Effect of ensemble averaging on amplitude and feature variabilities of Doppler spectrograms recorded in the lower limb arteries

L. Allard[1,2]   Y. E. Langlois[2]   L.-G. Durand[1]
G. O. Roederer[2]   G. Cloutier[1]

[1]Laboratoire de Génie Biomédical, Institut de Recherches Cliniques de Montréal, 110 avenue des Pins ouest, Montréal, Québec, Canada H2W 1R7
[2]Laboratoire de Recherches Non-invasives en Chirurgie Cardiovasculaire, Hôpital Hôtel-Dieu de Montréal, 3840 rue St-Urbain, Montréal, Québec, Canada H2W 1T8

Abstract—*The objective of the present study was to analyse the effect of averaging Doppler blood flow signals in lower limb arteries on amplitude and feature variabilities. Doppler signals recorded in 41 iliac and 35 superficial femoral arteries having different categories of stenosis were averaged over 1–15 cardiac cycles. Based on the relative decreasing rate of an index of variability, results indicated that amplitude variability of the spectrograms was exponentially reduced to 30, 6 and 1 per cent when averaging five, ten and 15 cardiac cycles, respectively. Nine diagnostic features were extracted from Doppler spectrograms and their variations from one cardiac cycle to the next quantified. Based on the relative decreasing rate of these variations, results indicated that feature variability was exponentially reduced to 30, 4 and 1 per cent when averaging five, ten and 15 cardiac cycles, respectively. The effect of averaging on the discriminant power of the features to separate the different categories of stenosis was also investigated by performing t-test analyses. Results showed that averaging between five and ten cardiac cycles provided the better discriminant power for most cases, whereas averaging over more than ten cardiac cycles was of little benefit. Based on the spectral analysis technique used in the present study, we conclude that averaging over ten cardiac cycles is sufficient for the analysis of Doppler spectrograms recorded in the lower limbs.*

Keywords—*Beat-to-beat amplitude variability, Beat-to-beat feature variability, Doppler ultrasound, Peripheral artery disease, Spectral analysis*

1 Introduction

SPECTRAL ANALYSIS has been shown to be a reliable technique to describe properties of the time-varying Doppler signals. The technique mostly used so far to generate the time/frequency representation (spectrogram) of Doppler signals is based on the computation of the short-time Fourier transform (PORTNOFF, 1980). These spectrograms are characterised by large random amplitude variations which are often called Doppler speckle (MO and COBBOLD, 1986). These amplitude fluctuations greatly affect the reliability of Doppler spectral diagnostic features used to assess the degree of arterial disease.

Ensemble averaging of spectrograms from different cardiac cycles has been widely used to reduce amplitude variability (GREENE et al., 1982; KNOX et al., 1982; CANNON et al., 1982; SHERRIFF et al., 1982; LANGLOIS et al., 1984; PHILLIPS et al., 1983; POOTS et al., 1986; HUTCHISON

and KARPINSKI, 1988; CLOUTIER et al., 1990; ALLARD et al., 1991). Recently, CLOUTIER et al. (1992) studied the effect of averaging on amplitude variability of cardiac Doppler spectrograms recorded in the left ventricular outflow tract. However, no study showed the effect of averaging Doppler spectrograms recorded in lower limb arteries on amplitude and feature variabilities as a function of the number of cardiac cycles. The effect of averaging on the discriminant power of diagnostic features has also never been investigated. These goals represent the main objective of the present study.

2 Materials and methods

2.1 Arterial segments analysed and data acquisition

Forty-one iliac arteries and 35 superficial femoral arteries were evaluated by ultrasonic duplex scanning and arteriography in 33 and 30 patients, respectively. These segments represent a subset of our patient database gathered during the initial development of a pattern-

recognition system used to classify disease in the lower limb arteries (ALLARD et al., 1991). Except for the young healthy volunteers, conventional biplane contrast angiographic studies were used as a reference to evaluate the degree of arterial stenosis in all segments analysed. Each angiographic film was read by an experienced angioradiologist and the degree of stenosis was assessed by caliper measurements on the arteriograms. The segments investigated belonged to one of the following categories of disease: 0–19 per cent, 20–49 per cent and 50–99 per cent diameter reduction. A fourth category composed of segments from young presumed normal subjects (PN) was also included in the study to characterise normal Doppler signals in healthy subjects. To minimise the influence of disease in proximal and/or distal segments, limbs carrying more than one stenosis, one of which being either greater than 50 per cent or greater than the one analysed, were not included in the study.

Pulse-wave (PW) Doppler examinations were performed with an Ultramark 8 Duplex scanner (Advanced Technology Laboratories) modified to allow the recording of the two quadrature Doppler signals on a four-channel audio tape recorder (TASCAM 22-4) for offline analysis. These signals were captured at the inputs of the sample-and-hold circuit of the Duplex scanner, before the spectrum analyser. The ECG signal was also recorded on an FM-carrier signal and used to synchronise the analysis of each cardiac cycle. Voice comments indicating the patient's name and arterial segments examined were recorded on the fourth channel.

A mechanically oscillating probe, operating at 5 MHz, was used for all Doppler recordings. The sample volume was either placed in the centrestream of the artery or where blood flow velocity was maximum. According to the manufacturer's specifications, the sample volume had a length of 1·5 mm in the direction of the axial beam. As the sample volume was much smaller than the diameter of the arterial segments studied, localised flow disturbance could be more easily detected and arterial wall movement rejected. For each segment examined, Doppler blood flow signals were recorded for a period of approximately 20 s to obtain at least 15 cardiac cycles. Segments were initially scanned to assess the presence or absence of stenosis. Representative blood flow signals were recorded at mid-course of the segment when no stenosis was found or at the site of most severe stenosis. The latter was identified by the technologist at the site of maximum flow disturbances, spectral broadening or increased Doppler frequencies. To standardise recordings, patients were asked to rest supine at least 30 min in a temperature-controlled environment (21–23°C). The high-pass filter used to remove possible artefacts from arterial wall motion was set at 100 Hz. For all recordings, the angle between the Doppler beam and blood vessel axis was maintained as close as possible to 60°.

2.2 Signal analysis

Doppler signal processing was performed on a 16 MHz IBM-PC/386 compatible personal computer. During tape playback, ECG and Doppler signals were digitised for 20 s with a 12-bit analogue-to-digital (A/D) convertor at sampling rates of 2 kHz and 20 kHz, respectively. Before digitisation, Doppler signals were low-pass filtered at 9 kHz with eighth-order Butterworth filters (−48 dB per octave) to prevent frequency aliasing. The cutoff frequency of 9 kHz was always greater than the $PRF/2$ used by the Doppler system, where PRF is the pulse repetition frequency. During analogue-to-digital conversion, the use of synchronised sample-and-hold circuits on each channel assured that the phase relationship between the quadrature Doppler signals was maintained. An algorithm for QRS detection was also used to synchronise the analysis of each cardiac cycle. The mean heart rate was computed and used to reject all beats having an interval differing by more than 10 per cent from the mean heart rate duration.

