**Original Article**

# A Mathematical Method for Constraint-based Cluster Analysis towards Optimized Constrictive Diameter Smoothing of Saphenous Vein Grafts

Thomas Franz[a,*], B. Daya Reddy[b] and Peter Zilla[a]

[a]Chris Barnard Division of Cardiothoracic Surgery,

[b]Department of Mathematics and Applied Mathematics,

University of Cape Town, Cape Town, South Africa

*Send correspondence to:

Thomas Franz, PhD

Cardiovascular Research Unit

Faculty of Health Sciences

University of Cape Town

Observatory 7935

South Africa

Tel: +27 21 406 6418

Fax: +27 21 448 5935

e-mail: thomas.franz@mweb.co.za

## Abstract

This study aimed at the cluster analysis of a moderately sized data set of human saphenous veins in search of the minimum number of diameter classes required for the constrictive smoothing of saphenous vein grafts. Mathematical algorithms were developed for data selection and transformation implementing constraints for each data object. Data classes, identified with interactive pattern evaluation, were facilitated in decision-tree algorithms to cluster the data objects without using intelligent systems. The developed method proved feasible for the analysis of data obtained from 100 patients. Two diameter classes were identified as the minimum requirement for the constrictive smoothing of the 118 vein grafts. The method offers potential for the analysis of larger data sets as the algorithms can easily be implemented in programming code. The diameter classes identified may potentially be used as cluster structure or, at least, as initial estimates for an intelligent system based cluster analysis.

**Keywords**: knowledge discovery from data; data mining; data classification; mathematical algorithm;

# Nomenclature

| **Symbol** | **Unit** | **Description** |
|---|---|---|
| $C_\mathrm{S}$ | % | Required constriction degree, ensuring complete smoothing of a vein |
| $C_\mathrm{R}$ | % | Applied constriction degree |
| $D_k$ | mm | Constrictive smoothing diameter (class label) |
| $D_i^+$ | mm | Maximum individual constrictive smoothing diameter for vein $V_i$ |
| $D_i^-$ | mm | Minimum individual constrictive smoothing diameter for vein $V_i$ |
| $l$ | - | Number of vein (data object) sub sets |
| $n$ | - | Total number of veins, or data objects, in analysis |
| $n_k$ | - | Number of veins, or data objects, in a sub set |
| $OD$ | mm | Outer diameter of vein |
| $OD_\mathrm{constr}$ | mm | Outer diameter of a vein after constriction |
| $OD_\mathrm{max}$ | mm | Maximum outer diameter of a vein |
| $OD_\mathrm{min}$ | mm | Minimum outer diameter of a vein segment |
| $p_\mathrm{o}$ | % | Permissible degree of distension of a vein segment |
| $p_\mathrm{u}$ | % | Permissible degree of constriction of a vein segment |
| $x_j$ | - | Position of *OD* measurement along a vein segment |
| $\delta_i$ | mm | Individual constrictive smoothing diameter range for a vein |
| $\Delta$ | mm | Constrictive smoothing diameter range for a set of veins |
| $\Delta_k$ | mm | Constrictive smoothing diameter range for a sub set of veins |

# 1. Introduction

With the steadily increasing potential to collect and store data, in particular large data sets, the analysis and exploration of such data has received wide attention and stimulated diverse research. Knowledge discovery from data (KDD) is a multidisciplinary field which focuses on the identification of meaningful patterns in data sets. Data mining constitutes a key element of the knowledge discovery process by employing intelligent methods to extract data patterns. Additional steps such as data cleaning, data selection and knowledge presentation are usually necessary for a successful data analysis [1]. Amongst the concepts of data mining are classification and cluster analysis. Classification aims at finding a model or function that describes the data and allows distinguishing data classes. Cluster analysis refers to the grouping of data into clusters, or classes, such that data objects in the same cluster are similar to one another but dissimilar to data objects of other clusters. A main difference between classification and cluster analysis is that the former requires knowledge of class labels whereas the latter suffices without class labels being known a-priori.

Knowledge discovery from large data sets and high-dimensional data often employs intelligent systems such as neural networks [2], fuzzy approaches [3, 4] and genetic algorithms and programming [5, 6] which can form hybrid methods with more analytical knowledge discovery techniques. The applications of KDD have been as multidisciplinary as its concepts. Finance institutions use classification techniques for risk assessments [1], specialized classification algorithms have been developed for the field of personalized medicine [7] whereas categorization frameworks are facilitated in forest ecology to analyze forest structures and describe their effects on composition, dynamics and function of the ecosystems [8]. KDD methods have generally employed intelligent system when dealing with large-scale data [6, 9, 10] whereas intelligent systems may not be required for the analysis small and intermediate data sets.

In this paper, we present a simple mathematical method for the clustering of small- to medium-sized sets of single-parameter data with the objective of identifying the minimum number of classes and their labels to satisfy a set of constraints. The feasibility of the proposed method is demonstrated using a data set comprising 118 saphenous veins obtained from 100 patients with the aim of finding the smallest number of diameter values that allow for the constrictive diametric smoothing of the vein grafts.

## 2. Methods

### 2.1 Data Acquisition

For 118 saphenous vein segments of 100 patients undergoing aorto-coronary bypass surgery, the outer diameter of was recorded every 2 cm along the harvested length using a vernier calliper during post-harvest in situ leakage-test distension. The measurement procedure was approved by the institutional review boards of the University of Cape Town and informed consent was obtained from all patients.

