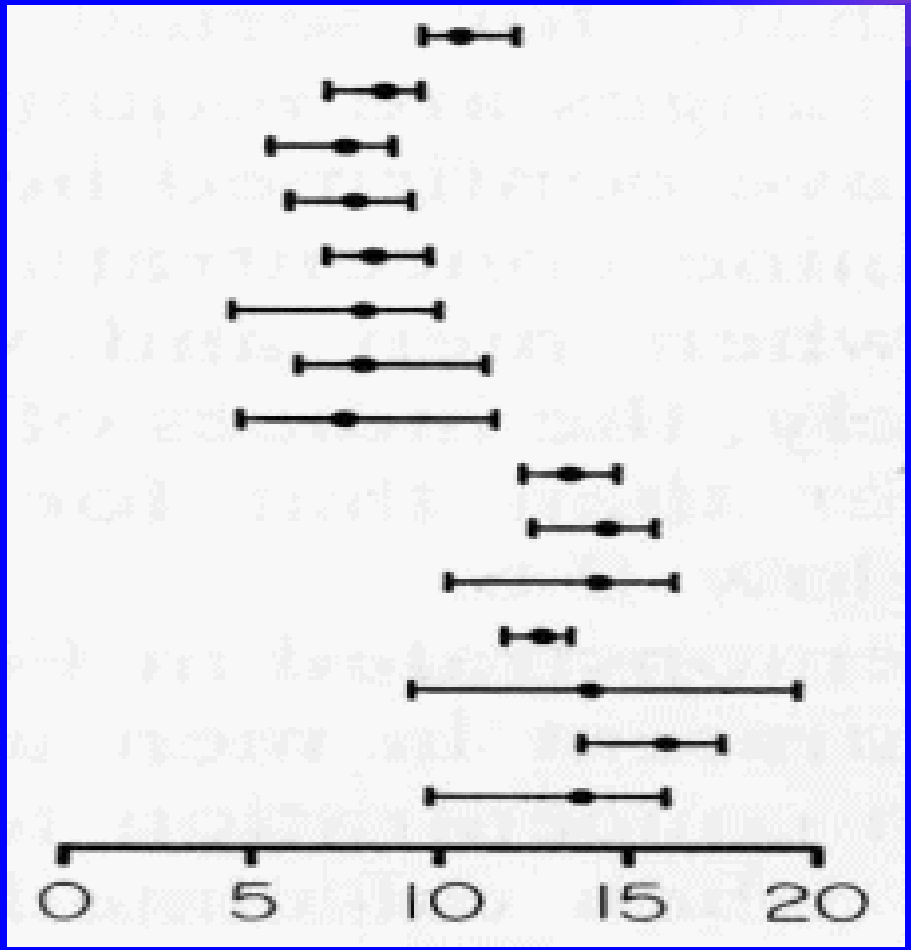
- *this is easily calculated as - CV_I/CV_G*
- *it is important to note that a low index of individuality means that the analyte has marked individuality whereas a high index of individuality means that the analyte has little individuality*
- *creatinine has II of 0.33, iron has II of 1.4*

Interpretation of index of individuality

- *When II is low [0.6], values for any individual span only a small part of the reference interval. Reference values will be of little utility, particularly when deciding whether changes in an individual have occurred. RCV preferred!*

- *In contrast, when II is high [1.4], values from an individual will cover much of distribution of the reference interval de. Thus, conventional reference values will be of significant value in many clinical settings.*

Ranges for 8 women and 7 men for urinary creatinine output



Individuality of urine creatinine

Group	*Within*	*Between*	*II*
Whole	*13.0*	*28.2*	*0.46*
Women	*15.7*	*11.0*	*1.42*
Men	*11.0*	*6.0*	*1.83*

This provides a scientific basis why stratification of reference values is often desirable.