A Hanning window of 10 ms was applied to the Doppler signals and a fast Fourier transform (FFT) algorithm used to compute a 256-sample power spectrum. By sliding the Hanning window over a period of 700 ms after the R-wave of the ECG with an increment of 5 ms, a Doppler spectrogram composed of 140 power spectra was produced. Fifteen mean power spectrograms were finally obtained by averaging, sequentially and in synchronisation with the R-wave of the ECG, 1–15 power spectrograms from the 20 s recording. A total of 1140 (15 cycles × 76 arterial segments) mean power spectrograms were thus generated and analysed. A typical mean power spectrogram of a presumed normal iliac artery averaged over five cardiac cycles is presented in Fig. 1.

2.3 Evaluation of the amplitude variability of Doppler spectrograms

The method used to evaluate the effect of spectrogram averaging on amplitude variability of Doppler spectrograms is described in detail in another paper (CLOUTIER et al., 1992). In summary, the method is based on the computation of an index of variability which represents the area under the absolute value of the derivative function of each spectrum of the Doppler spectrogram. These absolute areas were evaluated over frequencies ranging between 500 Hz and $PRF/2$ and cumulated over the 700 ms time interval. Two values of the index of variability were
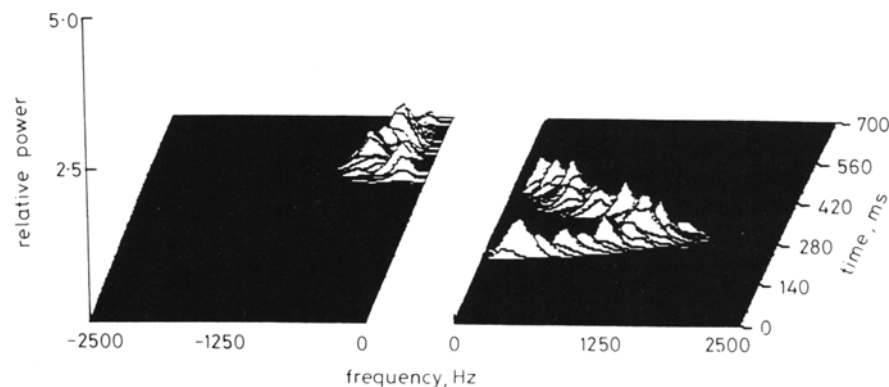

Fig. 1 Typical mean power spectrogram (power spectrum as a function of time and frequency) of a normal iliac arterial segment averaged over five cardiac cycles. Negative frequencies are associated with reverse blood flow and positive frequencies are associated with forward blood flow

thereby computed to quantify amplitude variability of positive (forward blood flow) and negative (reverse blood flow) frequencies of the Doppler spectrogram. For each arterial segment, the indices were computed over 1–15 cardiac cycles and then normalised by multiplying their values by 100 and dividing them by the maximum value found among the 15 values. The normalised indices of variability characterising the positive and the negative components of the spectrograms were averaged over 76 arterial segments.

2.4 Evaluation of the feature variability of Doppler spectrograms

Before extraction of the features, the spectral envelope delimited by the maximum and minimum frequency contours of each Doppler spectrogram were estimated using the following image-processing technique. Each Doppler spectrogram, initially in a floating-point format, was temporarily converted to a greyscale 8-bit image by mapping its amplitude range onto 256 grey levels. This operation performed some thresholding because low amplitudes of the spectrogram, essentially those associated with the background noise, were converted to a zero value. A binary image was then created from the 8-bit image and a median filter (3 × 3 window) was applied to it. The median filter had the property of preserving edges and suppressing residual noise consisting of any local fluctuations from the surrounding values of the pixels. The resulting binary image was used to define the spectral envelope and to mask the original Doppler spectrogram for feature extraction. This means that all samples of the original Doppler spectrogram located outside the spectral envelope were set to zero while those located within the spectral envelope stayed unchanged.

The features investigated in the present study are those having potential discriminant power, as already shown in a previous paper (ALLARD et al., 1991). As described in Table 1, a total of nine diagnostic features were extracted from the Doppler spectrograms: MXFS, MEFS and MNFS are, respectively, the maximum, mean and minimum frequencies at peak systole, MXFD is the maximum negative frequency during diastole, SEAP and SEAN are, respectively, the spectral envelope area of the positive and negative frequencies (CANNON et al., 1982), AVGMXF is the time-averaged frequency of the maximum positive frequency contour, SBI is the spectral broadening index at peak systole (KASSAM et al., 1982) and CV is the coefficient of variation of the frequency distribution at peak systole (KALMAN et al., 1985).

The effect of averaging Doppler spectrograms on feature variability was studied by analysing the variations of the value of each feature extracted from consecutive number of cardiac cycles. For each spectrogram, the variations of a given feature with averaging was quantified by computing the absolute difference between two values of the feature

obtained after averaging $r - 1$ and r cardiac cycles. This coefficient, denoted $AD(r)$, was computed as follows:

$$AD(r) = f(r) - f(r - 1) \qquad 2 \leqslant r \leqslant 15 \qquad (1)$$

where $f(r - 1)$ and $f(r)$ represent values of the feature after averaging $r - 1$ and r cardiac cycles, respectively. The resulting AD curve (AD values against the number of cardiac cycles) of each feature were then normalised to 100 per cent using:

$$NAD(r) = 100 \times AD(r)/\text{max}(AD) \qquad (2)$$

where $NAD(r)$ represents the normalised value of $AD(r)$ and $\text{max}(AD)$ the maximum AD value. A low value of the NAD coefficient means that the estimate of the feature is stable from one cardiac cycle to the next, whereas a high value means that it varies greatly from one cardiac cycle to the next. Because it is based on the absolute difference between two successive values of a feature, only 14 NAD values were obtained from the 15 cardiac cycles analysed. To study the global effect of averaging on feature variability, NAD coefficients were finally averaged over all segments and also over all features investigated.

2.5 Relative amplitude and feature variabilities

As proposed by CLOUTIER et al. (1992), the mean normalised indices of amplitude variability could be fitted to an exponentially decreasing model. The model characterised by an exponentially decreasing rate β_a was the following:

$$M_a(r) = A + Be^{(-\beta_a(r-1))} \qquad 1 \leqslant r \leqslant 15 \qquad (3)$$

where A and B are two parameters of the model and r is the number of cardiac cycles averaged. In the present study, this model was used to determine the exponentially decreasing rate β_a of the amplitude variability of both positive and negative components of the spectrograms. A similar model was also used to fit the NAD curve averaged over all features and segments:

$$M_f(r) = A + Be^{(-\beta_f(r-2))} \qquad 2 \leqslant r \leqslant 15 \qquad (4)$$

In a second step, β_a and β_f values were used to define indices sensitive only to the relative amplitude variability RAV and to the relative feature variability RFV. RAV and RFV indices were computed from the models of eqns. 3 and 4 where A was set to 0 per cent and B to 100 per cent:

$$RAV(r) = 100e^{(-\beta_a(r-1))} \qquad 1 \leqslant r \leqslant 15 \qquad (5)$$

$$RFV(r) = 100e^{(-\beta_f(r-2))} \qquad 2 \leqslant r \leqslant 15 \qquad (6)$$

These indices serve to quantify the relative amplitude and feature variabilities independently of their absolute range. For instance, a RAV value of 40 per cent means that amplitude variability has been reduced by 60 per cent and could be further reduced with averaging, while a value of 0 per cent means that no further reduction can be obtained with averaging.