### 2.2 Constraint-based Data Selection and Transformation

For each vein, the minimum and maximum outer diameter, $OD_{\min}$ and $OD_{\max}$, were identified. The required constriction degree, providing complete smoothing of the outer vein diameter by reducing $OD_{\max}$ to $OD_{\min}$, is:

$$C_S = \left( \frac{OD_{\max} - OD_{\min}}{OD_{\max}} \right) \cdot 100 . \tag{1}$$

The applied constriction degree, $C_R$, is defined as

$$C_R = \left( \frac{OD_{\max} - OD_{\text{constr}}}{OD_{\max}} \right) \cdot 100 , \tag{2}$$

where $OD_{\text{constr}}$ is the outer diameter of the vein after constriction.

Constraint 1: The applied constriction degree, $C_R$, was allowed to exceed the required constriction degree, $C_S$, but must not exceed a value of 50 %:

$$C_S \leq C_R \leq 50\% \, . \tag{3}$$

The maximum constriction degree $C_{R,\max}$ = 50 % proved feasible and provided the desired biological response in vein grafts, i.e. mitigation of intimal hyperplasia [11].

Constraint 2: The distension of a vein was not permitted.

With the above constraints, the individual smoothing diameter range for a vein can be defined as (see also Appendix)

$$\delta_i = \left[0.5 OD_{i,\max}, OD_{i,\min}\right] = [D_i^-, D_i^+] \tag{4}$$

with $i = 1, 2, \ldots, n$ for all veins considered for the analysis. $D_i^-$ and $D_i^+$ are the minimum and maximum individual smoothing diameter, respectively. Veins with $OD_{\max} > 2\, OD_{\min}$ were excluded from the analysis since complete smoothing is unfeasible with the constriction limit of $C_{R,\max}$ = 50 %.

Constraint 3: Any vein with $OD_{\min} < 3.0$ mm was excluded from the analysis following results of a related study which suggest the typical diameter of Internal Mammary Artery (IMA) grafts as the minimum diameter for saphenous vein grafts in the coronary position [11].

The group smoothing constriction diameter range, $\Delta$, which accommodates all $n$ veins, is defined as (see also Appendix)

$$\Delta \in \left[ \max_i D_i^-, \min_i D_i^+ \right] , \tag{5}$$

with the condition that the upper bound of the minimum individual smoothing diameter, $\max\limits_i D_i^-$, does not exceed the lower bound of the maximum individual smoothing diameter, $\min\limits_i D_i^+$:

$$\max_i D_i^- \leq \min_i D_i^+ \, . \tag{6}$$

In the case that Eq. (6) cannot be satisfied, a single constrictive smoothing diameter does not exist for the all $n$ veins. The set of veins with individual mesh diameter ranges $\delta_i$ needs to be divided in two or more sub sets with $n_k$ veins, such that Eq. (6) is satisfied for each sub set. The constrictive smoothing diameter range, $\Delta_k$, for sub set of veins is then derived from the maximum lower bound and minimum upper bound of individual constrictive smoothing diameters of the sub set:

$$\Delta_k \in \left[ \max_i D_i^-, \min_i D_i^+ \right]_k \qquad k = 1, 2, \ldots, l\,;\ i = 1, 2, \ldots, n_k\,. \tag{7}$$

Please refer to the appendix for a detailed description of the algorithm is.

### 2.3 Constraint-based Data Clustering

With the aim of minimising the number of constrictive smoothing diameters, i.e. classes, required to accommodate all veins, i.e. data objects, four class labels $D_1 < D_2 < D_3 < D_4$ were indentified interactively based on the set of individual constrictive smoothing diameter ranges $\delta_i$ with $i = 1, 2, \ldots, 118$ determined in the prior constraint-based data analysis. Two alternative decision-tree based clustering algorithms (CA) illustrated in Figure 1 were utilised to assign a vein to either the smallest (CA1) or largest (CA2) constriction diameter $D$ within the constrictive smoothing diameter range $\delta_i$ of that vein. The clustering of veins amongst the four classes (smoothing diameters) was evaluated using three parameters: 1) Number of veins assigned to each class; 2) Mean, minimum, and maximum smoothing constriction degree, $C_R$, for each class; 3) Mean constriction degree across all classes.

## 3. Results

### 3.1 Data Set

The data set obtained from 100 patients (age: 59.7 ± 8.5 years, weight: 80.3 ± 17.2 kg, gender: 64 male, 36 female) comprised 118 data objects (veins) with 1687 data points (*OD* measurements), i.e. 14.3 ± 4.7 data points per data object. The overall minimum and maximum $OD$ was 2.1 mm and 6.5 mm, and the mean $OD_{\mathrm{min}}$ and $OD_{\mathrm{max}}$ was 3.50 ± 0.61 mm and 4.77 ± 0.75 mm, respectively.

