Assessment of inflow and outflow stenoses using big spectral data and radial-based colour relation analysis on in vitro arteriovenous graft biophysical experimental model

Wei-Ling Chen1, Chung-Dann Kan2, Chia-Hung Lin3

1Department of Engineering and Maintenance, Kaohsiung Veterans General Hospital, Kaohsiung City 81362, Taiwan
2Department of Surgery, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan City 70101, Taiwan
3Department of Electrical Engineering, Kao-Yuan University, Kaohsiung City 82151, Taiwan
E-mail: kcd56@mail.ncku.edu.tw

Abstract: Dialysis vascular accesses are critical for dialysis therapy, but they frequently suffer from stenotic complications. Higher patency rates and thrombosis rates are a concern to nephrology nurses and patients. These complications are complex events, including inflow stenosis, outflow stenosis, and coexistence of both. Therefore, a biophysical experimental model is employed to mimic the various combinations of stenoses and dialysis circulation circuits on a virtual adult hand. Considering the suggested signal preprocessing specifications, auscultation method and frequency analysis technique are used to extract the key frequency components from sufficient big spectral data. Key frequency components, depending on the degree of stenosis (DOS) (positive correlation), are validated using multiple regression models with multiple explanatory variables and response variables. A new machine learning method, radial-based colour relation analysis, is employed to identify the level of DOS at the inflow and outflow sites. In contrast to the multiple linear regression and traditional machine learning method, the experimental results indicated that the proposed screening model had higher accuracy (hit rate), true-positive rate, and true-negative rate in clinical indication.

1 Introduction

Cardiovascular diseases (CVDs) and type two diabetes mellitus are major risk factors for morbidity and mortality in elderly patients and patients receiving or not receiving hemodialysis therapy. In Taiwan, CVDs are prevalent in elderly patients aged >65 years and in those with certain complications such as coronary artery disease, stroke, hypertensive heart disease, and peripheral arterial disease (PAD). Hemodialysis patients with CVDs and diabetes mellitus are at an increased risk of developing endothelial dysfunction in vascular accesses and atherosclerosis of peripheral arteries. These symptoms further lead to PAD or peripheral vascular stenosis in both lower and upper limbs. In addition, long-term traumatic puncture that is required for dialysis therapy results in vascular access stenosis, such as single stenosis and coexisting inflow and outflow stenoses. In clinical investigations, an arteriovenous graft (AVG) has a higher patency rate than an arteriovenous fistula [1, 2]. Stenosis development initiates at the outflow sites, such as venous side stenosis or stenoses near the graft-to-vein site. This can lead to arterial side stenosis and ultimately progression of coexisting stenosis, gradually leading to increased flow resistance and high-pressure pulsatile flow and to decreased vessel compliances. Acute problems result in coexistence of inflow and outflow stenoses, resulting in vibrations, turbulent flows, and murmur sounds in the stenotic vicinity [3–5]. However, obtaining only one measurement results in lack of information, because frequency sounds are dependent on the stenosis sites, multiple stenosis, monitoring sites, and degree of stenosis (DOS). Therefore, this study proposes the dual-channel auscultation method to increase the accuracy of stenosis detection in both inflow and outflow stenoses, as shown in Fig. 1.

In routine examinations, inspection, palpation, and auscultation methods are commonly recommended for monitoring and surveillance of vascular access function by the National Kidney Foundation Dialysis Outcomes Quality Initiative [6, 7]. However, the accuracy of inspection and palpation methods depends on the physician experience. Ultrasound examination with B-mode or A-mode operations is a useful and non-invasive technique with a higher accurate detection rate for stenosis diagnosis. It provides images and acoustic signals to visualise the presence of stenosis inside the vascular accesses and blood flow velocities [2, 8]. When a stenotic focus has >50% stenosis in the vascular accesses, the flow rate will be decreased and the flow velocity and resistance will be increased, resulting in hemodynamic or clinical abnormality. Through image examinations, >70% stenosis in a vascular access, percutaneous transluminal angioplasty or surgical treatment is required to dilate the stenotic lesion. However, its techniques and decision-making of patients’ diagnosis depend on the operation skill of the physician. In addition, they have limitations, including being not used to provide early warning for stenosis detection during dialysis treatment. Therefore, auscultation is a real-time and non-invasive screening method to detect the local conditions of hemodynamic motions using electronic stethoscopes for bruits murmur sounds [9–12]. The sound wave patterns can be captured in different data length of the streaming acoustic data. The phonoangiography (PAG) signals indicate the hemodynamic conditions in different DOS, such as the relationship of PAG to stenotic lesions, flow velocities, and flow rates [4, 5]. PAG signals can also be decomposed to time-frequency or frequency-based features for assessing the function of dialysis vascular accesses. Previous studies have reported that the changes in frequency and amplitude are dependent on the stenosis sites, monitoring sites, and DOS [9–12]. The characteristic frequencies indicate the spectral peaks between 25 and 800 Hz. Thus, the characteristic frequencies and power spectrum can be applied to separate the normal condition from the single and multiple stenosis with different acoustic data length and sampling frequency.