Table 1   Description of the diagnostic features

Features	Description
MXFD	maximum negative frequency at diastole, kHz
SEAN	spectral envelope area of negative frequencies, kHz s
AVGMXF	time-averaged frequency of maximum positive frequencies, kHz
CV	coefficient of variation of the frequency distribution at peak systole
MXFS	maximum positive frequency at peak systole, kHz
MEFS	mean positive frequency at peak systole, kHz
MNFS	minimum positive frequency at peak systole, kHz
SBI	spectral broadening index at peak systole
SEAP	spectral envelope area of positive frequencies, kHz s

2.6 Evaluation of the discriminant power of the features

The ultimate objective of feature extraction from Doppler spectrograms is to grade arterial stenoses in different categories. To have a high discriminant power, a given spectral feature must have mean values significantly different between two categories of stenoses. T-test analyses were performed to verify if averaging Doppler spectrograms increases the discriminant power of the features. These analyses were achieved for both groups of arterial segments by performing for each feature a t-test between the following categories of stenoses:

(a) 0–19 per cent against 20–49 per cent
(b) 0–19 per cent against 50–99 per cent
(c) 20–49 per cent against 50–99 per cent.

A typical result was composed of six curves (p-values against the number of cardiac cycles) for each feature.

3 Results

Table 2 shows the disease distribution in the iliac and superficial femoral arteries of the subjects studied. For the 41 iliac arteries, 11 came from young presumed normal volunteers, 11 had a 0–19 per cent diameter reduction, nine had a 20–49 per cent diameter reduction and ten had

Table 2   Disease distribution of the arterial segments

Segment	N	Percentage stenosis	Mean age
Iliac artery	11	0 (PN)	26 ± 2
	11	0–19	50 ± 12
	9	20–49	62 ± 6
	10	50–99	61 ± 10
Superficial femoral artery	11	0 (PN)	26 ± 2
	9	0–19	50 ± 12
	7	20–49	62 ± 4
	8	50–99	67 ± 2

a 50–99 per cent diameter reduction. For the 35 superficial femoral arteries, 11 came from young presumed normal volunteers, nine had a 0–19 per cent diameter reduction, seven had a 20–49 per cent diameter reduction and eight had a 50–99 per cent diameter reduction.
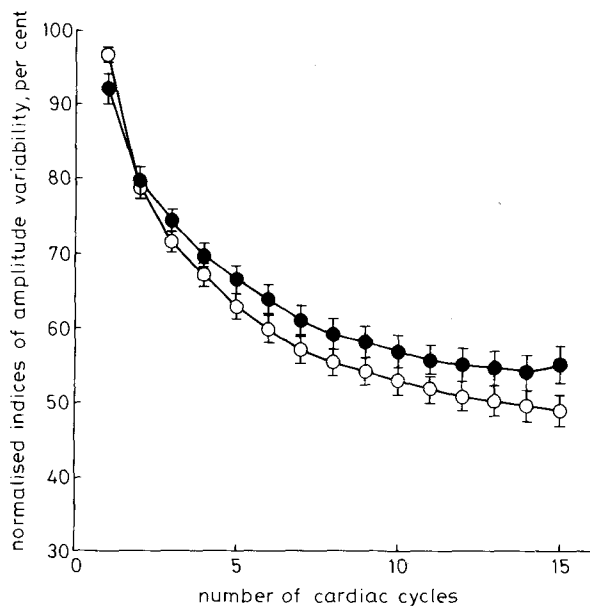

Fig. 2   Mean values of the normalised indices of amplitude variability (in per cent) of the positive (●) and negative (○) components of mean Doppler spectrograms as a function of the number of cardiac cycles. Standard errors of the mean are also presented

3.1 Amplitude variability of Doppler spectrograms

The mean values of the normalised indices of amplitude variability of the positive and negative frequencies of the spectrograms computed over the 76 arterial segments are presented in Fig. 2 as a function of the number of cardiac cycles. As seen on this figure, both curves decrease exponentially as the number of cardiac cycles increases. A plateau is reached around 13 cardiac cycles for the negative frequencies and above 15 cardiac cycles for the positive frequencies. Based on the model of eqn. 3, the exponentially decreasing rate β_a of the two curves was 0·32 for the positive frequencies of the spectrograms and 0·28 for the negative frequencies. In both cases, cross-correlation coefficients between the experimental data and the curve-fitting model were 99·0 per cent for the positive frequencies and 99·4 per cent for the negative frequencies. The RAV indices (eqn. 5) for five, ten and 15 cardiac cycles were, respectively, 27, 5 and 1 per cent for the positive component of the spectrograms and 33, 8 and 2 per cent for the negative component. These latter results are summarised in Table 3.

Table 3   Relative values of the amplitude variability (RAV) and feature variability (RFV) when averaging five, ten and 15 cardiac cycles

Number of cardiac cycles	RAV, per cent		RFV, per cent
	Positive ($\beta_a = 0\cdot32$)	Negative ($\beta_a = 0\cdot28$)	($\beta_f = 0\cdot40$)
5	27	33	30
10	5	8	4
15	1	2	1

3.2 Feature variability of Doppler spectrograms

Results presented in Fig. 3 show the NAD curve of each feature averaged over all segments. All NAD curves decrease exponentially according to the model of eqn. 4, as the number of cardiac cycles increases and a plateau is reached for many features around ten cardiac cycles. The low value of this plateau means that, in practice, no further significant reduction in feature variability could be obtained with averaging. This trend is illustrated in Fig. 4 where the nine curves of Fig. 3 have been averaged. As presented in Table 3, the exponentially decreasing rate β_f of that curve was 0·40 and the relative mean variability of all features, as expressed by the RFV index (eqn. 6), for five, ten and 15 cardiac cycles was, respectively, 30, 4 and 1 per cent. The cross-correlation coefficient between the results of Fig. 4 and the curve-fitting model was 91·1 per cent.

3.3 Discriminant power of the features

Results presented in Fig. 5 show the diagnostic power of the most discriminant features for the two groups of segments and for the three categories of stenoses as a function of the number of cardiac cycles. Only those features having a p-value $< 0\cdot05$ after averaging 15 cardiac cycles were presented (27 cases among 54). Indeed, features with a p-value $> 0\cdot05$ were not considered as discriminant and their behaviour with averaging was thus not analysed. As presented in Table 4, the behaviour of each of the 27 p-value curves were characterised into one of the following types of behaviour:

