
Downloaded from: http://sure.sunderland.ac.uk/id/eprint/9553/

Usage guidelines

Please refer to the usage guidelines at http://sure.sunderland.ac.uk/policies.html or alternatively contact sure@sunderland.ac.uk.
Data Mining in Vascular Surgery

Data Mining and Associated Analytical Tools as Decision Aids for Healthcare practitioners in Vascular Surgery

Reza Mofidi

This Commentary is submitted in partial fulfilment of the requirements of the University of Sunderland for the degree of PhD by existing Publications

March 2018
Dedication

To Cian, the bravest little guy I know
**Declarations**

In presenting this commentary for assessment, I declare that it is a final copy including any last revisions. I also declare that it is entirely the result of my own work other than where sources are explicitly acknowledged and referenced within the body of the text. This work has not been previously submitted, in whole or in substantial part, for any degree at this or any other institution.

**Acknowledgments**

I would like to thank my numerous co-authors and colleagues without whom it would not have been possible to complete the publications included in this commentary. I would also like to thank my supervisors Professor Scott Wilkes and Dr Kenneth McGarry for their advice, supervision and critical appraisal of this document.
Abstract

Vascular surgery is an increasingly data rich speciality. Planning treatment and assessing outcomes are highly dependent on objective assessment of number of imaging modalities including duplex ultrasound, CT scans and angiograms which are almost exclusively digitally created stored and accessed. Developments such as the national vascular registry mean that treatment outcomes are recorded scrutinised electronically. The widespread availability of data which is collected electronically and stored for future clinical use has created the opportunity to examine the efficacy of investigations and treatments in a way which has hitherto not been possible. In addition, new computational methods for data analysis have provided the opportunity for the clinicians and researchers to utilise this data to address pertinent clinical questions.

This narrative describes 9 studies which utilized data mining methodology to examine a range of different clinical problems in the fields of general and vascular surgery. These publications have been collectively cited 204 times in the literature. They study the natural history and outcomes of treatment for abdominal aortic aneurysms (AAA) in women, assessment of risk of failure following infrainguinal vein graft bypass surgery, accurate assessment of the degree of internal carotid artery stenosis using Duplex US and prediction of outcome following resection of oesophageal carcinoma. These studies utilise conventional statistical models such as linear, logistic and Cox’s regression analyses, as well as intelligent classification methods such as Classification and regression Trees (CART) and Multi-layered Perceptron (MLP) artificial neural network to address each clinical question.

Some of the clinical problems being addressed by this narrative such as AAAs in women
have only recently gained their due importance. Others such as being able to accurately assess the degree of ICA stenosis using Duplex US have gained prominence because of development of the Oxford stroke risk calculator and the realisation that minimising symptom to carotid endarterectomy time in reduces risk of stroke.

Above all this narrative illustrates how data mining is fast becoming an essential tool in understanding and developing the evidence base for vascular surgery. Understanding the main issues underlying data mining processes is important for success in the young but expanding fields of medical informatics and vascular surgery.
List of Publications on which this Commentary is based

1. Influence of sex on expansion rate of abdominal aortic aneurysms
   R Mofidi, VJ Goldie, J Kelman, ARW Dawson, JA Murie, RTA Chalmers
   Impact factor: 5.542
   Cited 81 times in the literature

   R Mofidi, SA Suttie, A Howd, ARW Dawson, GD Griffiths, PA Stonebridge
   British Journal of Surgery 2008: 95 (12), 1475-1479
   Impact factor: 5.542
   Cited 8 times in the literature

3. Significance of the early postoperative duplex result in infrainguinal vein bypass surveillance
   R Mofidi, J Kelman, O Berry, S Bennett, JA Murie, ARW Dawson
   Impact factor: 2.549
   Cited 41 times in the literature

4. Balloon angioplasty as the primary treatment for failing infra-inguinal vein grafts
   R Mofidi, M Flett, J Nagy, R Ross, GD Griffiths, S Chakraverty,
   Impact factor: 2.549
   Cited 30 times

5. The value of vein graft surveillance in bypasses performed with small-diameter vein grafts
   Impact factor: 1
   Cited 9 times in the literature

6. Development of a Decision Tree to Streamline Infrainguinal Vein Graft Surveillance
   OMB McBride, R Mofidi, GD Griffiths, AR Dawson, RTA Chalmers, PA Stonebridge
   Impact factor: 1
   Cited 5 times in the literature

7. Validation of a decision tree to streamline infrainguinal vein graft surveillance
   R Mofidi, OM McBride, BR Green, TGatenby, P Walker, S Milburn.
   Impact factor: 1
   Cited 3 times

8. Prediction of the exact degree of internal carotid artery stenosis using an artificial neural network based on duplex velocity measurements
   R Mofidi, TI Powell, A Brabazon, D Mehigan, SJ Sheehan, DP MacErlaine, TV Keaveny
   Annals of vascular surgery 2005: 19 (6), 829-837
   Impact factor: 1
   Cited 7 times

9. Prediction of survival from carcinoma of oesophagus and oesophago-gastric junction following surgical resection using an artificial neural network
   R Mofidi, C Deans, MD Duff, AC de Beaux, SP Brown
   European Journal of Surgical Oncology (EJSO) 2006: 32 (5), 533-539
   Impact factor: 3.244
   Cited 21 times
# Table of Contents

Chapter-1  Background and Introduction to the clinical contexts………………………………….1

*Introduction* ........................................................................................................................................1

*The Mission: Data mining Paradigms* ...............................................................................................6

*Data Mining Tools: (Methods by which datamining is implemented)* ........................................11

  Deterministic Analytical Models ....................................................................................................13

  Intelligent Analytical Models ..........................................................................................................13

  *Decision tree analysis* ..................................................................................................................13

  *Artificial Neural Networks* ..........................................................................................................15

*Data mining and decision support in healthcare* ...........................................................................18

*The Clinical Context for Included Publications* ............................................................................25

Chapter -2   Methods and Results (the CRISP-DM data mining Process) .................................39

2.1 Study of Associations Using Deterministic Classification .....................................................41

  2.1.1 Assessment of Growth rate of AAAs (Paper-1) ..................................................................41

  2.1.2 Assessment of outcome following repair of AAA in Scotland (Paper-2) .........................48

2.2 Sequence (Path) Analysis Using a Combination of Deterministic and Intelligent Methods ....54

  2.2.1 Maintaining patency of Infra-inguinal grafts (Papers-3,4,5,6,7) .......................................54

2.3 Classification using Intelligent Analysis ....................................................................................82

  Assessment of degree of ICA stenosis using ANN (Paper-8)

2.4 Prediction of Outcome using Intelligent Analysis ....................................................................89

  Prediction of outcome following resection of Oesophageal cancer using ANN (Paper-9)

Chapter -3   Discussion ......................................................................................................................96

Chapter -4   Concluding remarks ......................................................................................................115
Appendices

Appendix-1 (Book Chapter) The Evidence for Management of Abdominal Aortic Aneurysms: Lessons Learned from Randomised Controlled Trials.