### 3.2 Data Analysis and Clustering

With the constraints for the constriction of $C_{\mathrm{R,max}} \leq 50\%$ ($p_{\mathrm{u}} = 50\%$) and the distension of $p_{\mathrm{o}} = 100\%$, 117 veins satisfied the condition $OD_{\mathrm{i,max}} \leq 2OD_{\mathrm{i,min}}$ (Eq. A.6), namely $OD_{\mathrm{i,max}} < 2OD_{\mathrm{i,min}}$ for 115 veins and $OD_{\mathrm{i,max}} = 2OD_{\mathrm{i,min}}$ for 2 veins. The individual constrictive smoothing diameter range $\delta_{\mathrm{i}}$ based on the specified constraints is illustrated in Figure 2 for the entire set of 118 veins. The diameter range $\delta_{\mathrm{i}} \leq 0$ indicated the vein with $OD_{\mathrm{i,max}} > 2OD_{\mathrm{i,min}}$ . For this set of $\delta_{\mathrm{i}}$ ($i = 1, 2, \ldots 117$) with $\max_i D_i^- = 3.25$ mm and $\min_i D_i^+ = 2.10$ mm, a single constrictive smoothing diameter Δ did not exist since the necessary condition $\max_i D_i^- \leq \min_i D_i^+$ was not satisfied.

Four constrictive smoothing diameters were identified interactively as class labels for the clustering of the 117 veins: $D_1 = 3.0$ mm, $D_2 = 3.3$ mm, $D_3 = 3.6$ mm and $D_4 = 3.9$ mm. $D_1$ complied with the constraint for the minimum vein graft diameter $OD_{\mathrm{min}} \geq 3.0$ mm. Of the 117 veins, 14 veins were excluded from further analysis due to $D_i^+ < 3.0$ mm.

The clustering of the remaining 103 veins with algorithm CA1 succeeded with two classes: 95 and 8 veins were assigned to $D_1$ and $D_2$, respectively, whereas $D_3$ and $D_4$ were not populated (see Table 1). The mean constriction degree in the classes $D_1$ and $D_2$ was $C_{\mathrm{R}}(D_1) = 35.8 \pm 8.6$ % and $C_{\mathrm{R}}(D_2) = 47.7 \pm 1.4$ %. When employing clustering algorithm CA2, the distinction between three solutions comprising two ($D_1$, $D_2$), three ($D_1$, $D_2$, $D_3$), and four ($D_1$, $D_2$, $D_3$, $D_4$) classes was necessary. For the two-class solution, $n_1 = 27$ and $n_2 = 76$ veins were assigned to the diameter classes $D_1$ and $D_2$, respectively. Adding the third class, $D_3 = 3.6$ mm, resulted in re-assignment of 51 from $D_2$ to $D_3$, hence $n_1 = 27$, $n_2 = 25$ and $n_3 = 51$. For the four-class solution, 33 veins were reassigned from $D_3$ to $D_4 = 3.9$ mm ($n_1 = 27$, $n_2 = 25$, $n_3 = 18$, $n_4 = 33$).

## 4. Discussion

In this study, a simple mathematical method was developed for the constraint-based analysis and supervised clustering of small to medium-sized single parameter data sets. The method was motivated by the search for the minimum number of diameter classes, and their class labels, required for the constrictive smoothing of saphenous vein grafts of 100 patients.

The developed method comprises several elements of knowledge discovery from data, namely data selection, transformation, pattern evaluation and cluster analysis [1].

The data selection and transformation process served to extract the extreme, i.e. minimum and maximum, data points and to implement constraints on these data points for each individual data object. Pattern evaluation of the transformed data of all data objects provided the basis for supervised (interactive) identification of the number of classes required and their labels. Using the identified classes, the clustering was performed employing decision-tree algorithms. Two algorithms were proposed to allow for different prioritisation of the clustering while using the same class labels.

Many clustering algorithms require prior knowledge of the number of clusters to find [12]. Various methods for the estimation of that number have been suggested in the statistical literature. The supervised method of class label identification employed here, in conjunction with targeting moderately sized data sets, allowed for a simple clustering method without prior knowledge of the number of classes. The method also succeeded without the need for intelligent systems such as genetic algorithms, fuzzy techniques and machine learning which have been employed extensively for clustering of large and multidimensional data sets [3, 7-9, 13]. The proposed method was applied to a set of 118 data objects (human saphenous vein grafts) with a total of 1687 data points (measured diameter values). The evaluation of each data object under implementation of data object specific constraints (limits for constriction and dilation) and data set specific constraints (global minimum for vein graft diameter) resulted in exclusion of 15 data objects (i.e. 12% of the entire data set) unfit for the objective of the cluster analysis. For each of the fit data objects, the upper and lower thresholds of the data range were determined. Using interactive pattern evaluation, the minimum number of classes for the 103 data objects was found to be two. The subsequent clustering of the data was performed with the minimum of two classes as well as with three and four classes to evaluate the two alternative clustering algorithms and the resulting distribution of data objects in the different classes.

The data objects were most unevenly clustered when employing clustering algorithm CA1 which prioritised the lowest diameter class for data objects that qualified for more than one class ($D_1$: 92.2 %, $D_2$: 7.8 %, $D_3 = D_4 = 0$ %). The most even population of the data classes was obtained with four classes and prioritisation of the highest diameter class for multiple-class data objects ($D_1$: 26.2 %, $D_2$: 24.3 %, $D_3$: 17.5 %, $D_4$: 32.0 %). The even distribution resulted in a 24.3% decrease in grand mean constriction of the vein grafts ($C_R = 27.8 \pm 9.5$ %) compared to the uneven two-class distribution favouring low diameter classes ($C_R = 36.7 \pm 8.8$ %). The two- and three-class clusters with high diameter class prioritisation provided intermediate distribution patterns and overall constriction degrees.