Behaviour A: an exponential decrease of p-value curve

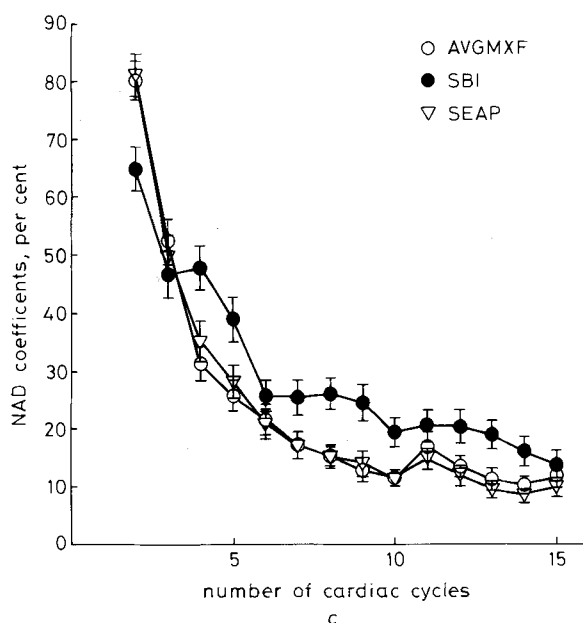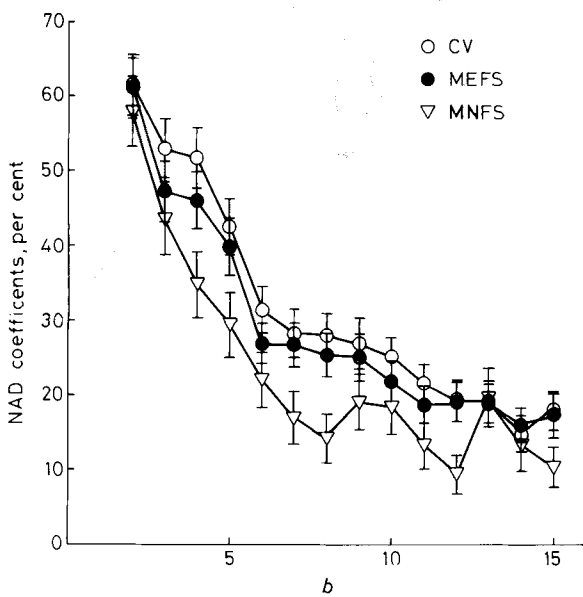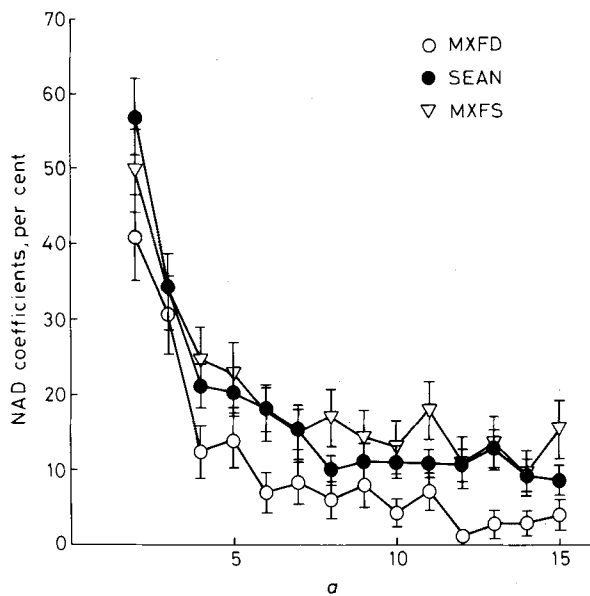**Fig. 4** *Mean values of the NAD coefficients of the nine features investigated as a function of the number of cardiac cycles. Standard errors of the mean are also presented*
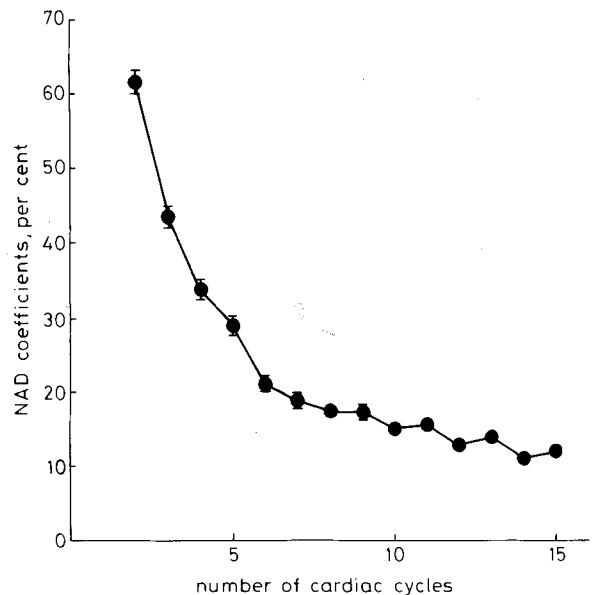
which reflects the increases in the discriminant power with averaging. A plateau is reached after approximately ten cardiac cycles.

Behaviour B: a decrease in p-values for a small number of cardiac cycles followed by a small increase and then by a plateau reached after approximately ten cardiac cycles.

Behaviour C: p-values relatively constant with averaging.

Behaviour D: an erratic p-value curve.

We observed that six (22 per cent) curves displayed behaviour A, two (7 per cent) behaviour B, 13 (48 per cent) behaviour C and six (22 per cent) behaviour D.

4 Discussion

4.1 Amplitude variability of Doppler spectrograms

It was found in the present study that averaging reduces amplitude variability of Doppler signals recorded in the lower limbs in an exponential way. A similar observation was made by CLOUTIER *et al.* (1992) when averaging cardiac Doppler signals recorded in the left ventricular outflow tract. Compared with the results of CLOUTIER *et al.* (1992), the reduction of amplitude variability observed in the present study was somewhat more rapid. One reason that could explain these differences may be attributed to the characteristics of cardiac Doppler signals which are quite different from those found in the lower limb arteries. Another explanation may be the fact that different portions of the cardiac cycle were analysed in both studies. CLOUTIER *et al.* (1992) characterised the forward blood flow only during systole ($\beta = 0.16$) and the reverse blood flow only during diastole ($\beta = 0.25$) while we characterised forward ($\beta_a = 0.32$) and reverse ($\beta_a = 0.28$) blood flow components of the spectrograms over the complete cardiac cycle (700 ms).