Appendix-2 (Paper): Influence of sex on expansion rate of abdominal aortic aneurysms


Appendix-4 (Paper): Significance of the early postoperative duplex result in infra-inguinal vein bypass surveillance

Appendix-5 (Paper): Balloon angioplasty as the primary treatment for failing infra-inguinal vein grafts

Appendix-6 (Paper): The value of vein graft surveillance in bypasses performed with small-diameter vein grafts

Appendix-7 (Paper): Development of a Decision Tree to Streamline Infra-inguinal Vein Graft Surveillance

Appendix-8 (Paper): Validation of a decision tree to streamline infra-inguinal vein graft surveillance

Appendix-9 (Paper): Prediction of the exact degree of internal carotid artery stenosis using an artificial neural network based on duplex velocity measurements

Appendix-10 (Paper): Prediction of survival from carcinoma of oesophagus and oesophago-gastric junction following surgical resection using an artificial neural network
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table-1</td>
<td>Variables included in the AAA surveillance clinic dataset</td>
<td>37</td>
</tr>
<tr>
<td>Table-2</td>
<td>Predictors of rapid aneurysm growth</td>
<td>42</td>
</tr>
<tr>
<td>Table-3</td>
<td>Factors associated with death from AAA</td>
<td>48</td>
</tr>
<tr>
<td>Table-4</td>
<td>Duplex US criteria identifying different categories of vein graft stenosis</td>
<td>55</td>
</tr>
<tr>
<td>Table-5</td>
<td>Patient demographics stratified by the VGD</td>
<td>55</td>
</tr>
<tr>
<td>Table-6</td>
<td>Factors affecting the need for re-intervention in infra-inguinal vein grafts.</td>
<td>60</td>
</tr>
<tr>
<td>Table-7</td>
<td>Factors affecting primary assisted patency after infra-inguinal bypass assessed using Cox’s multivariate regression analysis.</td>
<td>62</td>
</tr>
<tr>
<td>Table-8</td>
<td>Factors affecting amputation rates after infra-inguinal bypass assessed using Cox’s multivariate regression analysis.</td>
<td>62</td>
</tr>
<tr>
<td>Table-9</td>
<td>Cox regression analysis of potential factors associated with infra-inguinal vein graft bypass graft revision</td>
<td>65</td>
</tr>
<tr>
<td>Table-10</td>
<td>Indicators of accuracy of classification of the CART (decision tree) analysis for prediction of development of vein graft stenosis or occlusion within the first 2 years of follow up.</td>
<td>67</td>
</tr>
<tr>
<td>Table-11</td>
<td>Factors affecting primary assisted patency after infra-inguinal bypass surgery</td>
<td>76</td>
</tr>
<tr>
<td>Table-12</td>
<td>Factors affecting amputation rates after infra-inguinal bypass surgery</td>
<td>76</td>
</tr>
<tr>
<td>Table-13</td>
<td>The comparison between duplex velocity criteria and ANN at identification of &gt;50% and &gt;70% stenosis</td>
<td>83</td>
</tr>
<tr>
<td>Table-14</td>
<td>The pathological feature of the lesions based on the examination of the histological specimen</td>
<td>87</td>
</tr>
<tr>
<td>Table-15</td>
<td>The sample size and composition of training and evaluation datasets of ANN analysis</td>
<td>88</td>
</tr>
<tr>
<td>Table-16</td>
<td>Sensitivity Analysis for the variables used to construct the ANNs</td>
<td>90</td>
</tr>
</tbody>
</table>
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure-1</td>
<td>Graphic representation of a decision tree Precision tree™</td>
<td>11</td>
</tr>
<tr>
<td>Figure-2</td>
<td>A schematic of a simple artificial neural network</td>
<td>14</td>
</tr>
<tr>
<td>Figure-3</td>
<td>Visual representation of the processes involved in CRISP-DM</td>
<td>17</td>
</tr>
<tr>
<td>Figure-4</td>
<td>The steps involved in SEMMA data mining methodology</td>
<td>18</td>
</tr>
<tr>
<td>Figure-5</td>
<td>An overview of the steps involved in data mining (KDD) process</td>
<td>21</td>
</tr>
<tr>
<td>Figure-6</td>
<td>Serial Ultrasound examinations of the same asymptomatic AAA performed 5 years apart</td>
<td>22</td>
</tr>
<tr>
<td>Figure-7</td>
<td>Preoperative CT scan of a large intact AAA</td>
<td>24</td>
</tr>
<tr>
<td>Figure-8</td>
<td>Aetiology of vein graft stenosis is classified chronologically</td>
<td>25</td>
</tr>
<tr>
<td>Figure-9</td>
<td>Tight Proximal graft stenosis identified on Duplex US</td>
<td>28</td>
</tr>
<tr>
<td>Figure-10</td>
<td>Oxford stroke risk calculator</td>
<td>30</td>
</tr>
<tr>
<td>Figure-11</td>
<td>Carotid Duplex Ultrasound exam of right carotid bifurcation</td>
<td>31</td>
</tr>
<tr>
<td>Figure-12</td>
<td>CRISP-DM Process Applied to Study-</td>
<td>35</td>
</tr>
<tr>
<td>Figure-13</td>
<td>AP diameters of abdominal aortic aneurysm on initial presentation in men and women</td>
<td>40</td>
</tr>
<tr>
<td>Figure-14</td>
<td>Correlation between expansion rate and initial AP diameter of the abdominal aortic aneurysms</td>
<td>40</td>
</tr>
<tr>
<td>Figure-15</td>
<td>Difference in growth rates of abdominal aortic aneurysms between men and women</td>
<td>41</td>
</tr>
<tr>
<td>Figure-16</td>
<td>Growth rates of AAAs in men and women according to initial AP diameter</td>
<td>41</td>
</tr>
<tr>
<td>Figure-17</td>
<td>Annual rate of hospital admissions for abdominal aortic aneurysm in Scotland</td>
<td>45</td>
</tr>
<tr>
<td>Figure-18</td>
<td>Annual rates of admission to hospital with a primary diagnosis of AAA</td>
<td>46</td>
</tr>
<tr>
<td>Figure-19</td>
<td>Figure-19: Annual rates of admission to hospital with a diagnosis of Ruptured AAA</td>
<td>46</td>
</tr>
<tr>
<td>Figure-20</td>
<td>AAA repair in Scotland by modality (open versus endovascular repair)</td>
<td>47</td>
</tr>
<tr>
<td>Figure-21</td>
<td>An overview of the steps involved in data mining (KDD) process</td>
<td>52</td>
</tr>
<tr>
<td>Figure-22</td>
<td>The distribution of duplex related flow abnormalities along the length of the graft</td>
<td>55</td>
</tr>
<tr>
<td>Figure-23</td>
<td>Primary assisted patency over time in grafts enrolled into the vein graft surveillance program</td>
<td>57</td>
</tr>
<tr>
<td>Figure-24</td>
<td>Limb salvage over time in grafts enrolled into the vein graft surveillance program</td>
<td>57</td>
</tr>
<tr>
<td>Figure-25</td>
<td>Distribution of minimum preoperative VGD</td>
<td>58</td>
</tr>
<tr>
<td>Figure-26</td>
<td>Grafts requiring re-intervention stratified by the minimum preoperative internal diameter</td>
<td>59</td>
</tr>
<tr>
<td>Figure-27</td>
<td>Primary assisted patency of vein grafts stratified by preoperative # diameter the graft</td>
<td>60</td>
</tr>
<tr>
<td>Figure-28</td>
<td>Amputation rates stratified by diameter of graft</td>
<td>61</td>
</tr>
</tbody>
</table>
Figure-29: Tight Proximal graft stenosis before and after angiographic treatment  Page-64
Figure-30: Decision tree based on patient profile and first postoperative graft surveillance assessments  Page-66
Figure-31: Primary patency rate in the grafts that have been selected to be high-risk versus low-risk  Page-68
Figure-32: Amputation rate in the grafts that have been selected to be high-risk versus low-risk  Page-68
Figure-33 Decision tree based on patient profile and first postoperative graft surveillance assessments  Page-69
Figure-34 Kaplan Meier survival curve comparing primary assisted patency high vs low risk grafts  Page-70
Figure-35: Kaplan Meier survival curve comparing amputation rates high vs low risk grafts (validation)  Page-70
Figure-36: Angiograms of below knee popliteal bypass graft with stenosis treated with angioplasty  Page-73
Figure-37: Angiograms of a popliteal artery bypass graft with stenosis treated with angioplasty  Page-73
Figure-38 Kaplan-Meier survival curves of Primary assisted patency  Page-74
Figure-39 Kaplan-Meier survival curves of amputation rates  Page-76
Figure-40: Intra-arterial Digital Subtraction angiography of Left Carotid bifurcation  Page-77
Figure-41: Graphical representation of the Multi-Layered Perceptron (MLP)  Page-78
Figure-42: The training curve of the ANN used for banding the degree of ICA stenosis  Page-81
Figure-43: Angiographic degree of ICA stenosis (x axis) vs ANN prediction of ICA stenosis  Page-82
Figure-44: Bland-Altman analysis of comparison of degree of ICA stenosis by ANN and the angiography  Page-82
Figure-45: Kaplan–Meier survival curves based on the UICC classification  Page-87
Figure-46: Accuracy of the ANN survival at 1-year and 3-year vs. the number of input variables used  Page-91
Figure-47: Estimated time for which there is a 10% probability of reaching the threshold diameter  Page-94
Figure-48: Absolute risk reduction afforded by carotid endarterectomy  Page-106
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>Abdominal Aortic Aneurysm</td>
</tr>
<tr>
<td>ABPI</td>
<td>Ankle Brachial Pressure Index</td>
</tr>
<tr>
<td>ACAS</td>
<td>Asymptomatic Carotid Atherosclerosis Study</td>
</tr>
<tr>
<td>ACST</td>
<td>Asymptomatic Carotid Stenosis Trial</td>
</tr>
<tr>
<td>ADAM trial</td>
<td>Aneurysm Detection and Management trial</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AP</td>
<td>Antero-Posterior</td>
</tr>
<tr>
<td>CART</td>
<td>Classification and Regression Tree</td>
</tr>
<tr>
<td>CCA</td>
<td>Common Carotid Artery</td>
</tr>
<tr>
<td>CFR</td>
<td>Case Fatality Rate</td>
</tr>
<tr>
<td>CKD</td>
<td>Chronic kidney disease</td>
</tr>
<tr>
<td>COPD</td>
<td>Chronic Obstructive Pulmonary Disease</td>
</tr>
<tr>
<td>CRISP-DM</td>
<td>Cross Industry Standard Process for Data Mining CRISP-DM</td>
</tr>
<tr>
<td>CT</td>
<td>Computerised Tomography</td>
</tr>
<tr>
<td>ECST</td>
<td>European Carotid</td>
</tr>
<tr>
<td>EDV</td>
<td>End Diastolic Velocity</td>
</tr>
<tr>
<td>ETF</td>
<td>Exchange Traded Fund</td>
</tr>
<tr>
<td>EVAR</td>
<td>Endovascular Aneurysm Repair</td>
</tr>
<tr>
<td>HES</td>
<td>Hospital Episode Statistics</td>
</tr>
<tr>
<td>ICA</td>
<td>Internal Carotid Artery</td>
</tr>
<tr>
<td>ICD-9</td>
<td>International Classification of Diseases ninth revision</td>
</tr>
<tr>
<td>ISD</td>
<td>Information and Statistics Division (NHS Scotland)</td>
</tr>
<tr>
<td>KDD</td>
<td>Knowledge Discovery and Data Mining</td>
</tr>
<tr>
<td>KDOQI</td>
<td>Kidney Dialysis Outcomes Quality Initiative</td>
</tr>
<tr>
<td>LSA</td>
<td>Lothian Surgical Audit</td>
</tr>
<tr>
<td>MCP</td>
<td>McCulloch and Pitts model</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi Layered Perceptron</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MR Angiography</td>
<td>Magnetic Resonance Angiography</td>
</tr>
<tr>
<td>NASCET</td>
<td>North American Symptomatic Carotid Endarterectomy Trial</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>PACS</td>
<td>Picture Archiving and Classification Systems</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>PSV</td>
<td>Peak Systolic Velocity</td>
</tr>
<tr>
<td>PTA</td>
<td>Percutaneous Transluminal Angioplasty</td>
</tr>
<tr>
<td>SEMA</td>
<td>Sample, Explore, Modify, Model and Assess</td>
</tr>
<tr>
<td>SMR1</td>
<td>Scottish Morbidity Record (forms)</td>
</tr>
<tr>
<td>TNM</td>
<td>Tumour Node Metastasis</td>
</tr>
<tr>
<td>UICC</td>
<td>Union International Contra Cancrum</td>
</tr>
<tr>
<td>UKSAT</td>
<td>UK Small Aneurysm Trial</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
<tr>
<td>VGD</td>
<td>Vein Graft Diameter</td>
</tr>
<tr>
<td>VGST</td>
<td>Vein Graft Surveillance trial</td>
</tr>
<tr>
<td>VSQIP</td>
<td>Vascular Surgery Quality Improvement Program</td>
</tr>
</tbody>
</table>
Chapter-1 Background