For multiple stenosis screening, multiple explanatory variables are designed to estimate the multiple response variables. Big data
analysis and collection can extract the frequency-based features from a particular size of a data set. These available data sets are indeed large and correlate to the characteristics of specific ecosystems, such as disease prevention, healthcare, and business trends [13, 14]. Multidimensional big data require decision-making, machine learning, graph analytics, and visualisation to characterise the primary components of big data [15–17]. Therefore, intelligent machine learning with multiple linear regressions is carried out to map the relationship between additional explanatory variables and response variables, such as parallel computing models and multilayer neural networks. In this study, for various combinations of single stenosis and multiple stenoses in an AVG biophysical experimental model, we propose the practical measurements at the pre-stenosis and post-stenosis sites to obtain dual-channel PAG signals. Based on frequency features, the Burg autoregressive (AR) method [18–20] was used to extract key parameters in the frequency spectra, which were established training data for intelligent machine learning. Radial-based function is a kernel model with Gaussian functions that uses the relational measurement for pattern recognition. Then, colour relation analysis (CRA) [21, 22] is employed to separate the normal condition from higher DOS. The similarity and dissimilarity levels are parameterised into the specific membership grades using the fuzzification operations. CRA utilises the membership grades into ‘hue angle’ and ‘saturation value’ to identify the DOS by describing the perceptual colour relationships for DOS < 50%, 50% < DOS < 70%, and DOS > 70% screening. In contrast to the multiple linear regression [23, 24] and machine learning methods [25, 26], the proposed model overcomes the three limitations, including the determination of multilayer network structure, the
The measurement sites are also around the in vitro biophysical experimental model. Various conditions in an in vitro biophysical experimental model, are consistent with pressure and thresholds should be $< 600 \text{ ml/min}$. As shown in Fig. 2, a roller pump (40–80 rpm) is used to drive the blood-mimicking fluid (BMF) and is employed to supply the pulsatile flow from the radial artery to the vein on the virtual adult hand. Another roller pump provides the capillary circulation to mimic the peripheral arterial system. A heart rate of 80–120 beats/min is used to drive the BMF through different stenotic grafts, as shown in Fig. 2c. 

The DOS is defined as follows [9–12]:

$$\text{DOS}\% = \left[1 - \frac{D_s}{D_a}\right]^2 \times 100\% \quad (1)$$

where $D_s$ is the diameter of the normal access in the direction of the BMF flow and $D_a$ is the diameter of the stenotic access. The 50–95% of tube diameters are considered in this study, where the inner diameters of a normal access are considered to have a circular cross section with constant wall thickness, $D_n = 0.635 \text{ cm}$, for simulating arteries and veins. BMF is prepared from a mixture of water and glycerin with a hematocrit ratio of 38–62%, kinematic viscosity of $3.2 \times 10^{-5} \text{ m}^2/\text{s}$, density of $1090 \text{ kg/m}^3$, and temperature of $28^\circ \text{C}$. The biophysical model of an AVG consists of two semi-cylindrical blocks of polydimethylsiloxane with a refractive index of 1.40 [27], as shown in Fig. 2f. The roller pump is used to simulate the blood circulation, heartbeat, blood pressure, and pulsatile flow. It is employed to control the flow rate, pressure, and heart rate during dialysis treatment with a maximum flow rate of approximately 600 ml/min through the AVG access. For stenotic conditions, the flow threshold should be $< 600 \text{ ml/min}$. As shown in Fig. 2f, via the pressure and flow rate monitors, the blood pressures, 90–150 mmHg, and flow rates, 600–400 l/min through an AVG biophysical model, are confirmed that are adapted to simulate the various conditions in an in vitro biophysical experimental model. The measurement sites are also around the inflow sites and outflow sites. Then, the auscultation method is employed to acquire PAG signals via a data acquisition card (National Instruments$^\text{TM}$ DAQ card, analogue-to-digital converter with eight channels, sampling rate of 1 MHz, USA) or wireless (infrared transmission) data transfers to a tablet PC.

2.2 PAG signal acquisition and feature extraction

Considering the various combinations of single stenosis (0–95%) and multiple stenoses (50–95% against 50–95%), we designed 21 single stenosis (inflow or outflow stenosis) and 150 coexisting inflow and outflow stenosis experiments at the pre-stenosis and post-stenosis sites. PAG signals were obtained using dual-channel electronic stethoscopes (3M$^\text{TM}$ Littmann®, Model 4100, K051790, Minnesota, USA). As shown in Fig. 3, fast Fourier transform (FFT) algorithm was used with discrete Fourier transform (DFT) to transform the time-domain PAG signal into discrete frequency-domain representation, as follows [20]:

$$x[n] = \text{PAG}[n], \quad \Omega = \frac{2\pi m}{N}, \quad n = 0, 1, \ldots, N - 1 \quad (2)$$

$$y[n'] = \sum_{n=0}^{N-1} x[n]e^{-j\Omega n'}, \quad n' = 0, 1, \ldots, N - 1 \quad (3)$$

where $x[n]$ is a discrete time PAG signal (sampling data of original PAG signal, PAG[n]); $n$ and $n'$ are the sampling points; and $y[n']$ is a periodic and extended from frequency, $f = 0$, to sampling frequency, $f_s$. The DFT was defined in the region between 0 and $f_s$. The first half of the frequency range (from 0 to the Nyquist frequency, $f_s/2$) was sufficient to identify the component frequencies in the FFT data. After the DFT process, a PAG signal can be decomposed into a number of discrete frequencies, $y[n']$, in the specific range from 0 to 1000 Hz. The frequency spectrum was normalised with the maximum amplitude, $\Phi[n']$, as

$$\Phi[n'] = \left[\frac{\text{abs}(y[n'])}{\text{max} [\text{abs}(y[n'])]}\right], \quad n' = 1, 2, 3, \ldots, p \quad (4)$$

Then, the Burg AR method [10–12] was employed to smooth the frequency spectrum over a continuous range. To obtain the reliable characteristic frequencies, the AR model with AR order was determined to minimise the sum of the residual energies using the Levinson–Durbin recursion algorithm [18, 19]. Overall procedures of PAG signal processing and feature extraction are shown in Fig. 3.

