

# The value of the coefficient of variation in assessing repeat variation in ECG measurements

S. C. McLaughlin\*, T. C. Aitchison† and P. W. Macfarlane\*

\*Department of Medical Cardiology, †Department of Statistics, University of Glasgow, Glasgow, Scotland, U.K.

**Aims** The coefficient of variation is a popular measure for describing the amount of repeat variability present in ECG measurements from recording to recording. However, it can be misleading. The aim of the present study was to assess repeat variation (reclassification) in computer measured ECG criteria, i.e. positive to negative or vice versa, and compare this with the coefficient of variability.

**Methods and Results** Two ECGs were obtained from each of 295 patients, one day apart, and separately from a further 364 patients, several minutes apart. All patients were considered to be in a stable condition. Estimates of the coefficients of variation were obtained for a number of ECG parameters used in the diagnosis of left ventricular hypertrophy. Corresponding reclassification rates of relevant ECG criteria were also calculated.

Large coefficients of variation were observed in voltage parameters, e.g. R in V<sub>5</sub> (20% for day-to-day recordings

and 6% for minute-to-minute recordings) while the corresponding reclassification rates were 8% and 0% respectively. The repeat variation in the diagnosis of left ventricular hypertrophy was up to 5% for day-to-day recordings and up to 3% for minute-to-minute recordings based on several different criteria.

**Conclusion** A large coefficient of variation in a particular variable does not necessarily correspond to a high reclassification rate. A better measure of the impact of ECG variability for a particular measurement is obtained from its reclassification rate. In turn, this may have a minimal effect on the overall diagnosis of a particular abnormality.

(*Eur Heart J* 1998; 19: 342–351)

**Key Words:** Electrocardiogram, computer analysis, repeatability.

## Introduction

The problem of inconsistency between consecutive automated ECG interpretations is widely recognised<sup>[1,2]</sup>. There are several factors which contribute to this lack of agreement, a major contributor being the positioning of the electrodes<sup>[3]</sup>. When the electrodes are not correctly placed consistently, the resulting ECG measurements are unlikely to be compatible. Other contributory factors include the posture of the patient, proximity of ECG recording to mealtimes, inhalation of tobacco smoke and the recording equipment being used<sup>[4,5]</sup>.

It would be a tedious, if not an impossible task to ensure that all conditions remained constant while recording any given individual's ECG on consecutive occasions. However, it is important to appreciate that, although there may be substantial differences in ECG measurements leading to inconsistent computer interpre-

tations, such differences may not be visually indicative of the manifestation or the disappearance of a particular abnormality.

Consistency of the automated diagnosis of left ventricular hypertrophy has already been investigated in this laboratory and methods of improving the agreement in successive interpretations proposed<sup>[6]</sup>. The present paper assesses the ability of the coefficient of variation to quantify the effect of observed variability in those ECG parameters which are associated with the diagnosis of left ventricular hypertrophy in order to ascertain their susceptibility to variation. This metric has been used by various groups<sup>[7–9]</sup> to identify ECG parameters most susceptible to variation.

The aim of the present study was to examine the ability of the coefficient of variation to determine which ECG measurements are most likely to contribute to a lack of repeatability in the automated diagnosis of left ventricular hypertrophy. In addition, it was of interest to compare the coefficients of variation of ECG parameters with their corresponding reclassification rates, i.e. a change from positive to negative or vice versa, and hence the effect on the overall diagnosis of left ventricular hypertrophy. The primary interest lies in those ECG

Revision submitted 22 July 1997, and accepted 5 August 1997

*Correspondence:* Professor P. W. Macfarlane, University Department of Medical Cardiology, Glasgow Royal Infirmary, 10 Alexandra Parade, Glasgow G3 7ER, U.K.

**Table 1** Mean, maximum and minimum ages of the 295 patients on whom day-to-day ECGs were recorded

	Number	Mean age	Minimum age	Maximum age
Male	234	56	23	80
Female	61	58	29	76

parameters which are used by the Glasgow programme to diagnose left ventricular hypertrophy. However, coefficients of variation and reclassification rates are also calculated for several other commonly used ECG predictors of left ventricular hypertrophy.

## Methods

### Day-to-day ECGs

For the purposes of this study, 12-lead ECGs were recorded at least 24 h apart from 295 non-acute cardiac patients admitted for routine investigation to the cardiology wards of Glasgow Royal Infirmary over a 3-year period (Table 1). The patients were unselected, i.e. there was no attempt to select patients on the basis of history or echocardiographic findings that might suggest the presence of left ventricular hypertrophy. To represent realistic day-to-day recording conditions, no standardizing procedures were used. Thus, electrode positions were not marked, and no restrictions were placed on the recording technicians (i.e. the technician who recorded the ECG on the second occasion need not necessarily have been the same individual who recorded the initial ECG).

Table 1 illustrates that there is a predominance of males in the database. This is due to the fact that it is mainly men who are admitted to the University Department of Medical Cardiology for investigation, principally of ischaemic heart disease.

### Minute-to-minute ECGs

In order to obtain estimates of ECG measurement variability which could not be attributed to electrode positioning, ECGs were recorded twice within a few

**Table 2** Mean, maximum and minimum ages of the 364 patients on whom minute-to-minute ECGs were recorded

	Number	Mean age	Minimum age	Maximum age
Male	265	57	25	92
Female	99	58	36	77

same three year period, and none of whom was acutely ill (see Table 2). As for the previous group, the patients were unselected.

All of the ECGs were recorded using either a locally-designed and built electrocardiograph<sup>[10]</sup> or a MINGOREC 4 (Siemens-Elema, Stockholm, Sweden), both acquiring ECG leads simultaneously and digitizing them at 500 samples per second.