Introduction

Data mining is a search for valuable information in large volumes of data.\(^1\) Data mining techniques have contributed to every aspect of human endeavour including healthcare.\(^2\)\(^-\)\(^5\) In the twenty first century, health care practitioners require sophisticated decision support systems to support their clinical decisions and practice evidence based medicine.\(^6\)\(^-\)\(^8\)

Avedis Donabedian who is considered the father of healthcare improvement defined delivery of healthcare in terms of the triad of structures, processes and outcomes.\(^9\) Structure includes the material and human resources as well as the organizational structure within which care is delivered.\(^9\) Process denotes the steps involved in providing care and outcomes are effects of care on the health status of patients and populations.\(^10\) Care pathways involved in the delivery of care to patients with lower limb ischaemia, stroke or abdominal aortic aneurysm falls within the definition of the processes which are followed to deliver, safe and effective treatment to this cohort of patients. Care pathways provide a portfolio approach for delivery of care which require exactly the same steps to be followed in management of each of patient and have been credited with improving patient safety in delivery of healthcare.\(^9\)\(^,\)\(^10\) This has only become possible as the progressive improvements in the quality of preoperative investigations have provided the evidence based criteria which clinicians can work with (such as the Antero-Posterior diameter of Aorta in case of AAA, the degree of ICA stenosis or accurate staging of cancer) and randomized controlled trials which have created the evidence base for treatments. Individualizing patient risk in order to inform the clinicians and patients as to the relative importance and necessity of each step in
the care pathway is a compelling argument\textsuperscript{10}, especially with regards to identifying conditions which are likely to remain asymptomatic until they cause a seminal event, such as stroke, vein graft stenosis, or ruptured abdominal aortic aneurysm.

Obtaining and maintaining large databases which often include patient sensitive information and making the data available for legitimate data mining practices in one hand and complying with strict information governance practices in the other is a challenging endeavour with many pitfalls having been identified over the last 20 years. The adoption of clinical governance in the NHS has mandated the development of appropriate and reliable clinical data-sets for use in comparative audit; such data needs to be analysable in a meaningful way.\textsuperscript{11} Real-life data rarely if ever complies with the requirements of statistical analysis or data mining tools. It is may be inconsistent, contain redundant attributes or missing values. It may be in an unsuitable format for immediate analysis.\textsuperscript{11-14} Two types of databases are available in clinical practice.\textsuperscript{15-17} The first is the dataset acquired for research or as part of a registry of patients with a disease process. One example of this is the national vascular registry.\textsuperscript{16}

The other types of databases are hospital information systems or national databases. These information registries collect data for clinical use but without a specific research purpose. They too contain a wealth of information which is accurate and is collected in a systematic way. These types of information systems are often very complex with large numbers of datasets, and many attributes for each record. The examples of these databases include hospital laboratory systems, electronic patient records, hospital Picture Archiving and
Classification Systems (PACS). Other examples include the electronic health records of the patients registered with a general practitioner (Family Physician).¹⁸ These information systems collect accurate data for clinical use. This data is collected in various formats which is optimized for use by clinicians including free texts, codes and images and require significant pre-processing before it can be used for developing decision support systems.