While the size of the data set was small enough to allow for clustering without intelligent systems, the data source of 100 patients allows for a cautious indication that the validity of the diameter classes found extends to larger data sets. The analysis and clustering algorithms may be easily implemented in a programming code while utilising the four diameter classes proposed, at least as a starting point. The fact that two diameter classes ($D_3$ and $D_4$) were not populated when employing low-class prioritisation indicated the capacity of the method to accommodate data with a higher upper threshold, i.e. larger diameter values.

## 5. Conclusions

In combining the KDD elements of data selection, data transformation, pattern evaluation and decision-tree algorithms, a simple mathematical method was developed for the constraints-based cluster analysis of a moderately sized single parameter data set without the need for intelligent systems. The method was successfully applied to analyse a set of diametric data of 118 human saphenous veins obtained from 100 patients with the objective of identifying the minimum number of diameter classes that allows constrictive diametric smoothing of the entirety and majority, respectively, of vein grafts. The method offers potential for analysing larger data sets as the algorithms for data selection, transformation and clustering can easily be implemented in programming code. The diameter classes identified in this study can potentially be utilized as cluster structure for larger data sets or can, at least, serve as initial estimates for an intelligent system based cluster analysis.

## Acknowledgements

The authors wish to acknowledge Dr L Moodley, Dr W Lichtenberg, Dr J Scherman, Dr H Legodi, Dr G Mphahlele, and Dr I Taunyane of the Chris Barnard Division of Cardiothoracic Surgery, Groote Schuur Hospital, Cape Town, for performing the saphenous vein dimensional measurements. This research was mainly funded through a research collaboration grant by Medtronic Inc. (Minneapolis, MN, USA) to the University of Cape Town and a THRIP grant of the National Research Foundation (South Africa). The salary of TF was fully funded from the Medtronic research grant.

## References

[1] Han J, Kamber M. Data Mining: Concepts and Techniques. 2 ed. Amsterdam: Elsevier; 2006.

[2] Durbin B, Dudoit S, van der Laan MJ. A deletion/substitution/addition algorithm for classification neural networks, with applications to biomedical data. Journal of Statistical Planning and Inference 2008;138: 464-488.

[3] Hu Y-C, Chen R-S, Tzeng G-H. Finding fuzzy classification rules using data mining techniques. Pattern Recognition Letters 2003;24: 509-519.

[4] Paetz J. Supervised Neuro-fuzzy Clustering for Life Science Applications. In: Maglaveras N, Chouvarda I, Koutkias V, Brause R, editors. Biological and Medical Data Analysis. Berlin: Springer; 2006, p. 378-389.

[5] Zhang H, Ishikawa M. Bagging using hybrid real-coded genetic algorithm with pruning and its applications to data classification. International Congress Series 2007;1301: 184-187.

[6] Zhang Y, Bhattacharyya S. Genetic programming in classifying large-scale data: an ensemble method. Information Sciences 2004;163: 85-101.

[7] Moon H, Ahn H, Kodell RL, Baek S, Lin C-J, Chen JJ. Ensemble methods for classification of patients for personalized medicine with high-dimensional data. Artif Intell Med 2007;41: 197-207.

[8] Nadkarni NM, McIntosh ACS, Cushing JB. A framework to categorize forest structure concepts. Forest Ecology and Management 2008;256: 872-882.

[9] Vieira A, Barradas N. A training algorithm for classification of high-dimensional data. Neurocomputing 2003;50: 461-472.

[10] Xu P, Brock GN, Parrish RS. Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. Computational Statistics & Data Analysis 2009;53: 1674-1687.

[11] Zilla P, Human P, Wolf M, Lichtenberg W, Rafiee N, Bezuidenhout D, Samodien N, Schmidt C, Franz T. Constrictive external nitinol meshes inhibit vein graft intimal hyperplasia in nonhuman primates. The Journal of Thoracic and Cardiovascular Surgery 2008;136: 717-725.

[12] Mardia KV, Kent JT, Bibby JM. Multivariate Analysis. San Diego: Academic Press; 1980.

[13] Armand S, Watelain E, Mercier M, Lensel G, Lepoutre F-X. Identification and classification of toe-walkers based on ankle kinematics, using a data-mining method. Gait Posture 2006;23: 240-248.

# Appendix: Derivation of Algorithm for Constraint-based Data Selection and Transformation

A number of harvested veins $V_i = V_1, V_2, \ldots, V_i, \ldots V_n$ were considered where each vein has an outer diameter $OD_i(x_j)$ at the measurement points $x_j = x_1, x_2, \ldots, x_m$ equally spaced along the vein length. It is desired to find a constriction diameter $D$, to ensure complete diameter smoothing of all veins, with the following constraints:

a) The distension of a vein is permitted to a maximum of $p_\mathrm{o}$ % of the outer vein diameter at each measurement point:

$$D \le \frac{p_o}{100} OD_i\left(x_j\right) \qquad \text{with } i = 1, 2, \ldots, n\ , j = 1, 2, \ldots, m\ . \tag{A.1}$$

b) The constriction of a vein must not exceed $p_\mathrm{u}$ % of the outer vein diameter at each measurement point:

$$D \ge \frac{p_u}{100} OD_i\left(x_j\right) \qquad \text{with } i = 1, 2, \ldots, n\ , j = 1, 2, \ldots, m\ . \tag{A.2}$$