### Criteria assessment

The coefficient of variation is, to some extent, dependent on the magnitude of the measurement in question. Based on two repeated measures from each individual, this measure of reproducibility is calculated as the standard deviation of the differences between the repeated measurements divided by the average of the averages of the repeated measurements and is quoted as a percentage. If  $x_1$  and  $x_2$  are the first and second of the ECG parameters respectively, then the coefficient of variation (CV) may be calculated as:

$$CV = \frac{sd(x_{1i} - x_{2i})}{\left( \sum_{i=1}^n \frac{x_{1i} + x_{2i}}{2} \right) / n}$$

where  $n$  is the total number of repeated ECGs,  $i=1, \dots, n$  and  $sd$ =standard deviation.

The reclassification rate (RR) is defined as a change in a criterion from being positive to being negative, or vice versa. If  $b$  is the critical threshold value associated with ECG parameter  $x$ , and  $x_1$  and  $x_2$  are the first and second of the ECG parameters, respectively, then

$$RR = \frac{\text{no. of pairs of ECG recordings with } [(x_1 > b) \text{ and } (x_2 \leq b)] \text{ or } [(x_1 \leq b) \text{ and } (x_2 > b)]}{\text{Total number of repeated ECG recordings}}$$

minutes (without removal and subsequent replacement of the electrodes) from each of 364 cardiac patients who were admitted to Glasgow Royal Infirmary over the

It should be emphasized that the reclassification rate can be used to study individual criteria as well as overall classification. The latter would apply if  $x$

**Table 3** Brief outline of the Glasgow scoring system for the detection of left ventricular hypertrophy

Criterion	Number of points
Increased QRS voltage	≥ 2 points
ST-T changes	1–4 points
Increased P terminal force	2 points
Left axis deviation	2 points
Prolonged QRS duration	1 point
Delayed intrinsicoid deflection	1 point

Possible LVH=4 points, probable LVH=5 points, definite LVH=6 points or more.

represented an independent criterion for left ventricular hypertrophy, e.g. the Sokolow and Lyon index.

While calculation of coefficients of variation gives some insight into the problem of lack of repeatability, and to the magnitude of the variability in ECG measurements, it cannot provide information on the direct effects of such variation in ECG measurements upon the diagnosis of left ventricular hypertrophy. In order to investigate the reproducibility of the automated diagnosis, reclassification rates were calculated for each criterion contributing to a final outcome of left ventricular hypertrophy as well as for the overall diagnosis of left ventricular hypertrophy in the Glasgow programme.

Romhilt and Estes developed a system which relied both on voltage and non-voltage criteria for the ECG diagnosis of left ventricular hypertrophy<sup>[11]</sup>. A modified version of this point scoring system is used in the Glasgow laboratory<sup>[12]</sup> and a brief outline is given in Table 3. The final left ventricular hypertrophy score is calculated by summing the points recorded by the six individual components and a diagnosis of left ventricular hypertrophy is made if the total number of points ≥ 4. For some parameters, e.g. QRS voltage, criteria are categorized by sex and age<sup>[6]</sup>. Details of threshold values relating to the individual criteria are somewhat involved and can be found elsewhere<sup>[13]</sup>. As an example, the threshold values for R wave amplitude in aVL are 1.1, 1.2 and 1.3 mV for males aged 17–29, 30–39 and over 40 years respectively. Corresponding values for females are 0.9, 1.0 and 1.2 mV respectively (cf. Fig. 1(b)).

The problem of automated wave relabelling on repeat recordings was alleviated by the software always selecting the maximum of R or R' in the lateral leads, and similarly the maximum of S or S' in the septal leads. If a Qr complex changed into an rSr' configuration, then this would contribute to variation in the S wave reclassification in V<sub>1</sub> or V<sub>2</sub>.

Coefficients of variation and reclassification rates were also calculated for the Cornell voltage (RaVL+SV<sub>3</sub>) and for the 12-lead QRS voltage sum<sup>[14]</sup>.

It has also been demonstrated that the product of QRS duration and voltage can enhance the ECG diagnosis of left ventricular hypertrophy<sup>[15]</sup>. Thus, it was of interest to examine the repeatability of the 12-lead QRS product and the Cornell product. The Cornell voltage is a gender-specific criterion, with threshold values of 2800 µV and 2000 µV, respectively, for males and females. The threshold values used for the 12-lead QRS sum and for the voltage duration products were taken from Farb *et al.*<sup>[7]</sup>.

## Results

### Patient characteristics

The severity of left ventricular hypertrophy in the population studied was unknown, i.e. no echocardiograms were undertaken. However, an appreciation of the likely prevalence of left ventricular hypertrophy in the study cohort can be obtained from the histograms of the R wave amplitude in lead aVL, the Sokolow-Lyon index, and the Glasgow left ventricular hypertrophy score (Fig. 1). Histograms are provided for the 295 day 1 ECG recordings. The distributions of these parameters are similar for the 364 minute 1 ECG recordings.