In this study, stenosis combinations are considered at the pre-stenosis site and post-stenosis site, as shown in Fig. 4a, and all the 171 combinations of different stenotic segments are used to characterise the primary components of feature frequencies. PAG signals with different window lengths, $n = 0, 1, 2, ..., N – 1$, and the choice of AR order were suggested to achieve a good evaluation analysis by a clinician, including (i) matching the time-limited observation (10-s record for a consumer electronic stethoscope), (ii) reducing spectral broadening and spectral leakage, (iii) distinguishing the characteristic features in the specific ranges, and (iv) enhancing the frequency resolution at the observed flow lengths. The characteristic frequencies could be obtained to determine the spectral peaks between 200–300 and 500–600 Hz in two marginal distributions, symbolised as frequency-1 and frequency-2, respectively, as shown in Figs. 4b and c. Therefore, distinct peaks of power spectral densities (PSDs) can be found as key features for the stenosis screening at the pre-stenosis site and post-stenosis site, symbolised as four combinations (pre-PSD-1, pre-frequency-1), (pre-PSD-2, pre-frequency-2), (post-PSD-1, post-frequency-1), and (post-PSD-2, post-frequency-2). Multidimensional data sets of PSDs versus DOSs are visualised in three-dimensional (3D) feature space, as shown in Figs. 4d and e. Through experimental (big data) collections, four key explanatory variables (pre-PSD-1, pre-PSD-2) and (post-PSD-1, post-PSD-2) are employed to assess the inflow and outflow stenoses.

2.3 Radial-based colour relation analysis

For the high-dimensional recognition application, a non-linear regression pattern recognition scheme is used to estimate the pre-DOS and post-DOS, consisting of the radial functions and CRA [21, 26], as shown in Fig. 5. This model gives a sufficient number of pattern nodes to establish a non-linearity estimator with two or more input variables, corresponding pattern nodes keep on growing with addition or deletion. It could model a high-dimensional pattern mechanism with various combinations of training data are available, including various flow rates, single stenosis (50–95%), and multiple stenoses (50–95% against 50–95%). These patterns are usually groups of measurements or observations and can be classified, defining in a multi-dimensional pattern space.
**Fig. 3** Procedure of PAG signal processing and feature extraction

**Fig. 4** Stenosis combinations and frequency domain features

- **a** Combinations of single stenosis and multiple stenosis
- **b** Pre-frequency and post-frequency against pre-DOS
- **c** Pre-frequency and post-frequency against post-DOS
- **d** Pre-PSD-1 and post-PSD-1 against DOS
- **e** Pre-PSD-2 and post-PSD-2 against DOS

---

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
Fig. 5  Structure of RCRA based estimator

\(a\) RCRA based estimator for pre-DOS and post-DOS estimation
\(b\) HSV colour space

Similarity and dissimilarity, relational measurement is a manner for pattern recognition between the reference pattern, \(\Phi_r\), and other comparative patterns, \(\Phi_c\). Gaussian functions (radial-based function), \(G\), at pattern nodes are used to classify these patterns. Assume a reference pattern, \(\Phi_r(0) = \{\phi_r(0), \phi_r(0), \ldots\}\), and \(K\) comparative patterns, \(\Phi_c(k) = \{\phi_c(k), \phi_c(k), \ldots\}\), \(k = 1, 2, 3, \ldots, K\), can be represented as

\[
G(k) = \exp \left( -\frac{1}{2\sigma^2} \sum_{j=1}^{p} (\Delta \phi_j(k))^2 \right)
\]

(5)

\[
ED(k) = \sqrt{\sum_{j=1}^{p} (\Delta \phi_j(k))^2}, \quad \Delta \phi_j(k) = |\phi_j(0) - \phi_j(k)|
\]

(6)

where \(p\) is the dimensional space in one input pattern; \(ED(k)\) is the Euclidean distance (ED); \(\sigma\) is the standard deviation. The similarity degree is parameterised with Gaussian function, varying within between value 0 and 1. If input pattern, \(\Phi(0)\), is similar to any comparative pattern \(\Phi_r(k)\), the \(ED(k)\) will be small values (\(ED(k) \rightarrow 0, G(k) \rightarrow 1\)). Then, this index is used in CRA to measure the relationship between the \(\Phi_r(0)\) and \(\Phi_c(k)\). This concept is employed to screen the ‘similarity degree, \(G(k)\)’ among the template patterns.