Since the 1960s all NHS hospitals have been obliged to return standard forms containing patients' demographics, co-morbidities, surgical and other interventional procedures, and discharge diagnoses which are recorded in coded format, using internationally recognised International Classification of Diseases (ICD) system. The funding for these episodes of care is contingent upon the return of this data which is recorded in electronic format. Historically accuracy of this data related to the quality of coding in each hospital and tended to contain missing values, with change in legislation, introduction of healthcare quality of care benchmarking and payment by result funding methods, accuracy of national healthcare databases has improved significantly. Still some of the data collected may be imbalanced with regards to the class label of interest. An example of this would be a patient who suffers from heart disease but due to a smoking habit also suffers from chronic lung disease. The actual diagnosis resulting in hospital admission or cause of death would be diagnostic challenge in the individual case.¹⁹ Scrutiny of a national information system which contains a large cohort of similar patients creates a significant data imbalance and could affect the accuracy of the decision support system created from that data source.²⁰,²¹
I was introduced to data mining in healthcare over 20 years ago as a medical student when as a summer project I was involved in developing relational databases for day to day clinical use in patients suffering from sarcoidosis and cystic fibrosis. Subsequent to this I trained as a general and vascular surgeon in the Republic of Ireland, Edinburgh (Scotland) and Sydney, Australia. Parallel to my training in surgery I completed a diploma in medical informatics in the Royal College of Surgeons of Edinburgh. I was introduced to the modern concepts and processes of Knowledge Discovery and Data mining in healthcare (KDDM). Putting these concepts and processes into practice I developed a research portfolio which includes risk assessment models, management of acute pancreatitis using artificial neural networks and genetic algorithms and obtained a master’s degree in medical informatics from university of Bath.22-37 As my career branched out towards vascular surgery my publications also became specialized in that field. This narrative introduces the readers to data mining practices and provides a narrative on nine pertinent publications each of which highlight a data mining paradigm and how the author’s use of data mining has improved the knowledge or informed the debate in the field of vascular surgery. For the last eight years I have been practicing as a consultant vascular surgeon at James Cook University Hospital in Middlesbrough, United Kingdom. Vascular surgery is an increasingly data rich environment and is highly dependent on objective assessment of a number of imaging modalities including duplex ultrasound and is therefore a fertile area for the application of KDDM.
The Mission of Data mining: Data mining Paradigms

Data mining is achieved to fulfil one or 5 different clinical paradigms:38

- Association
- Sequence or Path analysis
- Classification
- Clustering
- Forecasting

These are the clinical missions of a data mining exercises. Depending on the data quality, available tools and the needs of the exercise these may be simple and austere or audacious and complex.

Association

The data mining paradigm that most clinicians are familiar with is studying the association between two or more variables unique to a group of patients. These associations can be examined only if accurate data pertinent to a clinical condition is available. Such data can be collected prospectively, or as is more and more the case be available from local, regional or national databases. The convention is to use statistical models which are appropriate to the clinical setting to disprove the ‘null hypothesis’ defined as the probability of any association between the two variables to be a chance event. More complicated statistical models such as logistic regression analysis can relate the relationship between two variables to other dependent variables. This paradigm is the most commonly used for data mining in vascular surgery, examples of which are in appendices-2 to 6.
Sequence or path analysis

The aim of data mining could be studying a chain of causality or effect of an intervention in changing natural history of a disease process. Sequence or path analysis studies a process and is a higher order of evidence than a simple association between two factors. These relationships may be intuitive such as a relationship between smoking and lung cancer, still as ill health which is related to behavioural and environmental factors happens over many years or indeed decades of exposure, strong evidence of the association only provides circumstantial evidence of causality. True chain of causality would involve evidence of a temporal association between a putative causative factor with the pre-disease states such as pre-cancerous changes and identification of the processes involved in turning healthy tissue to cancerous tissue. In the next chapter we examine how a path analysis can identify factors which are responsible for infra-inguinal vein graft failure and develop and implement a decision tree to aid with graft surveillance and identify a safe treatment modality.

Sequence or path analysis is used extensively in the study of a consecutive chain of activities that cohorts of internet users perform, improving the rates of return on stock market investments and in clinical psychology however its use for development of medical or surgical decision support aids is novel. Appendices 4 to 8 constitute a path analytical approach which involve a number of publications dealing with the same problem each attempting to answer a small but specific part of a question and together work towards building a comprehensive big picture which addresses issues regarding the full cycle of care (in this case identification and treatment of infrainguinal vein graft stenosis).
Clustering and classification are the other paradigms which utilize data mining. Clustering involves organizing a group of objects in such a way that the constituents of each group have more in common than those in other groups.\textsuperscript{45} Classification on the other hand is the simpler task of identifying which set of categories the observation belongs to.\textsuperscript{45,46} Classification is achieved through implementation of a ‘classifier’ which is a mathematical function or an algorithm that relates the value of the input variables to a particular outcome. The input variables can be categorical, ordinal, or linear and the classifier may be deterministic i.e. use a conventional statistical function or an intelligent classifier which identifies patterns within the data which are missed by conventional statistical models.\textsuperscript{46-52}

Classification in its simplest form is a two-step process. The first step involves developing training the data mining model to identify patterns within the data which can then be utilized to develop the classification rules required to make accurate prediction of the outcome variable being studied. This is predominantly by exposing the model to data datasets for which the classifier label “outcome” is known\textsuperscript{47,48}. The second step is evaluating the ability of the classification model in assigning the appropriate label to the unlabelled datasets or to “predict outcome”.\textsuperscript{47,48} By comparing the predicted and actual outcomes accuracy of the predictive model is assessed. Clustering on the other hand is not exposed to a training process ordering of patterns into groups or clusters of data\textsuperscript{50}. Cluster analysis is often not a single algorithm but a general task that needs resolution, data clusters are densities of data points within the dataspace, or a distribution. Clustering models could study the connectivity between data-points for example be hierarchical in nature (Taxonomy), represent a single data vector (for example K-mean distribution which relates
the data points to a central vector which may or may not be a member of that cluster),
distribution models (for example normal distribution), or identify densely packed areas in
the dataspace\textsuperscript{50} (graphical representation of data). Clustering process can be soft (each
object belongs to a certain cluster to a certain degree such as fuzzy logic algorithms) or hard
data points belong to a data cluster or not.\textsuperscript{51} Appendices 7 and 8 are examples of a
classification problems solved through the use of a classification and regression tree.