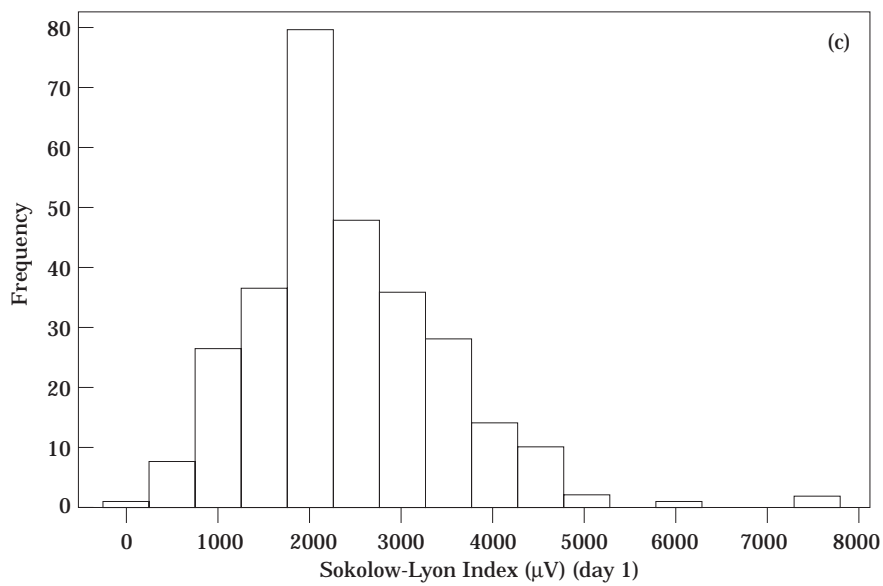
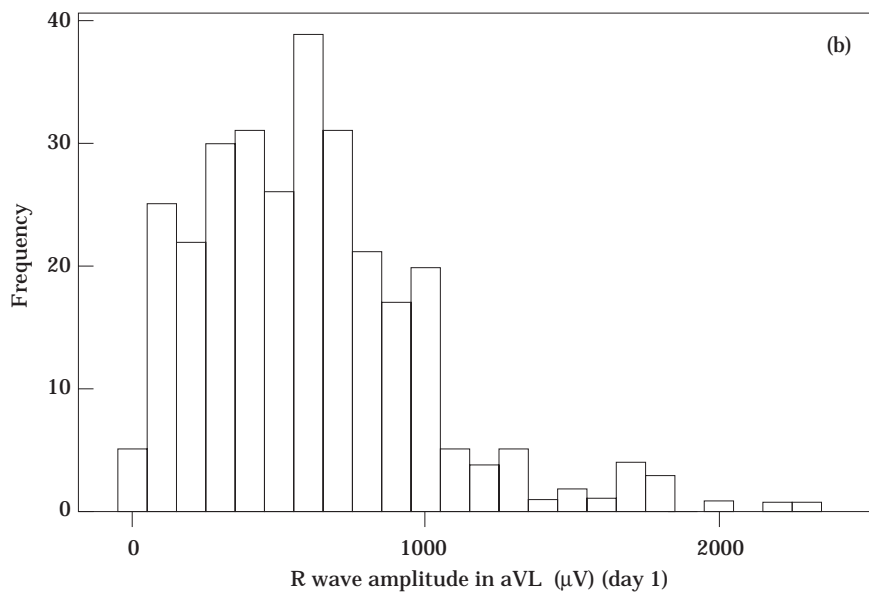
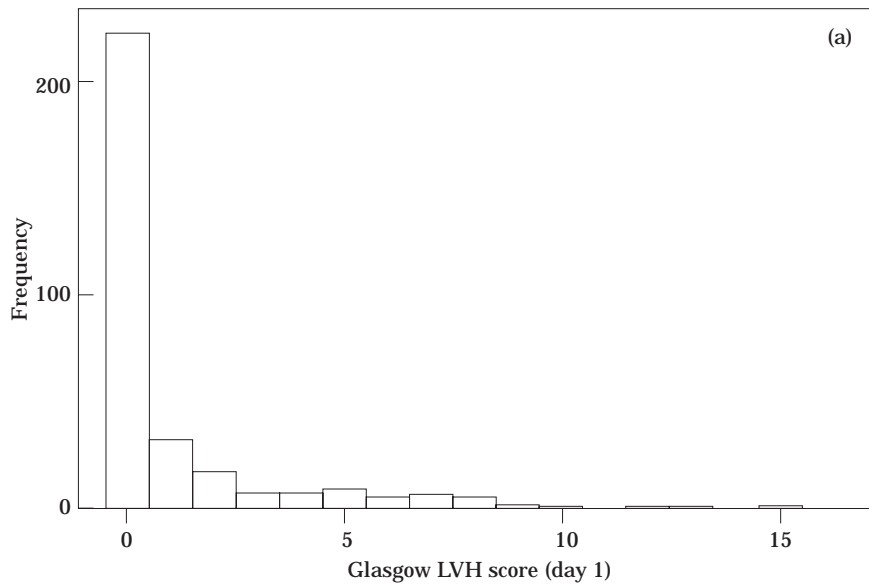
### Glasgow scoring system for left ventricular hypertrophy