4.2 Feature variability of Doppler spectrograms

The method used to estimate frequency contours is an important aspect of the present study because the reliability of the feature estimates may depend on it. However, we think that its influence may be negligible if a robust and stable frequency contour algorithm is applied.
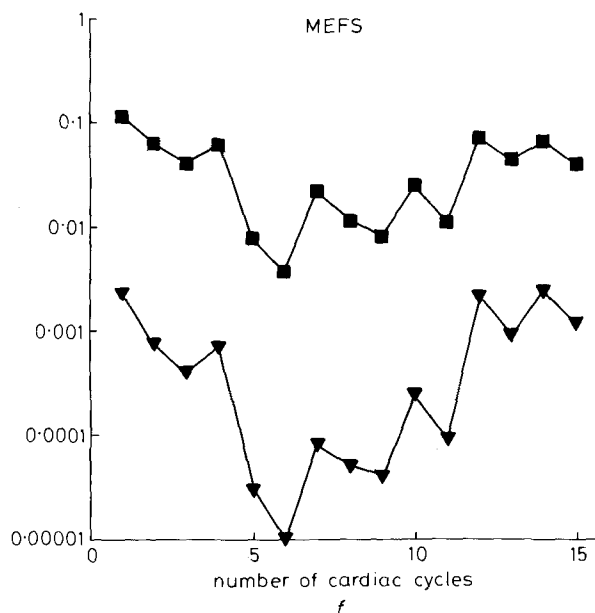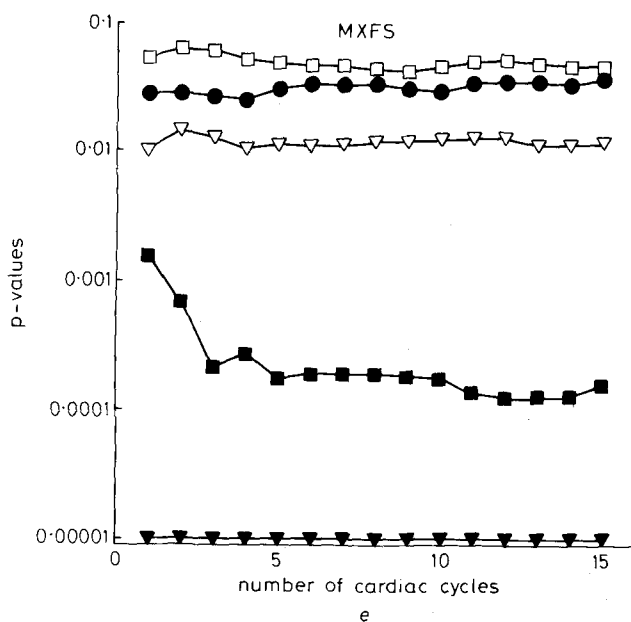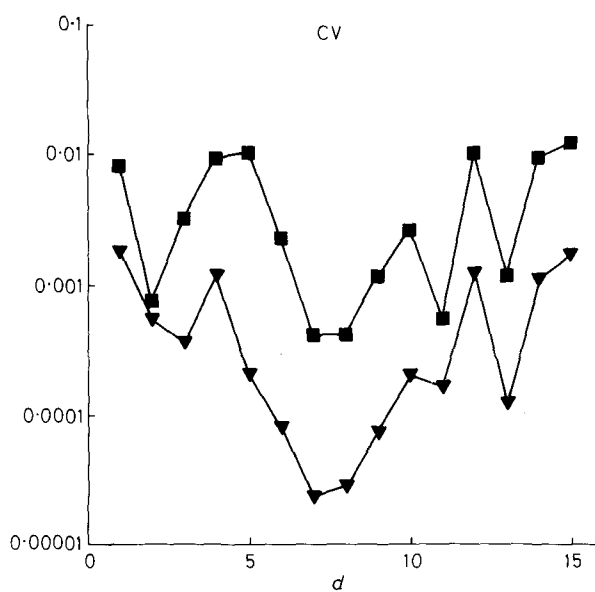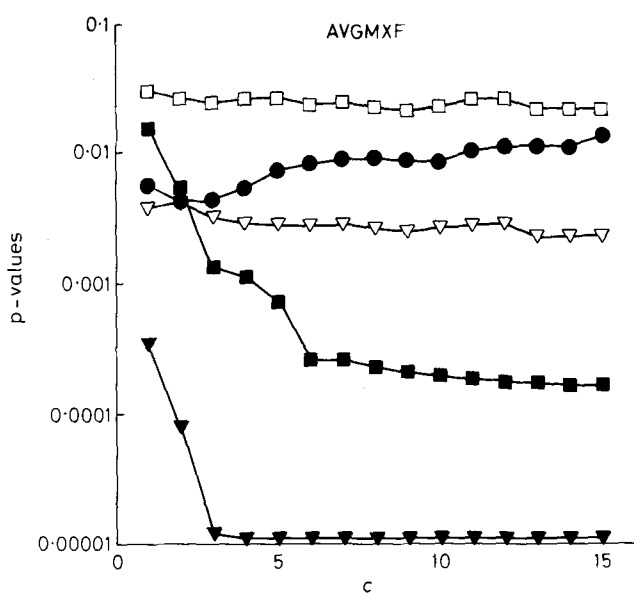

Fig. 3 *NAD coefficients of the nine features investigated as a function of the number of cardiac cycles. Standard errors of the mean are also presented*
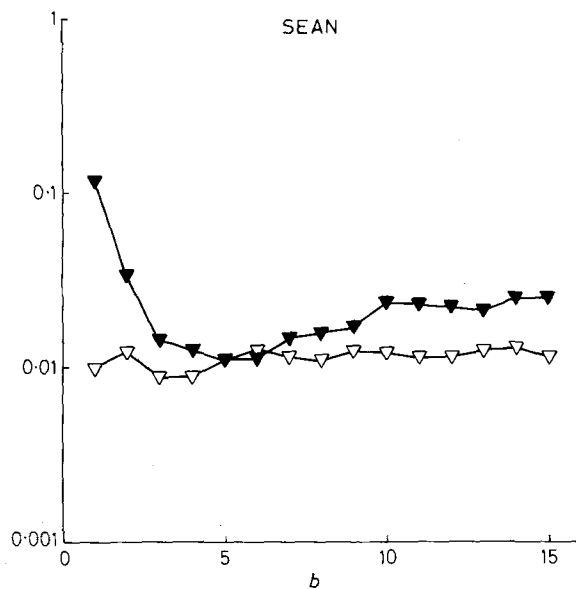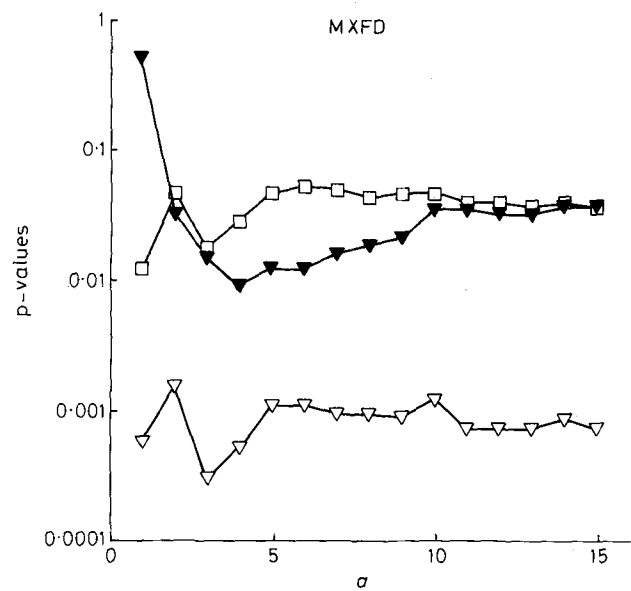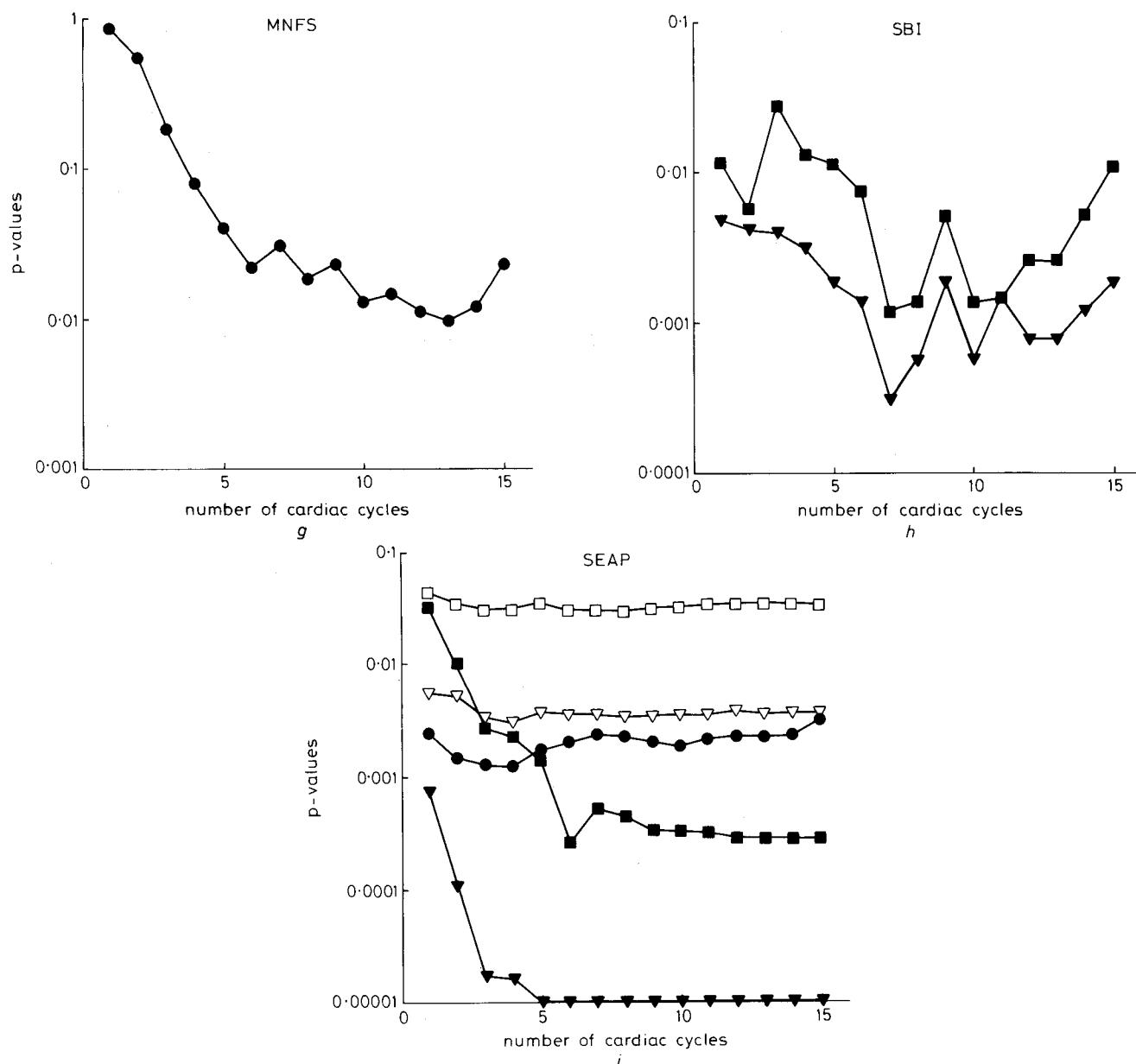
Fig. 5   *Caption opposite*