Three aggregate classes are used to assess stenosis detection, including ‘Class 1: DOS < 50\%’, ‘Class 2: 50\% < DOS < 70\%’, and ‘Class 3: DOS > 70\%’. Thus, overall indexes \(ED(k), k = 1, 2, 3, \ldots, K\), are converted to grey grades to compute the average grade for each class at arterial anastomosis site and venous anastomosis site, by

\[
y_m(k) = \frac{\sum_{j=1}^{K} w_{kj} \times G(k)}{\sum_{j=1}^{K} G(k)}
\]

(7)

where the values of \(w_{kj} \in [0, 1]\) are weighted connections for the three classes, \(j = 1, 2, 3\), at the pre-stenosis and post-stenosis sites, \(m = 1, 2\), associated with the output vector: \(Y = [y_{m1}, y_{m2}, y_{m3}]\). The connection weights are coded as binary values (0 /1), while \(w_{kj} = 1\) belongs to one of three classes. Then, the minimum and maximum average grades can be found, as

\[
\rho_{m_{\text{min}}} = \min [y_{m1}, y_{m2}, y_{m3}]
\]

(8)

\[
\rho_{m_{\text{max}}} = \max [y_{m1}, y_{m2}, y_{m3}]
\]

(9)

\[
\Delta \rho = \rho_{m_{\text{max}}} - \rho_{m_{\text{min}}}
\]

(10)

where \(\rho_{m_{\text{min}}} \neq \rho_{m_{\text{max}}}.\) Three average grades can be converted to primary colour grades, \(r\) (red), \(g\) (green), and \(b\) (blue), as

\[
g_m = y_{m1}, \quad b_m = y_{m2}, \quad r_m = y_{m3}; \quad m = 1, 2
\]

(11)

According to the HSV colour model [21, 22]. CRA model is defined as a transformation method from the RGB colour space and the HSV colour space, where primary colour grades, \(r_m \in [0, 1], g_m \in [0, 1], b_m \in [0, 1]\), coordinates in the RGB colour space. The hue angle, \(h_m \in [0, 360^\circ]\), can be found from the RGB colour space, and it can be defined as

\[
h_m = \begin{cases} 
60 \times \left( \frac{g_m - b_m}{\Delta \rho_m} \right), & r_m = \rho_{m_{\text{max}}}, \quad g_m \geq b_m \\
60 \times \left( \frac{r_m - b_m}{\Delta \rho_m} \right) + 120, & \rho_{m_{\text{min}}} \leq r_m \leq \rho_{m_{\text{max}}} \\
60 \times \left( \frac{r_m - g_m}{\Delta \rho_m} \right) + 240, & b_m \leq \rho_{m_{\text{max}}}, \quad r_m < b_m \\
60 \times \left( \frac{r_m - b_m}{\Delta \rho_m} \right) + 360, & \rho_{m_{\text{min}}} < r_m < \rho_{m_{\text{max}}}
\end{cases}
\]

(12)

where \(\rho_{m_{\text{min}}} < \rho_{m_{\text{max}}}, h_m = 0\). The saturation, \(s_m\), and the value, \(v_m\),
The hue angle, $h$, is generally normalised to lie between 0° and 360°, and $h_m$ has no geometric meaning when $P_{\text{min}}=P_{\text{max}}$ and saturation, $s_m$, is zero. The index, $h_m$, is used to identify the three classes, as four open intervals as follows: (1) $h_m=(0°, 180°)$ is green-series colour for $\text{DOS}<50\%$, (2) $h_m=(180°, 240°)$ is blue-series colour for $50\%<\text{DOS}<70\%$, and (3) $h_m=(0°, 60°)$ or $(240°, 360°)$ is red-series colour for $\text{DOS}>70\%$, as shown in Fig. 3b. Index $s_m$ is also employed to provide the confidence level. If its value is $>0.5$ and approaches 1.0, we have high confidence to confirm the possible class.

In addition, assessment results defined by the $(h, s, v)$ values in the HSV colour space, index $h_m$ with $s_m$ and $v_m$ varying between 0 and 1, can be calculated in the normalised colour space, as [21]

$$s_m = 1 - \frac{P_{\text{min}}}{v_m}, v_m = P_{\text{min}}, P_{\text{max}} \neq 0 \quad (13)$$

The hue angle, $h_m$, is generally normalised to lie between 0° and 360°, and $h_m$ has no geometric meaning when $P_{\text{min}}=P_{\text{max}}$ and saturation, $s_m$, is zero. The index, $h_m$, is used to identify the three classes, as four open intervals as follows: (1) $h_m=(0°, 180°)$ is green-series colour for $\text{DOS}<50\%$, (2) $h_m=(180°, 240°)$ is blue-series colour for $50\%<\text{DOS}<70\%$, and (3) $h_m=(0°, 60°)$ or $(240°, 360°)$ is red-series colour for $\text{DOS}>70\%$, as shown in Fig. 3b. Index $s_m$ is also employed to provide the confidence level. If its value is $>0.5$ and approaches 1.0, we have high confidence to confirm the possible class.

In addition, assessment results defined by the $(h, s, v)$ values in the HSV colour space, index $h_m$ with $s_m$ and $v_m$ varying between 0 and 1, can be calculated in the normalised colour space, as [21]

$$h_{m,h} = \frac{h_m}{360}, h_{m,h} \in [0, 1] \quad (14)$$

where index, $h_{m,h}=(1/6, 1/2), (1/2, 2/3), (2/3, 1)$, or $0, 1/6$; the four open intervals can be used in computer graphic application and human user interfaces by the colour map function to display the normalised angle $h_{m,h}$ as colours. The CRA is a visual method with colour codes to represent the DOS levels and has a flexibility inference mechanism in screening applications.