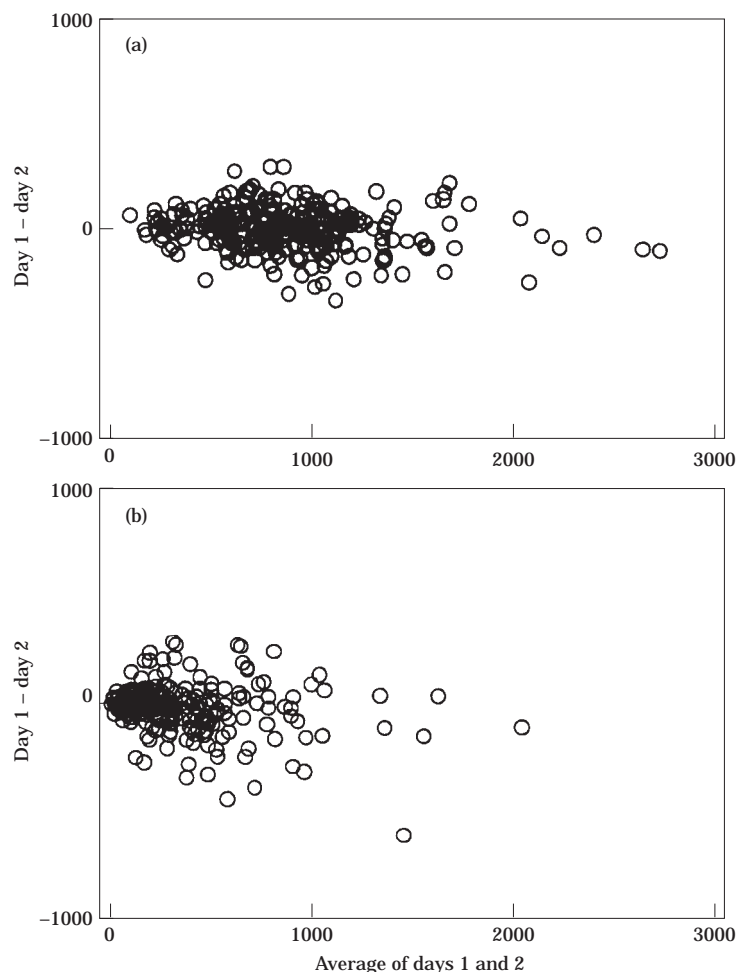
The data from both the day-to-day and minute-to-minute ECG recordings were used separately to obtain estimates of the coefficients of variation and reclassification rates for the ECG parameters used in the diagnosis of left ventricular hypertrophy in the Glasgow programme.

#### Coefficient of variation

**Voltage criteria.** Estimates of the day-to-day coefficients of variation for the voltage criteria are given in Table 4(a) both for single lead and combined lead voltage measurements. These values range from 13% for the amplitude of the R wave in lead I (RI) to 51% for the amplitude of the terminal negative P component in lead V<sub>1</sub>. A plot of the difference between the day 1 and day 2 measurements vs their average is given for the R wave amplitudes in leads I and III, respectively (Fig. 2(a) RI; Fig. 2(b) RIII), in order to provide a visual appreciation of the amounts of variability which can be expected from specific limb lead amplitudes. The data have been plotted on the same scale for both graphs. It is apparent that the differences between day 1 and day 2 readings for RIII are greater than those observed for RI, with respect to their corresponding mean values, resulting in a higher coefficient of variation.

**Figure 1(a)** Frequency histogram of the Glasgow left ventricular hypertrophy (LVH) score (day 1 ECG recordings). **(b)** Frequency histogram of the R wave amplitude in lead aVL (day 1 ECG recordings). **(c)** Frequency histogram of the Sokolow-Lyon Index (day 1 ECG recordings).





**Figure 2(a) Amplitude of the R wave in lead I (coefficient of variation=13%). (Day 1 – day 2) vs (day 1 + day 2)/2. (b) Amplitude of the R wave in lead III (coefficient of variation=36%). (Day 1 – day 2) vs (day 1 + day 2)/2.**

As expected, minute-to-minute variability is lower than day-to-day variability. However, values range from 4% for the amplitude of the S wave in leads  $V_1$  and  $V_2$  ( $SV_1$ ,  $SV_2$ ) to 34% for the terminal negative component of the P wave in lead  $V_1$  (Table 4(b)).

The presence of left ventricular strain, which consists of ST depression and T wave inversion, contributes significantly to a final diagnosis of left ventricular hypertrophy. The ST segment is a notorious area for measurement errors, and is therefore likely to be prone to lack of repeatability. Estimates of the day-to-day coefficients of variation for the ST and T wave measurements were evaluated, and ranged from 37% for the amplitude of the positive component of the T wave in lead I to 72% for the amplitude of the ST segment in lead I. The minute-to-minute values ranged from 18% for the amplitude of the positive component of the T wave in lead  $V_5$  to 64% for the amplitude of the ST segment in lead aVL.

**Non-voltage criteria.** The Glasgow programme uses a modified version of the Romhilt-Estes scoring system in

order to identify left ventricular hypertrophy<sup>[12]</sup>, using both voltage and non-voltage criteria. Estimates of the day-to-day coefficients of variation for the ECG parameters involved in the determination of abnormal P terminal force, prolonged QRS duration and delayed intrinsicoid deflection are given in Table 5(a). They range from 13% for the duration of the QRS complex in leads  $V_5$  and  $V_6$  to 62% for the duration of the negative component of the P wave in lead  $V_1$ . Plots of the difference between the day 1 and day 2 measurements vs their average for both ECG parameters are provided in Fig. 3.

Estimates of the minute-to-minute coefficients of variation (Table 5(b)) are somewhat lower than the corresponding day-to-day values (Table 5(a)). They range from 9% for the QRS duration in leads  $V_5$  and  $V_6$  to 49% for the duration of the negative component of the P wave in lead  $V_1$ .