Investigation of unusual proteins

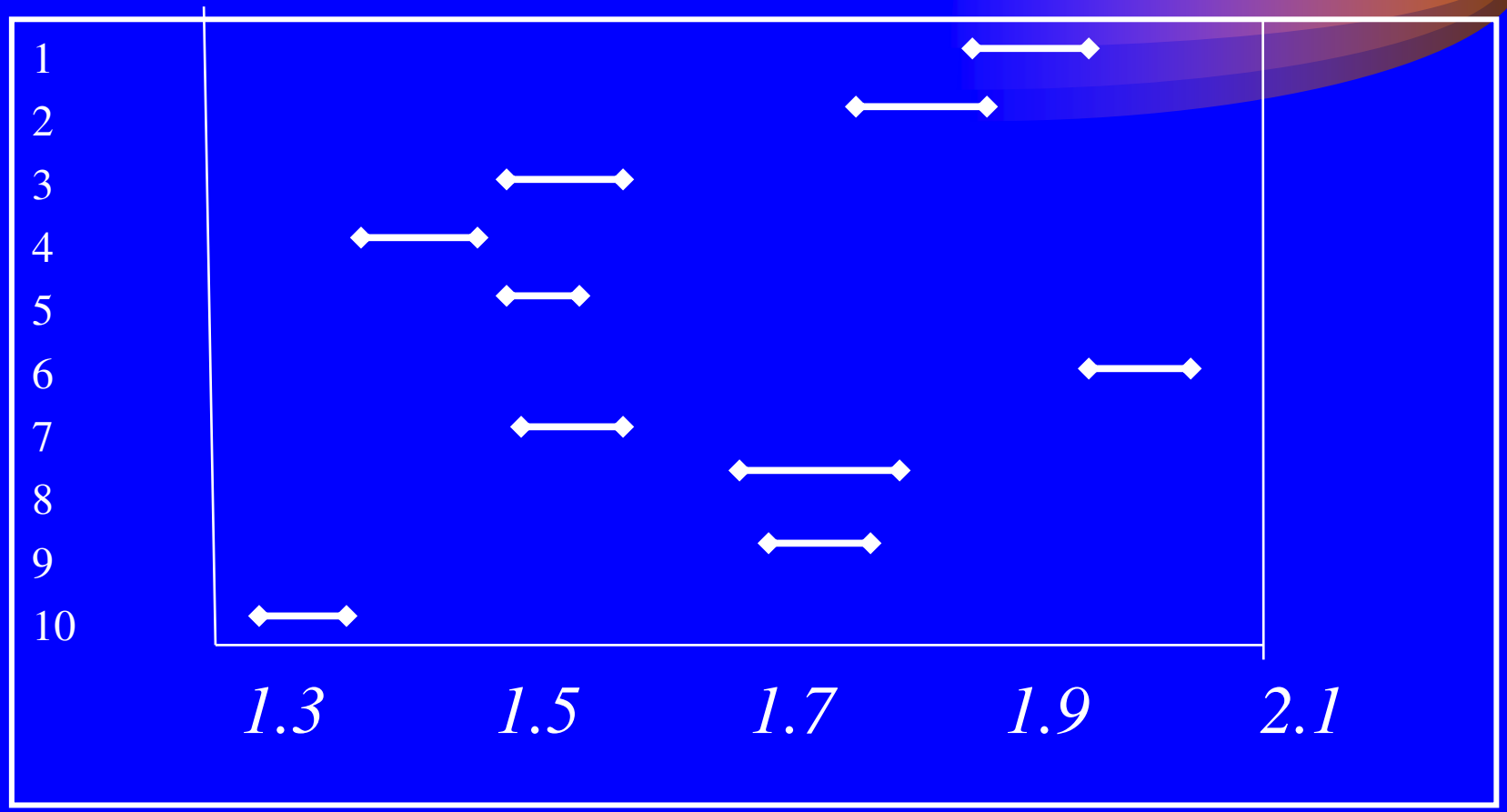
Is electrophoresis of serum to look for abnormal immunoglobulins an obsolete test? Proposal - measure IgG, IgA, IgM, kappa-chains [κ] and lamba-chains [λ] and calculate the heavy [IgG + IgA + IgM] to light chain [κ + λ] ratio and the κ/λ ratio to see if there was abnormal immunoglobulin and what type of protein was present.

Can biological variation data shed light on the clinical utility of this approach to diagnosis?

Individuality of immunoglobulins and derived ratios

Analyte	Within	Between	II
• IgG	4.4	13.0	0.34
• IgA	5.0	35.0	0.14
• IgM	5.9	48.5	0.12
• κ	4.8	15.3	0.31
• λ	4.8	17.3	0.28
• κ/λ ratio	0.7	12.1	0.06
• H/L ratio	4.2	4.8	0.87

Ranges for the κ/λ ratio



Individuality of hematological analytes

Analyte	Within	Between	II
• hemoglobin	2.8	6.6	0.42
• hematocrit	2.8	6.6	0.42
• MCV	1.3	4.8	0.27
• erythrocytes	3.2	6.1	0.52
• leukocytes	10.4	27.8	0.37
• platelets	9.1	21.9	0.42

Individuality

- *few analytes have II greater than 1.4*
- *most analytes have II less than 0.6*
- *reference values usually not of great utility - especially for monitoring*
- *hardly surprising that laboratory tests not good for case finding or screening*
- *stratification of reference values increases II*

Other uses of BV

Number of samples needed to ensure estimate is within a certain percentage of the homeostatic setting point with a predetermined probability –

$$n = (Z * [CV_A^2 + CV_I^2]^{1/2}/D)^2$$

where Z is the Z-score appropriate for the probability eg 1.96 for $P < 0.05$ and

D is the desired percentage closeness.