### 3 Experimental results and discussions

The proposed analytical methods, including DFT, Burg AR processes, and radial-based CRA, were designed on a tablet PC or an embedded system, using LabVIEW graphical programming software and MathScript real-time (RT) module (National Instruments™ Corporation, Austin, Texas, USA), as shown in Fig. 1. The MathScript RT module was a visual programming platform, integrated into an embedded system (custom-made applications). Its platform was used to program the proposed algorithms using a high-level, graphical and text-based programming language. The PAG signals in the experimental vessel accesses were obtained using dual-channel electronic stethoscopes (3MTM Littmann®) at the pre-stenosis and post-stenosis sites, as shown in Fig. 1c. Considering the measurement sites and stenosis combinations, radial-based CRA had four inputs, (pre-PSD-1#, pre-PSD-2#) and (post-PSD-1#, post-PSD-2#), and four outputs, $(h_{c,1}, s)$ and $(h_{c,2}, s)$, as shown in Fig. 5a. The stenosis combinations were

- normal and single stenosis: 0–95% of narrowed access at the pre-stenosis (inflow site) or post-stenosis site (outflow site),
- inflow stenosis with the progression of outflow stenosis: 50 and 95% of narrowed access at the inflow site and 50–95% of narrowed access at the outflow site,
- outflow stenosis with the progression of inflow stenosis: 50 and 95% of narrowed access at the outflow site and 50–95% of narrowed access at the inflow site.

Three stenosis combinations and measurement sites are shown in Fig. 6. There were three classes based on the degree of stenotic severity, $\text{DOS}<50\%$, $50\%<\text{DOS}<70\%$, and $\text{DOS}>70\%$, as shown by the graphical programming indication. Dual-channel electronic stethoscopes were used to capture PAG signals. Each soundtrack had a recording duration of < 10 s for inconsistent sounds or heart rate detection. To match the consumer electronic stethoscope’s time-limited observation, the data length was determined by the length of an input data vector from 30,000 to 60,000 samples. Hence, the streaming acoustic data with data length of 20–40 cycles (7.5–15.0 s) and a sampling rate of 4000 Hz were decomposed into discrete frequencies in the specific frequency ranges from 0 to 1000 Hz. The FFT algorithm transformed the time-domain data length and returned 256–1024 sampling points into discrete frequency-domain representation. Then, the frequency spectra were normalised with the maximum amplitude. The procedure of PAG signal processing is shown in Fig. 3. Hence, by setting the time-domain data length of 30,000–60,000 samples ($N=30,000–60,000$) and frequency-domain representation with 1024 sampling points ($n=1024$), the optimal AR order can be determined using the Levinson–Durbin recursion algorithm [10–12]. In this study, the choice of AR order was used to achieve a good evaluation analysis by a clinician, including: (i) matching the time-limited observation (2000 Hz sampling rate and 10-second record for an electronic stethoscope), (ii) reducing...
spectral broadening and spectral leakage [10], and (iii) enhancing the frequency resolution within the observed window (data length of 30,000 samples). We suggested the AR order = 14 for constructing the Burg AR model with AR coefficients, which could offer a good resolution analysis to extract the characteristic frequencies between 200–300 and 500–600 Hz, as shown in the graph analytics in Fig. 4. Big data collection can also be analysed for insights that lead to determination of the distinct peaks of PSDs at the pre-stenosis site and post-stenosis site, respectively. Given the 171 observation/measurement data, we could specify four explanatory variables, pre-PSD-1#, pre-PSD-2#, post-PSD-1#, and post-PSD-2#, and two response variables, pre-DOS and post-DOS, in the 3D feature space. At the characteristic frequencies, the four frequency features of the scatter data referring to DOSs are shown in Fig. 7.

To model the radial-based CRA with multiple variables, the collected data sets were used to establish the radial-based function network. Relational measurement with Gaussian functions was employed to screen the similarity between the input pattern and the template patterns. The similarity degree was transformed to colour perceptual representations for separating DOS% < 50% (true negative) from DOS% > 50% (true positive). In contrast to the multiple linear regression models and machine learning method, the accuracy (hit rate) and sensitivity analysis demonstrated the superiority of screening results over the conventional ones. Experimental tests and discussion under the various stenotic situations are detailed as follows.