#### *Reclassification rate*

Reclassification rates for each component contributing to the diagnosis of left ventricular hypertrophy can be

**Table 4(a) Coefficients of variation of ECG voltage measurements used in the diagnosis of left ventricular hypertrophy in the Glasgow programme (day-to-day ECG recordings)**

Measurement	Coefficient of variation (%)
RI	13
Sokolow-Lyon Index	15
SV <sub>1</sub>	16
RV <sub>6</sub>	17
RaVL	19
RV <sub>5</sub>	20
SV <sub>2</sub>	20
SIII	24
Lewis Index	28
SI	32
RIII	36
P amplitude V <sub>1</sub>	51

**Table 4(b) Coefficients of variation of ECG voltage measurements used in the diagnosis of left ventricular hypertrophy in the Glasgow programme (minute-to-minute ECG recordings)**

Measurement	Coefficient of variation (%)
SV <sub>1</sub>	4
SV <sub>2</sub>	4
RI	5
RV <sub>5</sub>	6
RaVL	7
RV <sub>6</sub>	7
SI	9
SIII	10
Lewis Index	12
Sokolow-Lyon Index	12
RIII	12
P amplitude V <sub>1</sub>	34

seen in Tables 6(a) and (b). Values are given for day-to-day recordings, and for minute-to-minute recordings. The ECG voltage measurements do not contribute to the diagnosis of left ventricular hypertrophy on an individual basis. They are combined in such a way that if any one of the voltage measurements exceeds its pre-defined threshold value, then the final left ventricular hypertrophy score is increased by an appropriate amount. Thus, although reclassification rates are provided for each of the individual ECG voltage measurements, it is the combined voltage reclassification rate which is of most interest.

The reclassification rates themselves do not imply that a change in diagnosis between recordings has occurred. In this context they are being used to describe a change in status of each individual criterion contributing to a diagnosis of left ventricular hypertrophy. Thus, the day-to-day reclassification rate for the overall voltage in Table 6(a) implies that abnormal QRS

**Table 5(a) Coefficients of variation of ECG non-voltage measurements used in the diagnosis of left ventricular hypertrophy in the Glasgow programme (day-to-day ECG recordings)**

Measurement	Coefficient of variation (%)
QRS duration V <sub>5</sub>	13
QRS duration V <sub>6</sub>	13
Intrinsicoid deflection V <sub>6</sub>	14
Intrinsicoid deflection V <sub>5</sub>	18
P duration V <sub>1</sub>	62

**Table 5(b) Coefficients of variation of ECG non-voltage measurements used in the diagnosis of left ventricular hypertrophy in the Glasgow programme (minute-to-minute ECG recordings)**

Measurement	Coefficient of variation (%)
QRS duration V <sub>5</sub>	9
QRS duration V <sub>6</sub>	9
Intrinsicoid deflection V <sub>6</sub>	13
Intrinsicoid deflection V <sub>5</sub>	15
P duration V <sub>1</sub>	47

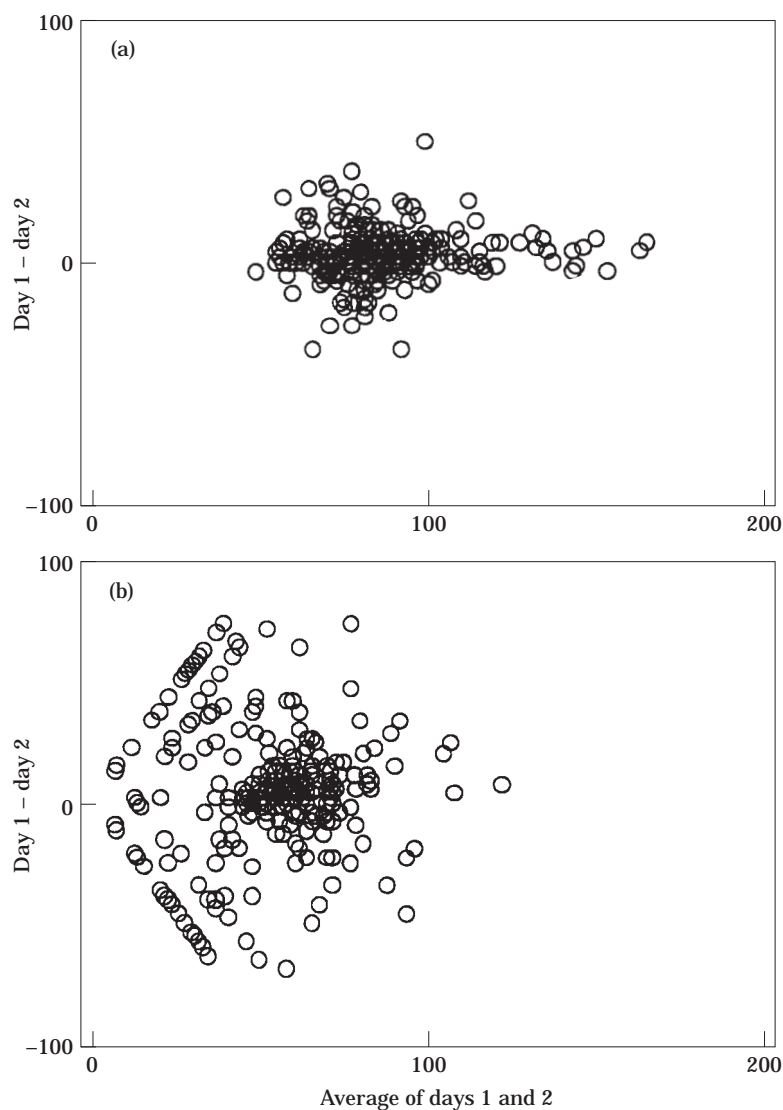
voltages were present in at least one of the QRS amplitude criteria in only one of a pair of ECGs in 14% of cases. In addition, the minute-to-minute reclassification rate for P terminal force (Table 6(b)) indicates that in 2% of cases, an abnormal P terminal force was observed in one ECG and not in the other.

The overall left ventricular hypertrophy reclassification rates do imply that there is a change in the diagnosis of left ventricular hypertrophy, i.e. from present to absent or vice versa from one recording to the next. This occurs in 5% of the day-to-day ECG recordings and in 3% of the minute-to-minute ECG recordings.