After combining and rewriting Eqs. (A.1) and (A.2), the condition for the constrictive smoothing diameter takes the form:

$$\frac{p_u}{100} OD_i\left(x_j\right) \le D \le \frac{p_o}{100} OD_i\left(x_j\right). \tag{A.3}$$

A constrictive smoothing diameter for an individual vein exists on the following conditions:

$$\max_j \frac{p_u}{100} OD_i\left(x_j\right) \le \min_j \frac{p_o}{100} OD_i\left(x_j\right) \tag{A.4}$$

which can be rewritten as

$$\frac{p_u}{100} OD_{i,\max} \le \frac{p_o}{100} OD_{i,\min} \tag{A.5}$$

and

$$OD_{i,\max} \le \frac{p_o}{p_u} OD_{i,\min}\ . \tag{A.6}$$

The range of the individual constrictive smoothing diameter, for a single vein $V_i$, is then the closed interval $\delta_i$:

$$\delta_i = \left[ \frac{p_u}{100} OD_{i,\max}, \frac{p_o}{100} OD_{i,\min} \right] = \left[ D_i^-, D_i^+ \right] \tag{A.7}$$

Figure A.1 illustrates for a single vein the outer vein diameter, minimum and maximum constriction diameter at each measurement point along the vein length as well as the range of the constrictive smoothing diameter $\delta_i$.

The constrictive smoothing diameter range Δ which accommodates a number of veins $V_1$, $V_2$, …,$V_\mathrm{i}$, … $V_n$, with individual constriction diameter ranges $\delta_1$, $\delta_2$, …, $\delta_i$, …, $\delta_n$, can be derived from the upper bound of the minimum individual constriction diameter and the lower bound of the maximum individual constriction diameter:

$$\Delta \in \left[ \max_i D_i^-, \min_i D_i^+ \right] \qquad \text{with } i = 1, 2, \ldots, n \tag{A.8}$$

with the condition

$$\max_i D_i^- \le \min_i D_i^+ . \tag{A.9}$$

Figure A.2 illustrates the individual mesh diameter ranges $\delta_i$ for a number of veins, ranked according to minimum and maximum constriction diameter, and the resulting group constriction diameter range Δ accommodating all veins.

<u>Exception 1:</u> An individual constriction diameter range, $\delta_i$, complying with Eqs. (A.1) and (A.2), cannot be found for a vein $V_i$ if the minimum of the constriction diameter exceeds the maximum of the constriction diameter:

$$\frac{p_u}{100} OD_{i,\max} > \frac{p_o}{100} OD_{i,\min} , \tag{A.10}$$

which can be rewritten as

$$OD_{i,\max} > \frac{p_o}{p_u} OD_{i,\min} . \tag{A.11}$$

Any vein which does not satisfy Eq. (A.6) will be excluded from further analysis.

Exception 2: A single constriction diameter range, $\mathit{\Delta}$, accommodating a number of veins $V_{\mathrm{i}}$ cannot be found if the upper bound of the minimum individual constriction diameter exceeds the lower bound of the maximum individual constriction diameter, see Eq. (A.9), i.e.:

$$\max_i D_i^- > \min_i D_i^+ \qquad i = 1, 2, \ldots, n\,. \qquad \text{(A.12)}$$

In such a case, it is required to divide the set of $n$ veins in two or more sub sets with $n_k$ veins in each sub set such that upper bound of the minimum individual constriction diameter does not exceed the lower bound of the maximum individual constriction diameter for each sub set:

$$\left(\max_i D_i^-\right)_k \le \left(\min_i D_i^+\right)_k \qquad \text{with } k = 1, 2, \ldots, l \text{ and } i = 1, 2, \ldots, n_k\,. \qquad \text{(A.13)}$$

The admissible range for the constriction diameter $\Delta_k$ for each sub set of veins is then

$$\Delta_k \in \left[\left(\max_i D_i^-\right)_k, \left(\min_i D_i^+\right)_k\right] \qquad \text{with } k = 1, 2, \ldots, l \text{ and } i = 1, 2, \ldots, n_k\,. \qquad \text{(A.14)}$$

Table 1. Distribution of veins and constriction degrees for two, three, and four classes obtained with clustering algorithms CA1 and CA2.