### 3.1 Results of multiple linear regressions

We chose two explanatory variables, pre-PSD-2# and post-PSD-2#, and two response variables, pre-DOS and post-DOS, to establish the multiple linear regression functions, where two explanatory variables had a significant relationship with the two response variables, p value < 0.001, as shown on the right-hand side of Fig. 7. Multiple linear regressions with two variables were used to analyse the scatter data, which among the explanatory variables were related to the response variables. Ordinary least square method was used to estimate the regression coefficients [23, 24]. This regression model is defined in terms of a few regression coefficients that were estimated from observation data. It took about 0.2508 s to model the forms of these relationships. Considering 4 bytes for each digital storage, the memory storage needed 3,420 bytes to store the explanatory and response data, including a regressor (3 × 171 matrix, 3 × 171 × 4 bytes) and two response vectors (2 × 171 × 4 bytes). For the 171 testing data, the positive correlation between the two explanatory variables and the two responses, as the coefficients of determination (R²), were 0.5288 and 0.7038 for pre-DOS and post-DOS estimations, respectively, as shown in Table 1. For single and coexisting inflow and outflow stenoses, the experimental results indicated > 85% of true-positive rate (0 failures for pre-DOS estimations and 16 failures for post-DOS estimations) and > 80% of accuracy (26 failures) for post-DOS estimations, as shown in Table 2. However, lower true-negative rates (13 failures) were obtained for DOS% < 50%. This indication had worse screening results for medical decision-making. Regression methods might provide good performance to the estimate the continuous response variables, in contrast to the discrete response variables used in classification applications. It did not provide a high confidence for assessing inflow and outflow stenoses. However, a sufficient quantity of data was available, and the positive correlation between the (pre-PSD-2#, post-PSD-2#) and (pre-DOS, post-DOS) had been validated.

### 3.2 Results of machine learning method

The generalised regression neural network (GRNN) [25, 26] was the adaptive machine learning method used to establish the multiple regression models with four or two explanatory variables and two response variables, as shown in Table 2. Its method includes a...
high-dimensional pattern mechanism and complex non-linear estimator to deal with the non-linear mapping techniques for prediction, interpolation, and classification applications. The measurement data was divided into two groups, 171 patterns were used as the training data and the other 171 patterns were used as testing data. The GRNN model could be directly determined using the presentation of 171 input–output pairs of training data. According to the training data, we have two or four input nodes in the input layer, 171 pattern nodes in the pattern layer, three nodes in the summation layer, and two nodes in the output layer. We had 2-171-3-2 and 4-171-3-2 topologies to construct the multiple non-linear regression models. Its iterative process took about < 10 s and < 250 iterative computations to reach the convergent condition with the prespecified tolerance value. Thus, the optimal model parameters, \( \alpha = 0.0100 \) and 0.0103, were obtained to minimise the mean squared error, \( e \leq 0.01 \). Using four frequency features as shown in Fig. 7, the prediction method was a kernel model with Gaussian distributions centred on observed scatter data for non-linear curve approximations in a high-dimensional feature space. For the 171 testing data, the experimental results indicated (i) > 95% of true-positive rate (DOS% > 50%), (ii) > 85% of true-negative rate (DOS% < 50%), and (iii) > 95% of accuracy, as shown in Table 2. The input frequency features referring to the estimated hue angles are shown in Figs. 8e and f. The radial-based CRA showed high confidence (four failures) for assessing inflow and outflow stenoses. For example, in case of coexisting inflow and outflow stenoses, with the narrowed degrees of 85 and 50%, a flow rate of 600 ml/min was used to drive the BMF through the AVG access by a roller pump with a pump speed of 40 rpm, and the in vitro AVG biophysical conditions were flow rate of 325.2 l/min and inflow pressure of 120 mmHg. The screening procedure of radial-based CRA method is detailed according to the following procedure:

Step 1 given the four frequency features, \([\text{pre-PSD}_{-1}, \text{pre-PSD}_{-2}] = [1.0000, 0.5855] \) and \([\text{post-PSD}_{-1}, \text{post-PSD}_{-2}] = [1.0000, 0.2280] \).

Step 2 computed the similarity degrees using equations, (5) and (6), as seen in Fig. 8g.

Step 3 converted the similarity degrees to the grey grades using equations, (7) and (11), and the primary colour grades could be determined, as \([g_1, b_1, r_1] = [0.6349, 0.2225, 0.4833] \) and \([g_2, b_2, r_2] = [0.0000, 0.6918, 0.3074] \).

Step 4 convert the primary colour grades to hue angles and saturations using equations, (12) and (13), as \([h_1, s_1] = [346.5542, 0.6349, 0.2225] \) and \([h_2, s_2] = [216.0088, 0.9618, 0.223] \).

Step 5 normalised the colour spaces using (14), as \([h_{1,2}, s_1] = [0.9626, 0.6349, 0.2225] \) and \([h_{1,2}, s_2] = [0.6000, 0.9608] \).

This finding confirmed that the proposed screening model could identify the level of inflow and outflow stenoses, and the saturation index, \( s \geq 0.5 \), denoted high confidence. Therefore, the higher degree of the lumen diameter in the AVG access resulted in indicating the patient needs to receive percutaneous transluminal angioplasty or surgical intervention. In this model, the radial-based function network required 2736 bytes in the memory storage unit. The CRA method had the

### Table 1: Related parameters of multiple regression models and proposed radial based CRA

| Regression model       | Explanatory variable | Response variable | Observation data | Model parameter | Time   |
|------------------------|----------------------|-------------------|------------------|-----------------|--------|
| multiple linear regression \([23, 24]\) | pre-PSD-2#, post-PSD-2# | pre-DOS#          | 171 combinations (K = 171) | \([b_{10}, b_{11}, b_{12}] = [0.6349, 0.2225, 0.0880] \) | 0.2508 s |
| non-linear regression model GRNN \([25, 26]\) | pre-PSD-2#, post-PSD-2# | post-DOS#         | \(\sigma = 0.0100\) | <10.00 s |
| radial based CRA | pre-PSD-1#, pre-PSD-2# | (h_{1,2}, s) | \(\sigma = 0.0500\) | 1.2778 s |