Fig. 5 *Plots of p-values of the most discriminant features as a function of the number of cardiac cycles for the two groups of segments and the three categories of disease: (a) MFXD; (b) SEAN; (c) AVGMXF; (d) CV; (e) MXFS; (f) MEFS; (g) MNFS; (h) SBI; and (i) SEAP*

iliac artery:
▽ *0–19 per cent against 50–99 per cent*
□ *20–49 per cent against 50–99 per cent*

superficial femoral artery:
● *0–19 per cent against 20–49 per cent*
▼ *0–19 per cent against 50–99 per cent*
■ *20–49 per cent against 50–99 per cent*

Table 4   Behaviours of the discriminant power of the diagnostic features between three categories of stenoses with averaging according to the p-value curves. Behaviours A, B, C and D are described in the text while NS corresponds to the features that were not significantly discriminant at a 0·05 level significance

Segment	MXFD	SEAN	AVGMXF	CV	MXFS	MEFS	MNFS	SBI	SEAP
Iliac artery									
0–19 per cent against 20–49 per cent	NS	NS	NS	NS	NS	NS	NS	NS	NS
0–19 per cent against 50–99 per cent	C	C	C	NS	C	NS	NS	NS	C
20–49 per cent against 50–99 per cent	C	NS	C	NS	C	NS	NS	NS	C
Superficial femoral artery									
0–19 per cent against 20–49 per cent	NS	NS	C	NS	C	NS	A	NS	C
0–19 per cent against 50–99 per cent	B	B	A	D	C	D	NS	D	A
20–49 per cent against 50–99 per cent	NS	NS	A	D	A	D	NS	D	A

Two general spectrum-by-spectrum thresholding approaches were used in the past to estimate frequency contours of spectrograms. The first used a threshold level set to some k multiple of the background noise level N_0 estimated from samples located in a pure noise region, e.g. near $PRF/2$ (Mo *et al.*, 1988; CLOUTIER *et al.*, 1990). With this approach, samples of each spectrum of the spectrogram were successively compared with the threshold level

to determine samples corresponding to the minimum and maximum frequencies. The major problem with this approach is the difficulty to correctly estimate N_0 and to set parameter k, which is function of the signal-to-noise ratio. Partial solution to that problem was obtained by using a parameter k adaptive to the specific conditions of each spectrogram (ALLARD et al., 1991).

The second method to estimate frequency contours of spectrograms is to use a threshold level set to some fraction of the amplitude level of the dominant frequency peak of the spectrogram. The difficulty with this second approach is that the threshold level is dependent on the maximum amplitude, which may vary from one spectrogram to another.

In summary, both approaches can either overestimate or underestimate the true maximum and minimum frequencies of a spectrum. Moreover, frequency contours of the spectrogram may contain many large spikes if noise level changes significantly over a cardiac cycle. To circumvent those problems, we treated Doppler spectrogram as greyscale image and used image-processing techniques for noise reduction and edge detection. In fact, we found in the present study that a two-dimensional median filter was a reliable and easily implemented method for the detection of spectrogram frequency contours.

Some considerations need to be presented concerning the index of variability used to characterise feature variability with averaging. In a simulation flow model where a large number of repeatable Doppler signal waveforms are available and where the true value of a given feature is known, the effect of averaging on feature variability can be assessed in terms of bias and variance of the feature. In that case, both bias and variance should tend towards zero when the number of cardiac cycles averaged tends towards infinity. The difficulty with clinical Doppler signals is that the true value of a given feature extracted from a spectrogram is unknown (and probably time-varying) and only a limited number of cardiac cycles are available. These considerations have prevented the use of the bias and variance to characterise the trend of our features with averaging.

The alternative we proposed in the present study was to quantify relative variations of the feature values from one cardiac cycles to the next, as expressed by the NAD coefficients. In a model flow simulation, as the number of cardiac cycles becomes very large, NAD coefficients should converge to zero indicating that the value of the feature tends towards its true value. In our experiment using clinical Doppler signals, it was found that NAD coefficients tend towards a non-zero value plateau due to beat-to-beat variations of the haemodynamics of the patients.

Fig. 3 demonstrates that all nine features investigated have an exponential decreasing trend with averaging. Moreover it can be seen from Figs. 2 and 4 that the exponential decreasing effect of averaging on feature variability is very similar to the one observed for amplitude variability. Taking into account the spectral-analysis technique used in the present study, these results strengthen the hypothesis that averaging over more than ten cardiac

cycles has no significant improvement on the reliability of the features extracted, as indicated in Table 3.