Forecasting outcomes is intuitively one of the most useful applications of data mining. Most
of the other data mining paradigms such as classification, sequence or path analysis or
association are indirectly linked to forecasting the likelihood of a particular outcome\textsuperscript{52}. This
may be identifying the risk associated with an adverse outcome such as cancer recurrence,
death or failure of a treatment. Enumeration of the level of risk allows a clinician to focus
risk mitigation strategies to cases (datasets) with the highest level of risk\textsuperscript{51-54}. Whilst the holy
grail of data mining is identification of the exact level of risk, for practical applications
stratification of cases to high or low risk or identifying the association between a key
variable and an outcome are more applicable to day to day clinical use\textsuperscript{46}. This is because a
good outcome is the goal of all forms of treatment and in many cases a bad outcome is a
self-fulfilling prophecy, i.e. a cancer which has a high risk of recurrence with treatment will
be associated with a shortened life expectancy if it is deemed untreatable by a clinician who
then decides not to treat the condition based on the prediction. Appendix 10 is an example
of a prediction paradigm where the outcome of surgical treatment for oesophageal cancer is
predicted using a multi-layered perceptron artificial neural network.
Data mining Tools:

Data mining tools are models used in order to implement data mining paradigms. They are broadly divided into deterministic and intelligent data analysis models. Deterministic models identify variables which impact the eventual output variable independently of each other and are incorporated into the model which are often clinical scoring systems. Deterministic models extract pertinent features of the data using conventional statistical models such as regression analysis. The independent predictors of an outcome or a diagnosis can then be used to construct an overall score which may be weighted. If the value of the overall score crosses a threshold value the outcome, diagnosis or disease severity is likely. Examples of deterministic classification models include; Glasgow severity scores for pancreatitis, P-POSSUM perioperative risk prediction model or the Oxford Stroke calculator.

Intelligent models combine statistical analysis and machine learning to extract hidden patterns and relationships which are not recognizable by conventional statistical analysis. Intelligent systems include Bayesian models, Decision trees, Fuzzy logic algorithms, artificial neural networks and genetic algorithms. These models simulate a trained observer by recognizing recurring patterns within the datasets which are not recognizable through conventional statistics. In this narrative decision tree analysis and multi layered perceptron artificial neural network models are used and therefore these data mining models are discussed in some detail preferentially. This does not minimise the value of other intelligent analytical systems such as Bayesian models and fuzzy logic algorithms.
The success of a data mining model at fulfilling the paradigm assigned to it is a very important attribute. If the mission is the study of an association between two variables then disproving a null hypothesis may be sufficient. If the mission is classification or prediction, assessment of sensitivity, specificity and positive and negative predictive values are used. Study of entropy (heterogeneity) is one way of assessing the usefulness of data mining model. A population of subjects with a binary attribute where 60% of subjects are=0 and 40% are=1. Entropy of this system without classification would be calculated by the equation:

$$E(s) = \sum_{i=1}^{c} -P_i \log_2 P_i \text{ or Entropy (0.4, 0.6) which equals 0.97}$$

A classifier which identifies subjects 0 and has 80% sensitivity and 80% specificity will divide them as follows:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>Heterogeneity</th>
<th>entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identified as 0</td>
<td>48</td>
<td>8</td>
<td>56</td>
<td>Entropy (0.85, 0.15)</td>
</tr>
<tr>
<td>Identified as 1</td>
<td>12</td>
<td>32</td>
<td>44</td>
<td>Entropy (0.27, 0.73)</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>40</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Entropy of the system using a classifier would be calculated using this equation:

$$E (\text{Classifier}) = \sum_{i=1}^{c} P(c)E(c) \text{ or } E (\text{Classifier}) = (0.19)+(0.41) \text{ or 0.60}$$

The net information gain using this classifier is 0.37. The closer to 0 the entropy value is, the more homogeneous the classified population are conversely; the closer to one the value is the higher the heterogeneity (uncertainty of the observation) will be. The net information gain is the value often used by intelligent classification models in selecting the pertinent variables, functions or topography to develop the most informationally efficient models.
Decision tree Analysis

Decision tree analysis is a decision support model which has been developed from the decision theory by the distinguished American statistician Leo Breiman. Decision theory is the study of optimal actions as determined by considering the probability and utility of different outcomes. A decision tree uses a tree like graph model to map observations about a target item to conclusions about the value of the target item (Figure-1). Where the target value has a finite set of values (ordinal or nominal) the process is known as a classification tree. If the Target value is a continuous variable it is called a regression tree. The umbrella term Classification and Regression Tree (CART) is used to group both processes. Once created and validates, these algorithms are simple to understand, requires little preparation, can use categorical as well as continuous variables and create a white box model which is transparent to the user and it is possible to validate using statistical analysis such as receiver operator characteristic test. CART models will be discussed in the context of infra-inguinal vein graft surveillance (Appendices 7and 8).

Figure-1: Graphic representation of a decision tree Precision tree™ Palisade corporation.
Artificial Neural Networks

Artificial neural networks are a family information processing algorithms, which are modelled on the way biological nervous systems process information. Artificial neural networks like their biological counterparts are composed of many highly interconnected processing elements (neurones) working in unison to solve specific problems. Artificial neural networks have an ability to drive meaning from imprecise complicated or incomplete data. In addition, ANNs exhibit the following features:

- Adaptive learning
- Self-organization
- Real time operation
- Fault tolerance (redundant information coding)

The neuron (node) is the basic computational unit of an artificial neural network. It receives a variety of inputs from other neurons through connections that resemble synaptic structures and have a binary output. This output is determined by the sum of the inputs as well as the synaptic strength (weight) attached to each input variable. A wide variety of neural network designs with varying degrees of complexity have been described. The simplest form of neural network is the multi-layered perceptron (MLP) Figure-2. This network topography was the one used in this narrative. This is made up of layers interconnected of processing units with some pre-processing of input variables. It often has a back propagation circuit which allows training of the ANN. Each of the input nodes in the first layer corresponds to a single input variable. Initial values are set for the ‘weights’ associated with each link in the network. Input data for which an output is known are
presented to the network. If the predicted output from the model does not equal the known output, the ‘weights’ within the network are changed so as to narrow this difference. This process continues until the prediction errors are minimized. In this way a trained neural network contains knowledge which is stored in the form of weighted connections between the neurones in the hidden layers.\textsuperscript{63} Once the network is trained and validated, it may be used on unseen data for prediction or classification purposes.\textsuperscript{63} In general, neural network analysis is more successful than traditional statistical techniques when the importance of prognostic variables is expressed as a complex unknown function of the value of the variable, when the prognostic impact of a variable is influenced by other prognostic variables, or when the prognostic impact of a variable varies over time. These conditions are found in complex biological systems. For the purposes of analysis, each variable used for outcome prediction can be regarded as a single dimension in a multidimensional space. Traditional deterministic statistical techniques are particularly suited to the analysis of data with a low dimensional complexity, small number of variables and linear separation between different classes. Artificial neural networks have been found to be significantly more accurate at predicting outcome than deterministic classification models in patients with cancer\textsuperscript{36,64}, critically ill patients\textsuperscript{65}, in the diagnosis of myocardial infarction\textsuperscript{65} and assessment of severity of acute pancreatitis.\textsuperscript{29}

The last 2 publications in this narrative discuss the use of MLPs artificial neural networks in measuring the degree of internal carotid artery stenosis (Appendix-9) and predicting the success of surgical treatment for oesophageal carcinoma (Appendix-10).
Figure-2: A schematic of a simple artificial neural network
Data mining and Decision Support in Healthcare

Evidence based practice has gained broad based support in healthcare. This has been driven though structural changes in delivery of healthcare, development of standardized evidence based processes called care pathways which have driven quality improvement in different fields of medicine. Establishment of the National Institute for Health and Care Excellence NICE and its Scottish counterpart SIGN which have placed evidence based guidelines for management of a diverse range of medical conditions within easy access of clinicians.