### Table 2: Experimental results

| Regression model | DOS | \(R^2\) | Accuracy (hit rate) | Sensitivity |
|------------------|-----|--------|---------------------|-------------|
|                |     |        | True positive rate  | True negative rate |
| multiple linear regression (two explanatory variables) | | | | |
| pre-DOs        | 0.5288 | 98.2% (171/3) | 100.0% (168/0) | 0.0% (3/3) |
| post-DOs       | 0.7038 | 84.8% (171/26) | 89.1% (148/16) | 56.5% (23/10) |
| GRNN with two/four explanatory variables | | | | |
| 1. post-DOs    | 0.9055 | 100% (171/1) | 100.0% (168/0) | 100.0% (3/3) |
| 1. post-DOs    | 0.8990 | 96.5% (171/6) | 97.3% (148/4) | 91.3% (23/2) |
| 2. post-DOs    | 0.9055 | 100% (171/0) | 100.0% (168/0) | 100.0% (3/3) |
| 2. post-DOs    | 0.8990 | 99.4% (171/1) | 99.3% (148/1) | 100.0% (23/0) |
| radial based CRA four explanatory variables | | | | |
| pre-DOs        | – | 97.6% (171/4) | 99.3% (148/1) | 86.9% (23/3) |
| post-DOs       | – | | | |

Note: (1) (Total/Failure); (2) DOS < 0.50 for Negative; (3) DOS > 0.50 for Positive; (4) Multiple Linear Regression: a. Pre-DOs = 0.8349 + 0.2225 \times \text{Post-PSD} + 0.1410 \times \text{Post-PSD} + 0.4833 \times \text{Post-PSD2}.

IET Cyber-Phys. Syst., Theory Appl., 2017, Vol. 2, Iss. 1, pp. 10–19

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
flexibility inference mechanism with colour perceptual representations and no iterative computations to update the model parameters in classification applications. The hue angle had a flexibility visual manner with colour codes, index $h_{\text{max}} \pm 60$, to realise the stenosis level. In contrast to other machine learning methods, the mechanism of the CRA method had the capability of self-regulation in the primary colour grades using similarity degrees, average grey grades, and minimum and maximum grey grades. Hence, this simple technique can be easily implemented in a tablet PC and in an intelligent embedded system. This assistant tool also integrated sufficient number of stenosis combinations and had individualised functions to evaluate the progression of inflow and outflow stenoses for patient self-healthcare demands.

Fig. 8 Experimental results

a and b Experimental results using colour perceptual representations for 171 testing data at pre-stenosis and post-stenosis sites

$e$ and $f$ Four input features refer to the hue angles at pre-stenosis and post-stenosis sites

g Similarity degree measurements

18 This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
4 Conclusion
To prevent the progression of vascular access stenosis, big spectral data and radial-based CRA were integrated into an assistant tool for detection of inflow or outflow stenosis. Under various stenosis combinations, a biophysical experimental model was established to produce complex symptomatic conditions with various roller pump speeds, heart rates, flow rates, and blood pressures. On a virtual adult hand, acoustic signals were produced and recorded by dual-channel electronic stethoscopes. Big spectral data analysis extracted the frequency-based features from a particular size of data set (171 measurement data) using the FFT algorithm and Burg AR processes. Through signal preprocessing, we obtained some suggestions to

- achieve a good resolution with selecting the 7.5–15.0 s of the streaming acoustic data and returning 1,024 sampling points into discrete frequency spectra,
- choice the AR order = 14, to find the distinguishing characteristic frequencies by Levinson-Durbin recursion algorithm and clinician decisions,

Therefore, the characteristic frequencies could be determined between 200–300 Hz and 500–600 Hz from the 171 testing data. Multiple linear regression was employed to validate the positive correlation between PSDs and DOs. Multidimensional frequency spectral data also characterised the primary components to separate correlation between PSDs and DOs. Multidimensional frequency analysis validated that these multidimensional spectral data were separable. A high-dimensional pattern mechanism could deal with the clinical study or home healthcare.

5 Acknowledge
This work was supported in part by the Ministry of Science and Technology, Taiwan, under contract number: MOST 105-2218-E-075B-001 and MOST 105-2221-E-006-087-MY2, duration: 1 March 2016–28 February 2017.