There is no suggestion of any substantial association between coefficients of variation and reclassification rate values in either Table 6(a) or (b) and this is clearly confirmed in Fig. 4.

### *Cornell-based criteria for left ventricular hypertrophy*

Table 7 provides coefficients of variation and corresponding reclassification rates for several more recently developed ECG criteria used in the identification of left ventricular hypertrophy. These include the Cornell voltage (R in aVL+S in V<sub>3</sub>), the Cornell product (the Cornell voltage multiplied by the QRS duration), the 12 lead QRS sum (sum of the QRS amplitudes in all 12 leads) and the 12 lead QRS product (the 12 lead QRS sum multiplied by the QRS duration). Again there is no



**Figure 3** (a) Duration of QRS complex in  $V_5$  (coefficient of variation=13%). (Day 1 – day 2) vs (day 1 + day 2)/2. (b) Duration of negative component of the P wave in lead  $V_1$  (coefficient of variation=62%). (Day 1 – day 2) vs (day 1 + day 2)/2.

evidence of a relationship between the coefficient of variation and the reclassification rate. Since each of these criteria is used on an individual basis to diagnose left ventricular hypertrophy, the term reclassification rates is appropriate.

## Discussion

The coefficients of variability were calculated for many of the ECG parameters used in the diagnosis of left ventricular hypertrophy. It was found that, in certain cases, repeatability was very poor. For example, the amplitude and duration of the negative component of the P wave in lead  $V_1$  displayed day-to-day coefficients of variation of 62% and 51% respectively (see Tables

4(a) and 5(a)). The corresponding minute-to-minute values were also large, being 47% and 34%, respectively (see Tables 4(b) and 5(b)), even when the recordings were obtained without removal and subsequent replacement of the electrodes. In addition, values for the ST and T wave measurements in the lateral leads I, aVL,  $V_5$  and  $V_6$  indicated poor reproducibility.

Based on the coefficient of variation, the Lewis Index performs poorly when compared to the Sokolow–Lyon Index using the day-to-day recordings (15% vs 29%). One reason for this is the apparent lack of reproducibility in the measurement of the R wave in lead III which, in terms of day-to-day repeatability, is substantially worse than other single lead voltage measurements with a coefficient of 36% (see Table 4). Repeatability of the S wave in leads I and III is also

**Table 6(a) Day-to-day ECGs. Reclassification rates and coefficients of variation for criteria relating to the automated diagnosis of left ventricular hypertrophy in the Glasgow programme**

Measurement	Reclassification rate (%)	Coefficient of variation (%)
Intrinsicoid deflection $V_5/V_6$	1	15
Sokolow-Lyon Index	1	15
P terminal force	2	55
R in lead I	4	13
S in lead $V_1/V_2$	4	18
R in lead aVL	5	19
Lewis Index	7	29
R in lead $V_5/V_6$	8	18
QRS duration $V_5/V_6$	8	13
Left ventricular strain	2	N/A
Combined voltage	14	N/A
Overall left ventricular hypertrophy	5	N/A

Rates are given in ascending order where corresponding coefficients of variation are available.

**Table 6(b) Minute-to-minute ECGs. Reclassification rates and coefficients of variation for criteria relating to the automated diagnosis of left ventricular hypertrophy in the Glasgow programme**

Measurement	Reclassification rate (%)	Coefficient of variation (%)
Intrinsicoid deflection $V_5/V_6$	0	14
R in lead $V_5/V_6$	0	7
R in lead I	1	5
S in lead $V_1/V_2$	1	4
R in lead aVL	1	7
Sokolow-Lyon Index	1	12
P terminal force	2	40
Lewis Index	2	12
QRS duration $V_5/V_6$	7	9
Left ventricular strain	2	N/A
Combined voltage	1	N/A
Overall left ventricular hypertrophy	3	N/A

Rates are given in ascending order where corresponding coefficients of variation are available.

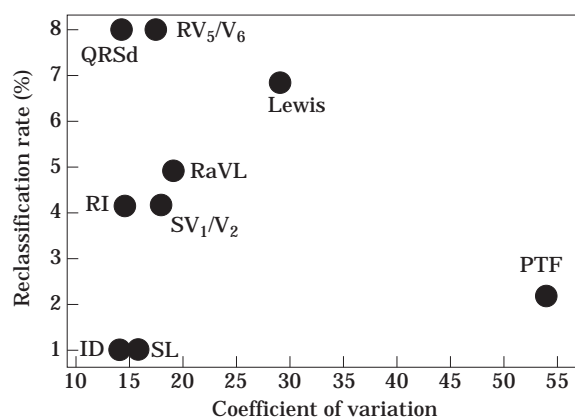
poor, with values of 32% and 24%. Other studies have also demonstrated relatively high day-to-day coefficients of variation of 39%<sup>[7]</sup> and 35.4%<sup>[8]</sup> for SIII.

The fact that the coefficient of variation associated with the R wave in lead III (36%) is three times that associated with the R wave in lead I (13%) may be attributed, in the majority of cases, to the amplitude of the R wave in lead I being greater than that in III in this particular group of individuals. It is interesting to note that van den Hoogen *et al.* made a similar observation (CV of RI=9.2%, CV for RIII=29.4%)<sup>[8]</sup>. In general, small measurements will tend to produce higher coefficients than larger measurements. The QRS axis on day 1 was less than 30° in 187 of the 295 pairs of ECGs which were recorded. In this particular subset, the amplitude of the R wave in lead I was substantially larger than the amplitude of the R wave in lead III (median values of 933  $\mu$ V and 149  $\mu$ V respectively) and the corresponding coefficients of variation were 11% and 42%. In the cases

where the QRS axis was in excess of 30° on day 1, the values for lead I and lead III were 19% and 27%, respectively. The closer agreement in values is due to the fact that the median values for the amplitudes of the R wave in leads I and III were more comparable in this subgroup, being 641  $\mu$ V and 443  $\mu$ V, respectively. Thus, results are highly dependent on the population studied.