Clinical decision support systems are electronic systems which are used to facilitate clinical decision making process using patient data, models, and structured decision processes. Decision support encompasses a growing group of disciplines including, psychology, operations research, data warehousing The trends in clinical decision support follow those in business fields, therefore major future contributions to clinical decision support are likely to come from data warehouses, integration with data mining, developments in qualitative modelling and “soft “computing. Examples of such a data warehouse will be demonstrated in the publications discussed in later chapters, specifically in the context of infra-inguinal vein bypass grafts (Appendix 7 and 8) and internal carotid artery stenosis and the risk of stroke (Appendix-9).
There are two commonly used standardized processes for data mining:

- Cross Industry Standard Process for Data mining CRISP-DM\textsuperscript{72}
- Sample, Explore, Modify, Model and Assess SEMMA\textsuperscript{73}

CRISP-DM is the predominant process for knowledge discovery. It was devised in the late 1990s as a European Union project led by 5 large corporations, SPSS Inc., Teradata, NCR Corporation, Daimler AG and OHRA. It was released in 1999 and since then it has become "de facto standard for developing data mining and knowledge discovery projects." CRISP-DM divides the process of data mining into 6 different phases which are illustrated in Figure-3.\textsuperscript{72}

The processes involved in CRISP-DM flow form generic tasks to specific ones. Generic tasks are those which are involved in any data mining such as addressing missing data. Specific tasks on the other hand include those processes required to deal with the task at hand and process instances which are specific items performed to achieve process definitions (objectives) defined by data mining project.\textsuperscript{73}

Arguably the best example of CRISP-DM use is reported by Catley et al. They used CRISP-DM to manage a multi-dimensional data stream in a neonatal intensive care. This is a which is a very challenging clinical field with regards to the wealth of data, the time dependency of the data available and the critical nature of the information gained.\textsuperscript{74}
Figure-3 Visual representation of the processes involved in CRISP-DM.\textsuperscript{72}
SEMMA is an acronym that stands for Sample, Explore, Modify, Model and Assess (Figure-4). It was developed by Statistical Analytics Systems (SAS) inc. The SEMMA model goes through a few self-explanatory steps. These steps are illustrated in figure-4. SEMMA is different to CRISP-DM as it focuses on the data mining process leaving business or the clinical context aspects out. For this reason and the fact that SEMMA is a proprietary product we elected to use CRISP-DM in the projects described in this thesis.

---

Figure-4 The steps involved in SEMMA data mining methodology
Data mining processes are usually dependent on user parameters and context; each data mining algorithm has its own theoretical assumptions. End-users may not have sufficient information about the parameter(s) and their selection; therefore, processes such as CRISP-DM engage with business and data for understanding prior to design of database and the need for validation and quality control. The data used for data mining has traditionally needed to be sourced from different sources such as hospital databases which are primarily used for billing purposes do not contain information of sufficient quality to be the sole source for data mining and are exposed to potential threats such as data imbalance and missing data. Data mining packages require extensive knowledge and expertise, which have meant that data mining has remained the domain of enthusiasts and experts. This narrative explores how these barriers can be overcome so that data mining can become the core technology for the practice of evidence based vascular surgery.
The Clinical Context for Included Publications

Abdominal Aortic Aneurysms

Over 7000 people die from ruptured AAA in the United Kingdom annually.\textsuperscript{76} The best treatment for AAA is elective repair of pre-symptomatic abdominal aortic aneurysms. Such a therapeutic strategy depends on effective identification of patients with AAA and the subgroup of patients in whom there is a real risk of aneurysm rupture. Most patients with AAAs are asymptomatic, timely identification of AAA may be achieved through targeted screening of the at-risk populations. Small abdominal aortic aneurysms can safely be followed by ultrasound based surveillance.\textsuperscript{77-82} Due to several large-scale randomised controlled trials a lot is known about AAAs and the timing of treatment, duration of follow up and treatment modalities employed.

There is a targeted national screening program for identification of patients with AAAs in the United Kingdom. This program is aimed at men above the age of 65 due to the higher incidence of AAA in this cohort of patients, consequently less is known about natural history of AAAs in women. Health care processes have the potential to generate vast quantities of data, much of which has been collected in clinical databases. Over the last two decades these databases have accumulated large quantities of information much of which is hidden in the relationships and patterns within the data.\textsuperscript{83} The two publications discussed are examples of how large clinical and administrative databases can be successfully mined to identify clinically significant associations using structured data mining methodology (Figure-5)\textsuperscript{84} whilst dealing with issue such as, noise, missing and incomplete datasets, class imbalance and validity of class labels.
Observational studies suggest that 25–41 per cent of AAAs larger than 5 cm will rupture within 5-years if they are not repaired. In the United Kingdom management of AAAs below this size (5·5 cm in maximum anteroposterior (AP) diameter) is guided by the UK Small Aneurysm trials and the US veteran affairs Aneurysm Detection and Management (ADAM) trials, which have revealed that early elective surgery conferred no long-term survival benefit over regular surveillance. Following the advent of Endovascular Aneurysm Repair (EVAR), further randomised controlled trials in Europe and the United States (ACE, PIVOTAL and have revealed that patients with AAAs which are less than 5cms do not benefit from endovascular repair compared with ultrasound based follow-up.

The UK Small Aneurysm Trial found that the risk of AAA rupture was independently associated with aneurysm size, current smoking, mean blood pressure and female sex. Growth rate is one of the factors used to identify patients who are at potential risk of aneurysm rupture. It is broadly accepted that growth rate of greater than 1 cm in one year indicates a risk of rupture and is an indication for AAA repair regardless of the size of AAA.
Figure-6 Serial Ultrasound examinations of the same asymptomatic AAA performed 5 years apart revealing the gradual increase in the maximum AP diameter from 4.44 cms to 5.98 cms (leading to its eventual repair).
The mean growth rate of AAA is variable depends on the population studied. Reported expansion rates vary widely, being up to four times higher in referral-based studies compared with community-based screening programmes which detect smaller AAAs. The main determinant of growth rate of an AAA is its initial maximum anteroposterior (AP) diameter. Other factors such as age, cigarette smoking and severe cardiac disease have also been associated with rapid expansion of AAA. Female sex may be an independent risk factor for rapid growth of AAA. A large population based study from Sweden suggests that a large number of female subjects with ectatic aorta (46%) go on to develop an AAA within 5 years of initial US examination. A recent publication by Sidloff et al has revealed that women continue to have a higher mortality from AAA repair despite the introduction of endovascular techniques and the institution of Vascular Quality Improvement Program.

The introduction of endovascular aneurysm repair (Figure-7) holds the promise of reducing the peri-procedural morbidity and mortality of treatment. In 2007-8 the impact of these advances on national practices and utilization of services were yet to be analysed in Scotland. Furthermore, no information was available at that time on whether heightened awareness of AAA in primary care has resulted in the desired outcome of reducing aneurysm rupture rates. This study is noteworthy as it analyses outcomes following treatment for AAA, just prior to the development and implementation of Vascular Surgery Quality Improvement Program (VSQIP) and nationwide introduction of population screening for AAA, which were measures developed to address deficiencies in care of patients with AAA. Comparing the results of Mofidi et al (Appendix-3) and Sidloff et al is sobering and suggests that the task of reducing mortality from AAA in women is far from accomplished.
Figure-7: Preoperative CT scan of a large intact AAA (Above). Completion angiogram of the same AAA following treatment with an endovascular stent EVAR (Below).

Infra-inguinal vein graft surveillance
For over 70 years infra-inguinal vein graft bypass surgery has remained the gold standard treatment for management of disabling intermittent claudication and limb threatening ischaemia\textsuperscript{94-95}. Vein grafts are prone to develop stenoses, which may precipitate their failure\textsuperscript{96-102}. Stenosis may develop from technical error, pre-existing vein abnormality or myo-intimal hyperplasia. Over time, these grafts are at risk from progression of atherosclerotic disease which can result in graft failure by reducing flow across it (Figure-8).