6 Reference
1 Asif, A., Gadalean, F.N., Merrill, D., et al.: ‘Inflow stenosis in arteriovenous fistulas and grafts: a multicenter, prospective study’, Kidney Int., 2005, 67, pp. 1986–1992
2 Manos, T.A., Sokolits, D.P., Giagini, A.T., et al.: ‘Local hemodynamics and intimal hyperplasia at the venous side of a porcine arteriovenous shunt’, IEEE Trans. Inf. Technol. Biomed., 2010, 14, (3), pp. 681–690
3 Stalder, A.F., Frydrychowicz, A., Ruse, M.F., et al.: ‘Assessment of flow instabilities in the healthy aorta using flow-sensitive MRI’, J. Magn. Reson. Imaging, 2011, 33, pp. 839–846
4 Chen, W.-L., Lin, Y.-H., Kan, C.-D., et al.: ‘Assessment of flow instabilities in in-vitro stenotic arteriovenous grafts using an equivalent astable multivibrator’, IET Sci. Meas. Technol., 2015, 9, (6), pp. 709–716
5 Chin, C.-H., Kan, C.-D., Chen, W.-L., et al.: ‘An equivalent astable multivibrator model to assess flow instability and dysfunction risk in in-vitro stenotic arteriovenous grafts’, Technol. Healthc. Care, 2016, 24, (2), pp. 295–308
6 KDOQI Clinical Practice Guidelines for Chronic Kidney Disease. Available at http://www.kidney.org/professionals/KDOQI/guidelines_ckd/p4_class_g1.htm
7 National Kidney Foundation: KDOQI clinical practice guidelines and clinical practice recommendations for 2006 updates: hemodialysis adequacy, peritoneal dialysis adequacy and vascular access’, Am. J. Kidney Dis., 2006, 48, (suppl. 1), pp. 1–322
8 Shen, C.C., Lin, C.H.: ‘Chirp-encoded excitation for dual-frequency ultrasound tissue harmonic imaging’, IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control, 2012, 59, (11), pp. 2420–2430
9 Lea, S.W., Fischer, P.F., Lotha, F., et al.: ‘Flow-induced ven-wall vibration in an arterio-venous graft’, J. Fluid Struct., 2005, 20, (6), pp. 837–852
10 Chen, W.-L., Chen, T., Lin, C.-H., et al.: ‘Phonangiographic with a fractional order chaotic system—a novel and simple algorithm in analyzing residual arteriovenous access stenosis’, Med. Biol. Eng. Comput., 2013, 51, (9), pp. 1011–1019
11 Chen, W.-L., Kan, C.-D., Lin, C.-H., et al.: ‘A rule-based decision-making diagnosis system to evaluate arteriovenous shunt stenosis for hemodialysis treatment of patients using fuzzy petri nets’, IEEE Biomed. Health Inf., 2014, 18, (2), pp. 703–713
12 Du, Y.-C., Chen, W.-L., Lin, C.-H., et al.: ‘Residual stenosis estimation of arteriovenous graft using a dual-channel phonangiography with fractional-order features’, IEEE J. Biomed. Health Inf., 2015, 19, (2), pp. 596–600
13 O’Donoghue, J., Herbert, J.: ‘Data management within mHealth environments: patient sensors, mobile devices, and databases’, J. Data Inf. Qual., 2012, 4, (1), pp. 5:1-5:20
14 Mirkes, E.M., Coats, T.J., Levesley, J., et al.: ‘Handling missing data in large healthcare dataset: a case study of unknown trauma outcome’, Comput. Biol. Med., 2016, 75, pp. 203–210
15 Huang, H.H., Liu, H.: ‘Big data machine learning and graph analytics: current state and future challenges’. 2014 IEEE Int. Conf. on Big Data, October 2014, doi: 10.1109/BigData.2014.7004471
16 Tejaviram, V., Selanki, H., Ravi, V., et al.: ‘Auto associative extreme learning machine based non-linear principal component regression for big data applications’, 2015 Tenth Int. Conf. on Digital Information Management, October 2015, doi: 10.1109/ICDIM.2015.7381854
17 Zheng, Y., Wu, W., Chen, Y., et al.: ‘Visual analytics in Urban computing: an overview’, IEEE Trans Big Data, 2016, 2, (3), pp. 276–296
18 Collob, C.: ‘Linear prediction and Levinson-Durbin algorithm, 2009. Available at http://ecollob.free.fr/technotes/
19 Syntax: arburg, 2016. Available at http://www.mathworks.com/help/signal/ref/arburg.html
20 EI-Overly, A.: ‘Matlab Tutorial’. Available at http://www.ee.nmt.edu/~cloesley/matlab/matlab.pdf
21 Lin, C.-H.: ‘Assessment of bilateral photoplethysmography for lower limb peripheral vascular occlusive disease using color relation analysis classifier’, Comput. Method Program Biomed., 2011, 103, (3), pp. 121–131
22 Wu, M.-J., Chen, W.-L., Kan, C.-D., et al.: ‘Dysfunction screening in experimental arteriovenous grafts for hemodialysis using fractional-order extender and color relation analysis’, Cardiovasc. Eng. Technol., 2015, 6, (4), pp. 463–473
23 Remcher, A.C., Christensen, W.F.: ‘Methods of multivariate analysis’ (Wiley Series in Probability and Statistics, 2012, 3rd edn.), ISBN: 978-0-470-17896-6
24 Cohen, J., Cohen, P., West, S.G., et al.: ‘Applied multiple regression/correlation analysis for the behavioral sciences Hillsdale’ (Lawrence Erlbaum Associates, NJ, 2nd edn.)
25 Kan, C.-D., Chen, W.-L., Lin, C.-H., et al.: ‘Optimal flow adjustment of vено-venoartarial extracorporeal membrane oxygenation with an adaptive prediction model: cannula sizes screening and pump speeds estimation’, IET Sci. Meas. Technol., 2016, 10, (3), pp. 177–184
26 Specht, D.E.: ‘A general regression neural network’, IEEE Trans. Neural Netw., 1991, 2, (6), pp. 568–576
27 Knuttel, A., Boehlau-Godau, M.: ‘Spatially confined and temporally resolved refractive index and scattering evaluation in human performed with optical coherence tomography’, J. Biomed. Opt., 2000, 5, (1), pp. 83–92