When the coefficient of variation was considered, the Lewis Index was the poorest of the voltage criteria both from day-to-day and from minute-to-minute (see Table 4). This suggests that its influence on reclassification rate might be high. However, it is interesting to note that the 'more robust' maximum R wave amplitude in lead  $V_5$  or  $V_6$ , in terms of the coefficient of variation, produces a slightly worse reclassification rate than the Lewis Index (8% compared with 7% for day-to-day ECG recordings). The Lewis index remains the worst of the minute-to-minute voltage measurements in terms of reclassification rates, although the observed rate of 2% is





**Figure 4** Coefficient of variation vs reclassification rate for voltage and non-voltage criteria.

**Table 7(a) Day-to-day ECGs. Reclassification rates and coefficients of variation for commonly used ECG predictors of left ventricular hypertrophy**

Measurement	Reclassification rate (%)	Coefficient of variation (%)
12 lead QRS sum	2	7
Cornell voltage	2	17
Cornell product	2	20
12 lead QRS product	5	11

Reclassification rates are given in ascending order.

**Table 7(b) Minute-to-minute ECGs. Reclassification rates and coefficients of variation for commonly used ECG predictors of left ventricular hypertrophy**

Measurement	Reclassification rate (%)	Coefficient of variation (%)
12 lead QRS sum	1	3
12 lead QRS product	1	7
Cornell voltage	2	4
Cornell product	2	8

Reclassification rates are given in ascending order.

lower than the substantial overall coefficient of variation of 12% originally suggested.

Changes in the Sokolow–Lyon Index produced very few classifications either from day-to-day or from minute-to-minute, with both exhibiting rates of 1%. This is in contrast with a previous study which observed day-to-day and minute-to-minute reclassification rates of 10% and 5% respectively<sup>[8]</sup>.

The coefficients of variation calculated for the ST and T wave measurements were poor. However, the impact on reclassification is less striking. In 2% of the day-to-day recordings and in 2% of the minute-to-minute recordings, left ventricular strain was reported on one occasion and not on the other. However, the presence of left ventricular strain greatly affects the diagnosis

of left ventricular hypertrophy and contributes a substantial amount to the final left ventricular hypertrophy score. Therefore, any discrepancies for this category, no matter how few, can have an influence on the reproducibility of the diagnosis of left ventricular hypertrophy.

As discussed previously, both the amplitude and the duration of the negative component of the P wave in lead V<sub>1</sub> were shown to exhibit high coefficients of variation. However, the day-to-day and minute-to-minute reclassification rates of 2% were low. The criterion relating to an abnormal P terminal force in lead V<sub>1</sub> requires that both the amplitude and the duration exceed predefined critical values. This helps overall stability since, if the amplitude fluctuates enormously from recording to recording but the duration remains safely on one side of its critical threshold value on both occasions, then the overall decision on the presence or absence of an abnormal P terminal force would not alter. Therefore, it is feasible that two quite differing amplitudes could provide consistent interpretations, given that the duration remained relatively stable, or vice versa.

The QRS duration in leads V<sub>5</sub> and V<sub>6</sub> displays rather poor reclassification rates of 8% (day-to-day) and 7% (minute-to-minute). The low coefficients of variation associated with these measurements would seem to suggest good reproducibility. However, a difference in QRS duration of just 1 ms may be enough to cause a reclassification if the initial reading was of the order of 100 ms (the critical value). A sizeable proportion of the cohort of patients (19% for day-to-day recordings, 20% for minute-to-minute recordings) exhibited a QRS duration in excess of 90 ms in lead V<sub>5</sub> or lead V<sub>6</sub> on at least one occasion. This would suggest that there is a reasonable chance for reclassifications to arise, particularly in those cases where the QRS duration is very close to the critical value of 100 ms. However, this particular component of the left ventricular hypertrophy criteria does not influence the final outcome as much as, for example, presence of left ventricular strain or an abnormal P terminal force. Thus, the effects on the reproducibility of the diagnosis of left ventricular hypertrophy are likely to be minimal, as indicated in Table 6.

The intrinsicoid deflection displays very low reclassification rates from day-to-day (1%) and from minute-to-minute (0%) and, in isolation, is not likely to cause many problems in terms of repeatability.

It is important to note that reclassification associated with each of the six individual components of the Glasgow left ventricular hypertrophy scoring system does not imply that the overall classification of left ventricular hypertrophy has changed. The fact that the number of points scored by one of the six components differs from recording to recording may result in the overall left ventricular hypertrophy score changing for example from 0 to 1 (no left ventricular hypertrophy on both occasions) or from 6 to 7 (left ventricular hypertrophy on both occasions). Thus, in such cases, the computer diagnosis of left ventricular hypertrophy will be consistent from recording to recording.

Reclassification rates for the overall diagnosis of left ventricular hypertrophy from day-to-day and from minute-to-minute are provided (Table 6). The Glasgow programme produces reports of no left ventricular hypertrophy, possible, probable or definite left ventricular hypertrophy depending on the value of the left ventricular hypertrophy score<sup>[11]</sup>. If outcomes of possible, probable and definite left ventricular hypertrophy are taken to be positive, and a diagnosis of no left ventricular hypertrophy is taken to be negative, then the corresponding overall left ventricular hypertrophy reclassification rates are 5% (day-to-day) and 3% (minute-to-minute).