Evidence of these problems is recognisable by duplex ultrasound scanning, a technique acknowledged for its accuracy in identifying and grading stenotic lesions that threaten graft patency\textsuperscript{101}. Duplex scanning has been widely used for graft surveillance, the rationale being that correction of stenotic lesions is likely to improve graft patency and limb salvage rates (Figure-9)\textsuperscript{102-106}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{Aetiology of vein graft stenosis is classified chronologically.}
\end{figure}
The wisdom of duplex vein graft surveillance has been questioned by the Vein Graft Surveillance (VGST) trial. It revealed no significant difference in graft patency or limb salvage in the medium term between patients who were followed up clinically and those who underwent vein graft surveillance. The finding of VGST threatened vein graft programs which were in place in the most vascular surgery departments in the developed world. Although there were some flaws in the design of the VGST, due to its size, the fact that it was a randomised controlled trial its and multi-centre nature, it was unlikely to be repeated to account for these methodological issues.

Several preoperative factors have been found to influence the outcome of infra-inguinal vein bypass. While many of these factors are endogenous to the patients and beyond the direct control of the surgeon, the size and quality of the venous conduit available for bypass is considered important factors affecting the outcome of infra-inguinal bypass. Although the use of smaller veins does not invariably result in graft failure, grafts with a diameter <3.5 mm are at increased risk of failure. Under these circumstances, the use of composite vein, arm vein or prosthetic grafts rather than uninterrupted small-calibre long saphenous veins are advocated all of which add to the complexity of the operation and are by themselves risk factors for graft failure.

Operative techniques such as patch angioplasty, interposition graft or replacement graft remain the gold standard for the management of vein graft stenosis. These techniques require the availability of an additional autogenous vein and carry the risk of morbidity and mortality associated with re-operation. In addition, jump and interposition grafts used to treat vein graft stenosis may by themselves be at risk of development of vein graft
stenosis\textsuperscript{114-116} The application of angioplasty in the treatment of stenoses in other venous conduits such as haemodialysis arterio-venous fistulae\textsuperscript{117-119} had led to the introduction of this technique for the management of threatened bypass grafts.\textsuperscript{120-124} Conventional wisdom indicates that a segment of vein which has been afflicted by neointimal hyperplasia causing a tight stenosis is no longer a suitable conduit for infra-inguinal bypass and needs replacing, on the other hand neointimal hyperplasia is often very focal in nature and represents the ideal lesion for endovascular intervention. The efficacy of this treatment modality had been questioned and at the time of this publication no consensus existed about its primary use in the management of vein graft stenosis.\textsuperscript{123-127}

The four studies in this narrative utilized data mining methodology to examine vein graft surveillance programs in the east of Scotland Vascular network. These studies have helped to inform the debate regarding vein graft surveillance, specifically with regards to identification of grafts which are at high risk of failure. In addition, one of the publications examined the efficacy of endovascular techniques for treatment of vein graft stenosis.
Figure-9: Tight Proximal graft stenosis identified on Duplex US (Above) angiographic finding of a stenotic proximal vein graft stenosis in a different graft (Below).
Assessment of Degree of Internal carotid artery stenosis

Randomised controlled trials have confirmed the value of carotid endarterectomy over best medical treatments for patients who have significant stenosis in the extracranial carotid artery. This benefit is most pronounced for patients who have suffered from recent ipsilateral transient ischaemic attacks, minor stroke or ischaemic ocular symptoms attributable to the stenosed artery. Together these studies have randomised 5893 patients with over 33,000 years of follow up. A meta-analysis of the individual patients from these studies (NASCET, ECST, VA309) revealed that the degree of ICA stenosis measured using digital subtraction angiography was a key determinant of risk of further neurological events and by extension the benefit each patient may obtain from timely carotid endarterectomy. It revealed that surgery was harmful in patients with less than 30% stenosis, of no benefit in those with 30–49% stenosis, of some benefit for 50–69% stenosis, and highly beneficial for those with 70% or more stenosis without near-occlusion. Data from these studies has been used to develop risk calculators (Figure-10) which can be used to individualise risk of developing further neurological events. A major component of this risk calculator is the degree of ICA stenosis. NICE guidelines suggest that these risk calculators are incorporated into every day clinical decision making. In effect 25 years after publication of NASCET and ECST accurate assessment of degree of ICA stenosis is as relevant as ever.
All pertinent randomised controlled trials had utilized digital subtraction angiography to measure degree of ICA stenosis (or confirm presence of ICA occlusion). Intra-arterial angiography has a significant morbidity and a low but significant mortality rate.\textsuperscript{134-136} As technology has developed and experience has broadened duplex ultrasound scanning has been found to provide sufficient information for clinical decision making in patients with ICA stenosis.\textsuperscript{137,138} Assessment of degree of ICA stenosis involves combining the information obtained from Peak Systolic Velocity (PSV), End Diastolic Velocity (EDV) at the site of stenosis, PSV and EDV at the common carotid artery (CCA) or distal ICA, the ratios between ICA and CCA velocities, subjective assessment of the wave form as well as the β-mode US appearance of stenotic lesions\textsuperscript{139-142}(Figure-11). Summing up all this information requires
significant degree of observer interpretation\textsuperscript{143} making carotid duplex examination prone to inter-observer variability and potential of introducing error.\textsuperscript{144}

Figure-11: Carotid Duplex Ultrasound exam of right carotid bifurcation (left) together with MR angiogram image (right) both showing a tight ICA stenosis.
Cut off values for these measurements are selected using receiver operator characteristic (ROC) curves are used to define a certain degree of stenosis.\textsuperscript{139-146} The combination of these measurements is based on empirical selection with not enough statistical data to support which values best represent the degree of ICA stenosis particularly in presence of confounding variables such as contralateral ICA occlusion.\textsuperscript{139-146} These criteria vary between laboratories both in the method of calculation and in the range of degrees of stenosis used to describe each stratification or band.\textsuperscript{139,140} Traditionally banding of the degree of ICA stenosis using Duplex criteria usually involves 20-percent bands. These broad bands do not allow adequate classification of ICA stenosis to provide all the information needed for clinical decision making in patients with symptomatic carotid stenosis\textsuperscript{139-146} or in patients with asymptomatic stenosis who are offered carotid endarterectomy.\textsuperscript{139-146} A small numbers of criteria exist which stratify the degree of ICA stenosis into 10% bands.\textsuperscript{140} The vascular society of the United Kingdom and Ireland and the Society for Vascular Technologists in the United Kingdom have put forward comprehensive reporting standards for carotid duplex US examinations which include a standardised criteria.\textsuperscript{140} However there is evidence to suggest that different Duplex criteria are still in use and that the difference between the different methods of interpreting the Doppler waveform generated by the Duplex US examination (Duplex criteria), (Figure-11) has clinical significance.\textsuperscript{139}
Predicting Outcome following surgery with curative intent for treatment of Oesophageal cancer

Union International Contra Cancrum (UICC) TNM classification system is the most widely used method of staging for patients with oesophageal and OG junction carcinoma. It is purely based on anatomical extent of the disease, therefore, its ability as a predictive model is limited to the information obtained from imaging modality, whilst anatomical extent of the disease is of value, a number of other factors, which are amenable to preoperative assessment, are known to influence survival following resection with curative intent. These include tumour length measured on contrast swallow, tumour differentiation and the presence of submucosal lymphocytic infiltration as well as involvement of the resection margins. Addition of these variables to a predictive model for oesophageal carcinoma based on conventional statistical models would significantly increase its complexity and limit its clinical utility. Artificial Neural Networks have been increasingly adapted in medicine as decision support aids and to predict outcome from using complex multivariable prognostic indicators.
Chapter-2  Methodology and Results