| | CA1 | CA2 | | |
|---|---|---|---|---|
| | 4 classes | 2 classes | 3 classes | 4 classes |
| Excluded | | | | |
| n | 14 | 14 | 14 | 14 |
| $D_1$ [mm] | 3.0 | | | |
| $n$ | 95 | 27 | 27 | 27 |
| $C_R$ [%] | 35.8±8.6 | 29.6±9.8 | 29.6±9.8 | 29.6±9.8 |
| $C_{R,min}$ [%] | 9.1 | 9.1 | 9.1 | 9.1 |
| $C_{R,max}$ [%] | 50 | 47.4 | 47.4 | 47.4 |
| $D_2$ [mm] | 3.3 | | | |
| $n$ | 8 | 76 | 25 | 25 |
| $C_R$ [%] | 47.7±1.4 | 33.7±8.4 | 31.2±9.4 | 31.2±9.4 |
| $C_{R,min}$ [%] | 45.9 | 13.2 | 13.2 | 13.2 |
| $C_{R,max}$ [%] | 49.2 | 49.2 | 47.6 | 47.6 |
| $D_3$ [mm] | 3.6 | | | |
| $n$ | 0 | N/A | 51 | 18 |
| $C_R$ [%] | | | 29.0±8.4 | 26.8±9.4 |
| $C_{R,min}$ [%] | | | 14.3 | 14.3 |
| $C_{R,max}$ [%] | | | 44.6 | 44.6 |
| $D_4$ [mm] | 3.9 | | | |
| $n$ | 0 | N/A | N/A | 33 |
| $C_R$ [%] | | | | 24.3±8.4 |
| $C_{R,min}$ [%] | | | | 9.3 |
| $C_{R,max}$ [%] | | | | 40 |
| Overall | | | | |
| $n$ | 103 | 103 | 103 | 103 |
| $C_R$ [%] | 36.7±8.8 | 32.6±9.0 | 29.7±9.0 | 27.8±9.5 |
| $C_{R,min}$ [%] | 9.1 | 9.1 | 9.1 | 9.1 |
| $C_{R,max}$ [%] | 50 | 49.2 | 47.6 | 47.6 |

## Figure Captions

Figure 1. Decision tree charts of the clustering algorithms employed to assign a data object (vein) to one class (smoothing diameter) of a four-class solution. Algorithms CA1 and CA2 resulted in a vein to be assigned to the smallest and largest smoothing diameter, respectively, suitable for smoothing diameter range of the vein.

Figure 2. Graph showing the individual constrictive smoothing diameter ranges, $\delta_i$, for the entire set of 118 veins based on the conditions of distension being prohibited and constriction of $C_{R,max} \leq 50\%$. The negative smoothing diameter range indicates the vein excluded from analysis due to the lower smoothing diameter bound exceeding the upper smoothing diameter bound.

Figure A-1. Graph illustrating the outer vein diameter $OD_i(x_j)$, upper bound $p_o/100\ OD_i(x_j)$ and lower bound $p_u/100\ OD_i(x_j)$ of the constrictive smoothing diameter at measurement points $x_j$ along the length of a harvested vein $V_i$ . The individual constrictive smoothing diameter range $\delta_i$ is indicated.

Figure A-2. Graph showing individual mesh diameter ranges $\delta_i$ for $n = 14$ veins ranked according to the minimum and maximum individual smoothing diameter $D_i^-$ and $D_i^+$ (range, minimum and maximum smoothing diameter are indicated for vein #6). The constrictive smoothing diameter range $\Delta$ accommodating all $n$ veins is determined by the grand maximum of the minimum individual constrictive smoothing diameter (max $D_i^-$) and the grand minimum of the maximum individual constrictive smoothing diameter (min $D_i^+$).