4.3 Discriminant power of the features

Theoretically, it is expected that the discriminant power of the features will increase with averaging because more accurate features can be extracted from spectrograms having less amplitude variability. However, based on the p-value curves of Fig. 5, it can be seen that the effect of averaging on the discriminant power of the features are characterised by different trends. These results suggest that, in practice, features with less variability from one cardiac cycle to the next do not necessarily imply that the discriminant power of the feature increases.

One reason to explain this observation could be the difficulty, during some examinations, to record Doppler signals that are reliable and representative of the stenosis as evaluated by angiography. More specifically, the results of Fig. 5 show that all p-value curves of the features extracted from the iliac artery are relatively constant with averaging whereas those from the superficial artery are characterised by different trends which depend on the feature itself and the categories of the stenosis differentiated. It is also interesting to observe that the p-value curves associated with the three features (CV, $MEFS$ and SBI) characterising the frequency distribution at peak systole are erratic. This could be explained by the fact that the mean value and standard deviation of these features extracted from only one spectrum at peak systole change significantly from one cardiac cycle to the next. A solution to minimise these statistical variations should be the averaging over more than one spectral line at peak systole.

We conclude that, for most features, averaging between five and ten cardiac cycles provides the best discriminant power, while averaging over more than ten cardiac cycles is of little benefit. Averaging over ten cardiac cycles can thus be considered an optimal choice for the analysis of the Doppler spectrograms recorded in the lower limbs, as described in the present study.

As a reference, Table 5 shows the mean values and standard deviations computed over ten cardiac cycles of the nine features extracted from spectrograms of the iliac and the superficial femoral arteries. The analysis of this table shows that features extracted from the normal segments are slightly different than those extracted from segments with a 0–19 per cent lesion. Moreover, the latter are characterised by a much larger standard deviation. In fact, based on the coefficient of variation (standard deviation/ mean value), variability of the features around their mean value increases as a function of the degree of stenosis. These differences were observed for any given number of cardiac cycles.

The p-values corresponding to the data of Table 5 are presented in Table 6. The underlined values are those corresponding to the features being significantly different ($p < 0.05$) between two categories of disease. It can be seen that some diagnostic features are more discriminant than

Table 5   Mean value ± standard deviation of diagnostic features when averaging ten cardiac cycles

Segment	MXFD, kHz	SEAN, kHz s	AVGMXF, kHz	CV	MXFS, kHz	MEFS, kHz	MNFS, kHz	SBI	SEAP, kHz s
Iliac artery									
PN	−639 ± 191	9·01 ± 5·06	907 ± 217	0·174 ± 0·047	2628 ± 604	1779 ± 508	668 ± 492	32·8 ± 3·9	83·4 ± 23·5
0–19 per cent	−760 ± 408	12·32 ± 12·22	1061 ± 452	0·240 ± 0·101	3203 ± 1411	2147 ± 1035	604 ± 590	34·1 ± 9·8	101·6 ± 54·6
20–49 per cent	−530 ± 398	6·90 ± 9·74	1383 ± 519	0·326 ± 0·138	3767 ± 651	2239 ± 327	347 ± 334	39·9 ± 8·0	151·5 ± 57·3
50–99 per cent	−234 ± 165	1·31 ± 2·76	2747 ± 1553	0·422 ± 0·292	5789 ± 2710	3317 ± 2071	617 ± 1216	42·2 ± 18·7	325·9 ± 216·0
Superficial femoral artery									
PN	−1080 ± 419	22·26 ± 13·10	1157 ± 115	0·118 ± 0·039	3509 ± 388	2766 ± 389	1385 ± 560	21·4 ± 3·7	86·4 ± 24·1
0–19 per cent	−833 ± 470	13·19 ± 9·64	857 ± 214	0·188 ± 0·101	2726 ± 546	1926 ± 519	720 ± 489	30·2 ± 8·6	72·2 ± 26·7
20–49 per cent	−525 ± 349	6·89 ± 8·04	1364 ± 437	0·268 ± 0·046	3973 ± 1414	2591 ± 876	190 ± 42	34·2 ± 2·6	159·8 ± 62·2
50–99 per cent	−371 ± 320	3·38 ± 5·53	3319 ± 921	0·519 ± 0·174	7822 ± 1443	3952 ± 1151	664 ± 548	49·5 ± 9·6	415·0 ± 126·4

Table 6 *Distribution of p-values of nine diagnostic features between three categories of stenoses when averaging ten cardiac cycles*

Segment	MFXD	SEAN	AVGMXF	CV	MXFS	MEFS	MNFS	SBI	SEAP
Iliac artery									
0–19 per cent against 20–49 per cent	0·220	0·296	0·156	0·125	0·285	0·800	0·262	0·169	0·062
0–19 per cent against 50–99 per cent	0·001	0·012	0·003	0·066	0·012	0·113	0·974	0·225	0·003
20–49 per cent against 50–99 per cent	0·046	0·099	0·023	0·381	0·044	0·142	0·529	0·743	0·032
Superficial femoral artery									
0–19 per cent against 20–49 per cent	0·169	0·187	0·008	0·074	0·028	0·078	0·013	0·263	0·002
0–19 per cent against 50–99 per cent	0·034	0·023	<0·0001	0·0002	< 0·0001	0·0002	0·825	0·0006	<0·0001
20–49 per cent against 50–99 per cent	0·390	0·337	0·0002	0·003	0·0002	0·024	0·041	0·001	0·0003

others and also that disease discrimination seems easier in the superficial femoral artery than in the iliac artery. Moreover, some categories of disease are differentiated more easily. Indeed, it appears that discrimination between the 20–49 per cent against 50–99 per cent or the 0–19 per cent against 50–99 per cent categories is easier than the one between the 0–19 per cent against 20–49 per cent. This conclusion can also be drawn from Table 5 by observing that, between the 0–19 per cent and 20–49 per cent categories, the mean values of most features are not significantly different, especially for the iliac artery. This explains the difficulties encountered in a previous study in which these same features and a pattern-recognition system (ALLARD *et al.*, 1991) were used to separate the 20–49 per cent category from the two other categories.

5 Conclusion

In clinical practice, the reduction of amplitude and feature variabilities of Doppler spectrograms is very important. Indeed, we believe that smoother spectrograms can facilitate the extraction of discriminant features and thus significantly improve the diagnostic accuracy. In the present study, it was demonstrated that the ensemble averaging technique reduces amplitude and feature variabilities according to an exponentially decreasing curve.

It was also demonstrated that averaging between five and ten cardiac cycles provides the best discriminant power for the most diagnostic features. Because the averaging of only a limited number of Doppler spectrograms are required to reduce amplitude and feature variabilities and probably improve diagnostic accuracy, clinical practice could certainly take advantage of the benefits from ensemble averaging technique.

References

ALLARD, L., LANGLOIS, Y. E., DURAND, L.-G., ROEDERER, G. O., BEAUDOIN, M., CLOUTIER, G., ROY, P. and ROBILLARD, P. (1991) Computer analysis and pattern recognition of Doppler blood spectra for disease classification in the lower limb arteries. *Ultrasound in Med. & Biol.*, **17**, 211–223.