An example - cholesterol

- *within-subject variation CV_I= 6% maximum imprecision - NCEP CV_A= 3%*
- *estimate to be within 5% of the true homeostatic setting point with 95% probability*

  $n = (Z * [CV_A^2 + CV_I^2]^{½}/D)^2 = (1.96 * [3^2 + 6^2]^{½}/5)^2 = 7$

decreasing the number of samples required done by -

- *lowering probability*
- *increasing window of acceptability*
- *decreasing imprecision*

Selection of best sample

Selecting the best sample to collect: microalbumin

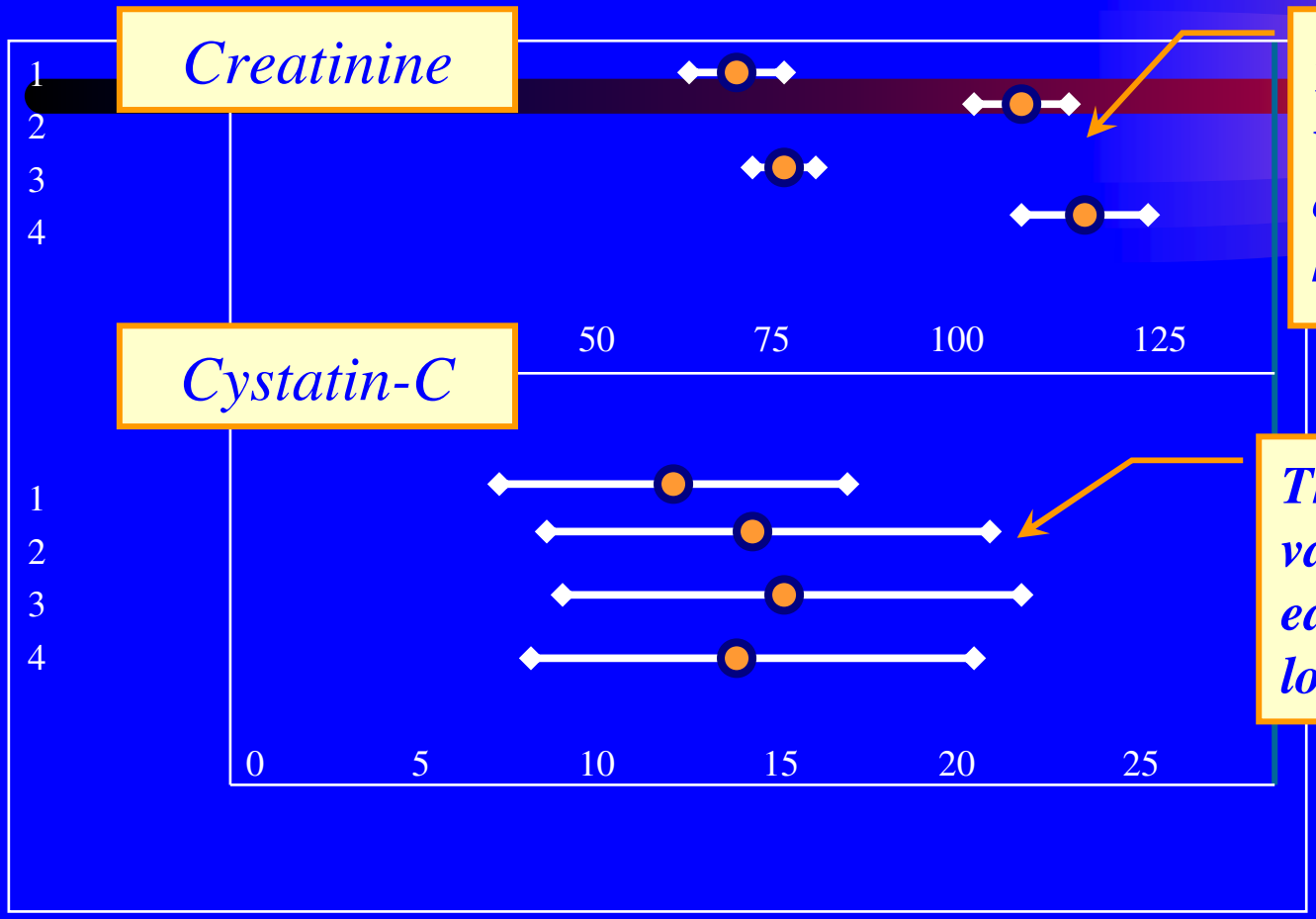
Sample type	CV_I	CV_G
First morning	*36*	*35*
Random spot	*86*	*61*
24 hour as concentration	*61*	*53*
24 hour as output	*70*	*55*

Comparison of tests

Analyte	CV_I	CV_G
cystatin-C	13.3	8.1
creatinine	4.9	18.2

consider individuality and RCV - cystatin-C will be better in diagnosis [less individual] but creatinine better in monitoring [smaller RCV].

Comparison of creatinine and cystatin-C



There is MUCH variation between each individual - high individuality

There is LITTLE variation between each individual - low individuality

Generation and application of data on biological variation

Usually not necessary to generate your own data- use existing database - then apply in -

- *setting quality specifications,*
- *assessing the significance of changes in serial results from an individual,*
- *deciding the utility of conventional population based reference values, and*
- *other uses in quality management.*