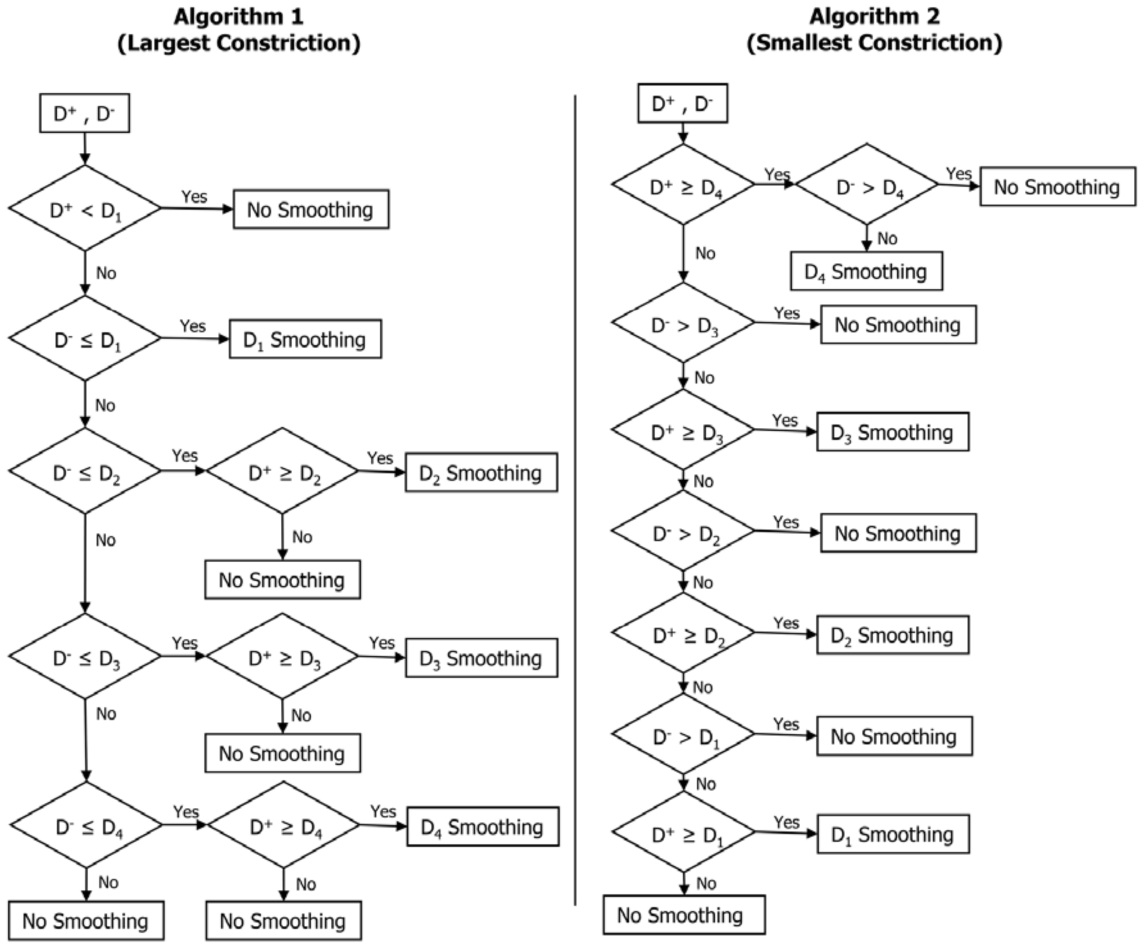

Figure1. Decision tree charts of the clustering algorithms employed to assign a data object (vein) to one class (smoothing diameter) of a four-class solution. Algorithms CA1 and CA2 resulted in a vein to be assigned to the smallest and largest smoothing diameter, respectively, suitable for smoothing diameter range of the vein.

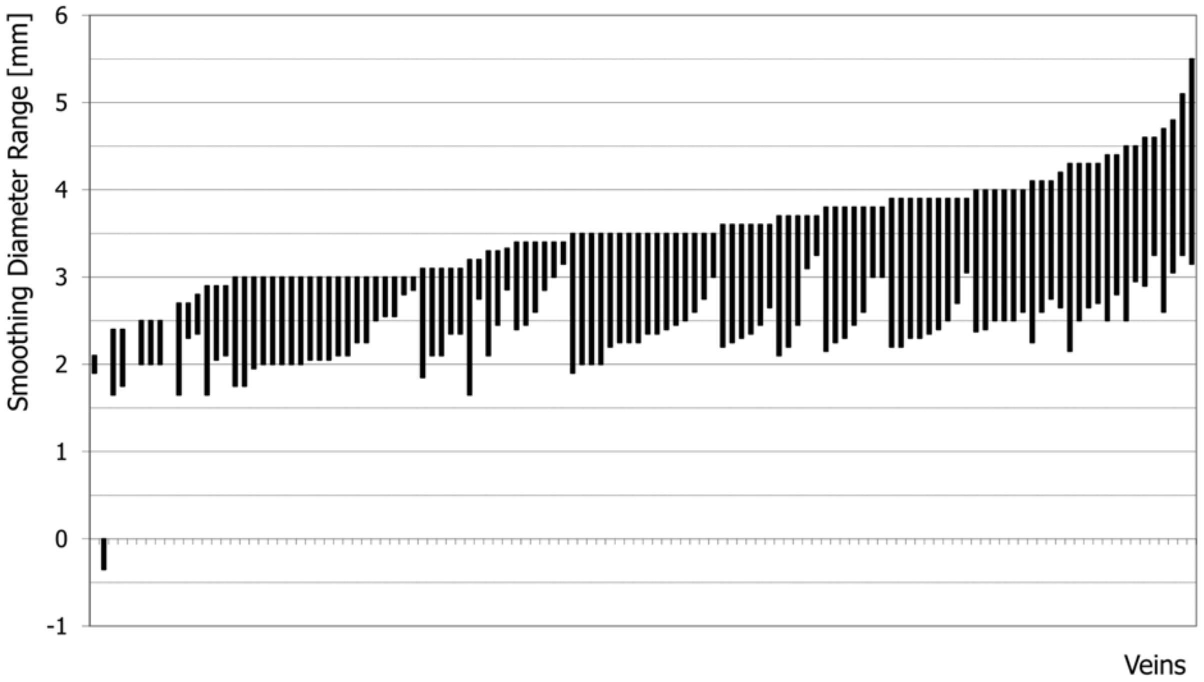

Figure 2. Graph showing the individual constrictive smoothing diameter ranges, $\delta_i$, for the 118 veins based on the conditions of distension being prohibited and constriction of $C_{R,max} \leq 50\%$. The negative smoothing diameter range indicates the vein excluded from analysis due to the lower smoothing diameter bound exceeding the upper smoothing diameter bound.