It is important to remember that, no matter how small the difference in two ECG measurements from recording to recording, inconsistent interpretations may arise as a result. If such differences cross a critical threshold value separating 'normal' from 'abnormal', then the magnitude of the discrepancy is irrelevant. To overcome this problem, we have commenced work on a new approach to scoring<sup>[16]</sup>.

The repeatability of a variety of other widely used criteria for the ECG identification of left ventricular hypertrophy was also examined (Table 7). The day-to-day and minute-to-minute coefficients of variation for the Cornell voltage were 17% and 4% respectively. Other studies have observed day-to-day and minute-to-minute values of 24.8%, 2.9%<sup>[7]</sup> and 10.3%, 4.6%<sup>[9]</sup>, respectively. Reclassification rates for the combined ECG voltage measurements of Cornell voltage and 12 lead QRS sum were comparable to those observed for the Sokolow Lyon index. For example, day-to-day reclassification rates were 2%, 2% and 1% respectively, while minute-to-minute reclassification rates were 2%, 1% and 1% respectively.

## Conclusions

ECG measurements which demonstrate substantial coefficients of variation may not have a significant impact on the repeatability of the automated diagnosis of a particular abnormality. Large values of the coefficient of variation are more likely to be associated with ECG measurements of small magnitude. In most situations, identification of a particular cardiac abnormality is usually associated with significant increases in ECG measurements, whether they be amplitudes or durations. Examples of this are the association of high voltages with cases of left ventricular hypertrophy, and prolonged durations in cases where conduction disturbances exist. It is not until ECG measurements are in close proximity to critical threshold values that lack of repeatability of interpretation becomes a problem.

The magnitude of the coefficient of variation values calculated from repeated ECG measurements should not be interpreted as a measure of suitability for the exclusion of a particular parameter from a

diagnostic decision on the basis of lack of repeatability. Coefficients of variation cannot be used to predict reclassification rates. It is the variability in the vicinity of any critical value in a diagnostic programme using threshold-based criteria which will determine the influence of day-to-day (or minute-to-minute) variability of a particular parameter upon the diagnostic programme.

## References

- [1] Machado HB, Emilia M, Silva P, Lourenco L. Reproducibility of diagnostic interpretative statements in a leading ECG computerised system. *Proceedings of Melecon 1991*; 1: 793-95.
- [2] Tuinstra CL. Reproducibility of ECG interpretation statements. In: Willems JL, van Bommel JH, Zywiets C, eds. *Computer ECG analysis: Towards Standardisation*. Amsterdam: North Holland 1986: 341-6.
- [3] Willems JL, Poblete PF, Pipberger HV. Day-to-day variation of the normal orthogonal electrocardiogram and vectorcardiogram. *Circulation* 1972; 45: 1057-64.
- [4] Graybiel A, Starr RS, White PD. Electrocardiographic changes following the inhalation of tobacco smoke. *Am Heart J* 1938; 15: 89-99.
- [5] Sears G, Manning GW. Routine Electrocardiography: Postprandial T-wave changes. *Am Heart J* 1958; 56: 591-7.
- [6] McLaughlin SC, Aitchison TC, Macfarlane PW. Improved repeatability of 12 lead ECG analysis using continuous scoring techniques. *J Electrocardiol* 1993; 26 (Suppl): 101-7.
- [7] Farb A, Devereux RB, Kligfield P. Day-to-day variability of voltage measurements used in electrocardiographic criteria for left ventricular hypertrophy. *J Am Coll Cardiol* 1990; 15: 618-23.
- [8] van den Hoogen JPH, Mol WH, Kowssolea A, van-Ree JW, Thien T, van Weel C. Reproducibility of electrocardiographic criteria for left ventricular hypertrophy in hypertensive patients in general practice. *Eur Heart J* 1992; 13: 1606-10.
- [9] de Bruyne MC, Kors JA, Hoes AW, Grobbee DE, van Bommel JH. Minute-to-minute and day-to-day variability of electrocardiographic measurements in an unselected population of older adults. In: van Oosterom A, Oostendorp TF, Uijen GJH, eds. *Proceedings of the 22nd International Congress on Electrocardiography*, Nijmegen, The Netherlands, 1995: 232-33.
- [10] Watts MP, Shoat DB. Trends in electrocardiographic design. *J Institution Electronic Radio Eng* 1987; 57: 140-50.
- [11] Romhilt DW, Estes EH. A point scoring system for the electrocardiographic diagnosis of left ventricular hypertrophy. *Am Heart J* 1968; 75: 752-8.
- [12] Huwez FU. Electrocardiography of the left ventricle in coronary artery disease and hypertrophy. PhD Thesis, University of Glasgow, 1990.
- [13] Milliken J, Macfarlane PW, Lawrie TDV. Enlargement and hypertrophy. In: Macfarlane PW, Lawrie TDV, eds. *Comprehensive Electrocardiology*. Oxford: Pergamon Press, 1989: 631-70.
- [14] Okin PM, Roman MJ, Devereux RB, Kligfield P. Electrocardiographic identification of increased left ventricular mass by simple voltage-duration products. *J Am Coll Cardiol* 1995; 25: 417-23.
- [15] Molloy TJ, Okin PM, Devereux RB, Kligfield P. Electrocardiographic detection of left ventricular hypertrophy by the simple QRS voltage-duration product. *J Am Coll Cardiol* 1992; 20: 1180-6.
- [16] McLaughlin SC, Aitchison TC, Macfarlane PW. Methods for improving the repeatability of automated ECG analysis. *Methods Inf Med* 1995; 34: 272-82.