CANNON, S. R., RICHARDS, K. L. and ROLLWITZ, W. T. (1982) Digital Fourier techniques in the diagnosis and quantification of aortic stenosis with pulsed-Doppler echocardiography. *J. Clin. Ultrasound*, **10**, 101–107.

CLOUTIER, G., LEMIRE, F., DURAND, L.-G., LATOUR, Y. and LANGLOIS, Y. E. (1990) Computer evaluation of Doppler spectral envelope area in patients having a valvular aortic stenosis. *Ultrasound in Med. & Biol.*, **16**, 247–260.

CLOUTIER, G., ALLARD, L., GUO, Z. and DURAND, L. G. (1992) Effect of averaging Doppler spectrograms on the reduction of

their amplitude variability. *Med. & Biol. Eng. & Comput.*, **30**, 177–186.

GREENE, F. M. Jr, BEACH, K., STRANDNESS, D. E. Jr, FELL, G. and PHILLIPS, D. J. (1982) Computer based pattern recognition of carotid arterial disease using pulsed Doppler ultrasound. *Ultrasound in Med. & Biol.*, **8**, 161–176.

HUTCHINSON, K. J. and KARPINSKI, E. (1988) Stability of flow pattern in the *in vivo* post-stenotic velocity field. *Ibid.*, **14**, 269–275.

KALMAN, P. G., JOHNSTON, K. W., ZUECH, P., KASSAM, M. and POOTS, K. (1985) *In vitro* comparison of alternative methods for quantifying the severity of Doppler spectral broadening for the diagnosis of carotid arterial occlusion disease. *Ibid.*, **11**, 435–440.

KASSAM, M. S., COBBOLD, R. S. C., JOHNSTON, K. W. and GRAHAM, C. M. (1982) Method for estimating the Doppler mean velocity waveform. *Ibid.*, **8**, 537–544.

KNOX, R. A., GREENE, F. M., BEACH, K., PHILLIPS, D. J., CHIKOS, P. M. and STRANDNESS, D. E. Jr (1982) Computer based classification of carotid arterial disease: a prospective assessment. *Stroke*, **13**, 589–594.

LANGLOIS, Y. E., GREENE, F. M. Jr, ROEDERER, G. O., JAGER, K. A., PHILLIPS, D. J., BEACH, K. W. and STRANDNESS, D. E. Jr (1984) Computer based pattern recognition of carotid artery Doppler signals for disease classification: prospective validation. *Ultrasound in Med. & Biol.*, **10**, 581–595.

MO, L. Y. and COBBOLD, S. C. (1986) 'Speckle' in continuous wave Doppler ultrasound spectra: a simulation study. *IEEE Trans.*, **UFFC-33**, 747–753.

MO, L. Y., YUN, L. C. M. and COBBOLD, R. S. C. (1988) Comparison of four digital maximum frequency estimators for Doppler ultrasound. *Ultrasound in Med. & Biol.*, **14**, 355–363.

PHILLIPS, D. J., GREENE, F. M. Jr, LANGLOIS, Y. E., ROEDERER, G. O. and STRANDNESS, D. E. Jr (1983) Flow velocity patterns in the carotid bifurcations of young, presumed normal subjects. *Ibid.*, **9**, 39–49.

POOTS, J. K., JOHNSTON, K. W., COBBOLD, R. S. C. and KASSAM, M. (1986) Comparison of CW Doppler ultrasound spectra with the spectra derived from a flow visualization model. *Ibid.*, **12**, 125–133.

PORTNOFF, M. R. (1980) Time-frequency representation of digital signals and systems based on short-time Fourier analysis. *IEEE Trans.*, **ASSP-28**, 55–68.

SHERRIFF, S. B., BARBER, D. C., MARTIN, T. R. P. and LAKEMAN, J. M. (1982) Use of principal component factor analysis in the detection of carotid artery disease from Dopler ultrasound. *Med. & Biol. Eng. & Comput.*, **20**, 351–356.

Authors' biographies
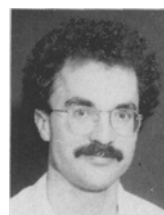
Louis Allard was born in Québec City, Québec, Canada, in 1961. He received the B.Sc.A degree in Engineering Physics from Laval University, Québec, in 1985 and the M.Sc.A degree in Biomedical Engineering from the Ecole Polytechnique, University of Montréal in 1988. He is presently a research associate at the Biomedical Engineering Laboratory of the Clinical Research Institute of Montreal and at the Cardiovascular Research Laboratory of the Hôtel-Dieu de Montréal Hospital. His major interests are digital signal processing, image processing and pattern recognition applied to echo-Doppler signals.
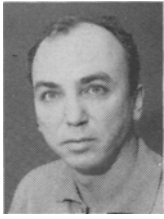
Yves Langlois was born in Montréal, Canada, in 1950. He is Associate Professor of Surgery at the University of Montreal and an active cardiovascular surgeon at the Hôtel Dieu de Montréal Hospital. He is also Director of the Cardiovascular Research Laboratory at Hôtel Dieu and Senior Researcher at the Biomedical Engineering Laboratory of the Clinical Research Institute of Montreal. His major interests are cardiovascular physiology and the use of complex biological signals in their clinical application for the detection and quantification of arterial and cardiac diseases.
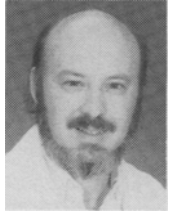
Louis-Gilles Durand was born in St-Jean de Matha, Québec, Canada, in 1949. He is Director of the Biomedical Engineering Laboratory, Clinical Research Institute of Montreal, Research Assistant Professor in the Department of Medicine, University of Montreal and Visiting Professor at the Institute of Biomedical Engineering, Ecole Polytechnique, Montreal, Canada. He has B.Sc, M.Sc. and Ph.D. degrees in Electrical Engineering from the Ecole Polytechnique, University of Montreal. In 1975 he set up a biomedical engineering service at the Clinical Research Institute of Montreal. His major interests are digital signal processing of physiological signals, computer modelling of physiological systems and medical instrumentation.

Born in Montreal, Dr Ghislaine Roederer first obtained a B.Sc. in Physiology from McGill University, Canada, graduated from the University of Montreal Medical School and has recently submitted a Ph.D. theses in Biomedical Sciences. Her research training has focused on the use of noninvasive techniques in vascular disease evaluation (Seattle, Washington and New York City) and lipoprotein disorders (Montreal, Quebec). Her current interests lie in the pursuit of the characterisation of atherosclerosis and the effect of medical and/or surgical interventions on its evolution.

Guy Cloutier was born in Trois-Rivières, Québec, Canada, in 1961. He received the B.Eng. degree in Electrical Engineering from the Université du Québec à Trois-Rivières in 1984, the M.Sc. and Ph.D. degrees in Biomedical Engineering from the Ecole Polytechnique, Université de Montréal in 1986 and 1990, respectively. Currently, Mr Cloutier is a research fellow of the Natural Sciences & Engineering Research Council of Canada and pursues postdoctoral research at the Pennsylvania State University, USA. His principal research interests are blood flow characterisation, ultrasonic tissue characterisation, and digital signal processing and pattern recognition of Doppler echocardiographic signals.