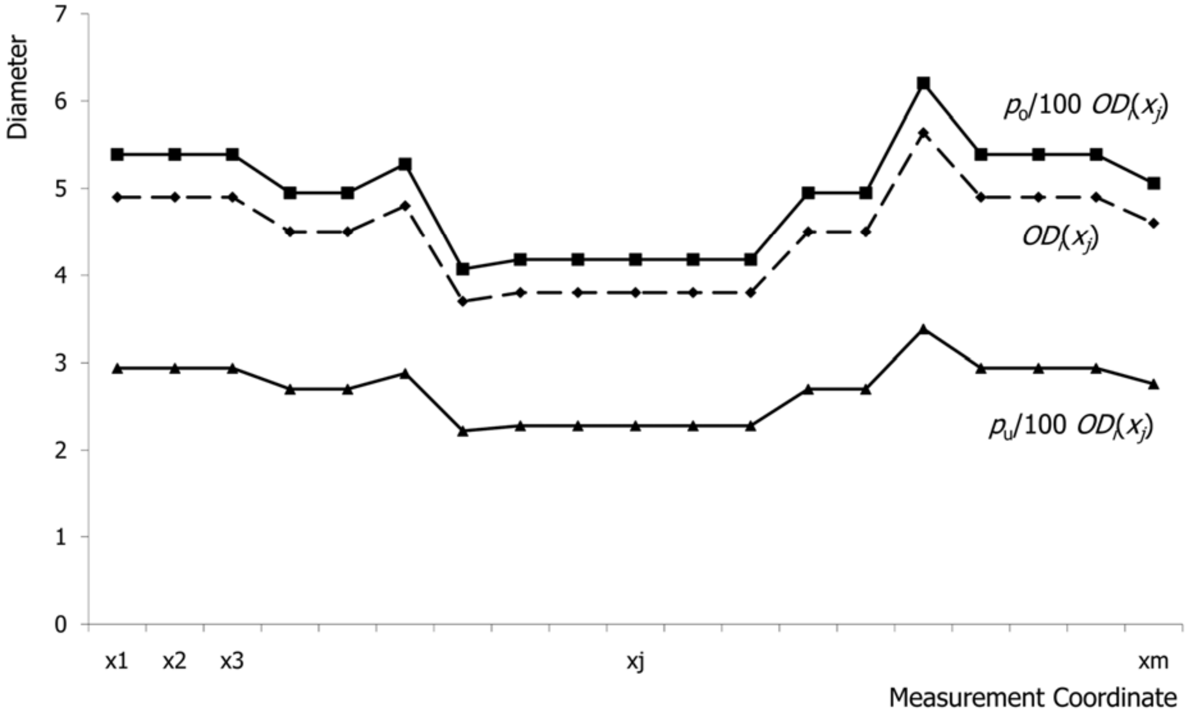

Figure A-1. Graph illustrating the outer vein diameter $OD_i(x_j)$, upper bound $p_o/100\ OD_i(x_j)$ and lower bound $p_u/100\ OD_i(x_j)$ of the constrictive smoothing diameter at measurement points $x_j$ along the length of a harvested vein $V_i$ . The individual constrictive smoothing diameter range $\delta_i$ is indicated.

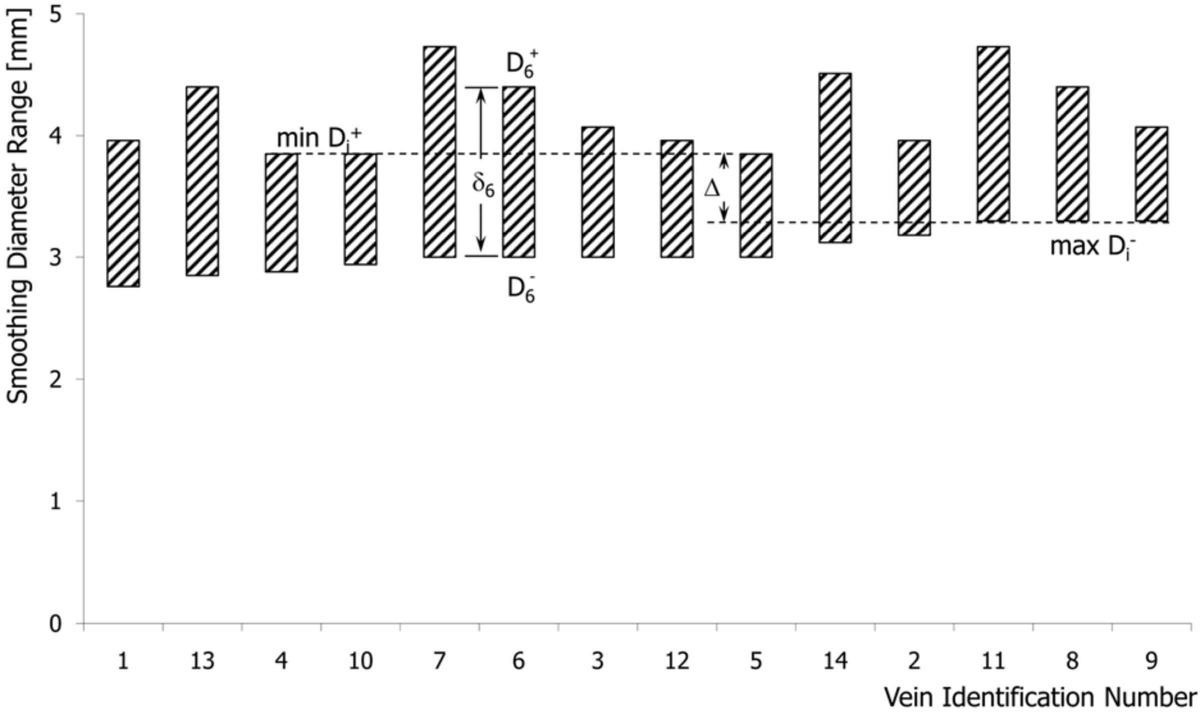

Figure A-2. Graph showing individual mesh diameter ranges $\delta_i$ for $n = 14$ veins ranked according to the minimum and maximum individual smoothing diameter $D_i^-$ and $D_i^+$ (range, minimum and maximum smoothing diameter are indicated for vein #6). The constrictive smoothing diameter range $\Delta$ accommodating all $n$ veins is determined by the grand maximum of the minimum individual constrictive smoothing diameter (max $D_i^-$) and the grand minimum of the maximum individual constrictive smoothing diameter (min $D_i^+$).