This narrative describes the contributions of 9 publications to the fields of general and vascular surgery. They cover a diverse set of clinical problems and provide unique solutions through the use of data mining techniques. Whilst it would not be practical to discuss these in the context of one global methodology there is a lot of commonality between them, not least is the use of CRISP-DM processes to implement them. However, the differences between the publications mandate that they are discussed separately when it is necessary to highlight their unique feature, such as data source or data analysis models. This chapter is divided into 4 different work-streams:

2.1 Study of Associations Using Deterministic Classification

2.2 Sequence (Path) Analysis Using a Combination of Deterministic and Intelligent Methods

2.3 Classification using Intelligent Analysis

2.4 Prediction of Outcome using Intelligent Analysis

Due to the fact, that CRISP-DM is an integrative data mining process (Figure-12); methodology and results are discussed in one unified chapter.
Figure-12: CRISP-DM Process Applied to Study-2
2.1 Study of Associations Using Deterministic Classification

2.1.1 Assessment of Growth rate of AAAs (Paper-1, Appendix 2)

This publication sets out to assess the factors that influence the growth rate of AAAs. Hence the basic data mining paradigm explored was study of associations between putative factors including gender and rate of growth of AAAs. This was achieved through analysis of a database of all patients who have been enrolled and followed up by an aneurysm surveillance clinic for 20 years. The model used for data mining was CRISP_DM model which has already been described in the previous chapter. The methodology of the study will be discussed under the 5 main headings of business understanding, data understanding, data preparation, data analysis and evaluation. The last step of the CRISP-DM which is the deployment of findings of each work-stream is discussed in later chapters.

2.1.1.1 Business Understanding

Between January 1985 and August 2005 all patients who had been diagnosed with an AAA at the Royal Infirmary of Edinburgh and not considered for early aneurysm repair were assessed serially by abdominal ultrasonography. This included all patients with small AAA and patients who either refused or were not considered for elective aneurysm repair at the time of first presentation. The data on each patient attending the aneurysm surveillance clinic was collected prospectively in a Microsoft® Access™ database. The database had been administered by the same person (Dr J Kelman) since its inception. As the administrator of the database she was responsible for integrity and quality of the data collected. The quality of the data had been regularly audited.183
The maximum AP diameter abdominal aortic aneurysm was recorded at enrolment and during each subsequent visit. Patients were followed up until they died; the AAA was repaired or they were discharged from the clinic. All patients were cross referenced against the Lothian Audit database\textsuperscript{166} to see if they had undergone operative interventions pertinent to their AAA.

### 2.1.1.2 Data understanding and preparation

The dataset consisted of 17 variables of which 8 were continuous variables the rest were categorical variables (Table-1)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data Type</th>
<th>% Missing values</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHI number</td>
<td>Continuous</td>
<td>0</td>
</tr>
<tr>
<td>Date of enrolment</td>
<td>Continuous</td>
<td>0</td>
</tr>
<tr>
<td>Date last followed</td>
<td>Continuous</td>
<td>0</td>
</tr>
<tr>
<td>Outcome</td>
<td>Nominal</td>
<td>0.04%</td>
</tr>
<tr>
<td>Date of birth</td>
<td>Continuous</td>
<td>0</td>
</tr>
<tr>
<td>Age</td>
<td>Continuous</td>
<td>0</td>
</tr>
<tr>
<td>Postcode</td>
<td>Nominal</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>Binary</td>
<td>0</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Binary</td>
<td>29%</td>
</tr>
<tr>
<td>COPD</td>
<td>Binary</td>
<td>30.60%</td>
</tr>
<tr>
<td>Ischaemic Heart disease</td>
<td>Binary</td>
<td>29.73%</td>
</tr>
<tr>
<td>Smoking</td>
<td>Binary</td>
<td>15.67%</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Binary</td>
<td>18.30%</td>
</tr>
<tr>
<td>Statin Therapy</td>
<td>Binary</td>
<td>57%</td>
</tr>
<tr>
<td>Initial AAA AP diameter</td>
<td>Continuous</td>
<td>0.00%</td>
</tr>
<tr>
<td>Final AAA AP diameter</td>
<td>Continuous</td>
<td>1.8%</td>
</tr>
<tr>
<td>Growth rate</td>
<td>Continuous</td>
<td>0.63%</td>
</tr>
</tbody>
</table>

Table-1: Variables included in the AAA surveillance clinic dataset

Issues Regarding the Available dataset were addressed under 3 different settings (Missing data, class imbalance and validity of class labels), due to the historical nature of data and
the long duration of follow up, minimal pre-processing of available data was performed and input variables which were not considered to exhibit sufficient validity for inclusion into the analytical model were excluded.

2.1.1.2.1 Missing data: Due to the historical nature of data it was not possible to source missing values from other organisational databases. Variables which had more than 20 percent missing values were excluded from analysis. Although the authors acknowledge that 20% missing data is an arbitrary value.

2.1.1.2.2 Class imbalance: As post-mortem examination concerning the cause of death was unavailable in a significant number of cases, and “it was believed that knowledge of the presence of AAA may have biased the recorded cause of death”, the risk of rupture of AAA was not calculated in this study. Although the authors acknowledge that this may have been extra cautious in retrospect as this data was readily available.

2.1.1.2.3: Validity of class labels: History of hypertension, smoking and renal failure were collected as nominal variables at the time of enrolment into the AAA surveillance clinic. As there had been no attempt at defining the diagnoses of hypertension or renal failure the observations were not thought to be robust enough for inclusion into the model. For example, the history of smoking was collected as a binary variable without collecting the duration or quantity of smoking (e.g. in pack years). The history of smoking after enrolment into AAA surveillance had not been collected either. These observations together with the significant number of missing values (16%) resulted in that value also being excluded from the model. The authors acknowledge that whilst accepting limitations imposed by the
validity of class labels, hypertension and smoking were two variables which could have been included and are putative contributors to growth rates of AAAs.

2.1.1.3 Modelling

The association between the initial diameter of an AAA and its growth rate was assumed to be linear. Therefore the model created was linear regression analysis. The growth rate of AAA was selected to be the dependent variable whilst age, sex and initial aneurysm diameter the independent variables. The use of linear regression analysis to assess expansion of AAAs is likely to be associated with upward bias\textsuperscript{168}, however as this would affect aneurysm growth rates in men and women equally, it is unlikely to influence comparisons based on sex.

2.1.1.4 Findings

Over 20 years, 1255 patients who had been diagnosed with AAA were entered into an ultrasound based surveillance programme (824 men and 431 women). They were followed for a median of 30 months (6–185) or six consecutive ultrasound examinations (range 2–37). Median AP diameter on initial AAA examination was 41 (range 25–83) mm. AAAs were significantly larger at initial presentation in men compared with women (Figure-13, P<0.001).
Figure-13 Distribution of anteroposterior diameter of abdominal aortic aneurysm on initial presentation in men and women.

The median growth rate of AAAs was 2.79 (range: 4.80–37.02) mm per year. There was a significant variation in the growth rate of AAAs. The growth rate of AAAs was higher in women than in men (P<0.01), (Figure-15).

Figure-14 Correlation between expansion rate and initial maximum anteroposterior diameter of the abdominal aortic aneurysm $R^2 = 0.22$, (P < 0.001).
Figure-15: Box and whisker plot illustrating the difference in growth rates of abdominal aortic aneurysms between men and women.

Figure-16: Growth rates of abdominal aortic aneurysms in men and women according to initial anteroposterior diameter.
Weighted linear regression analysis revealed that initial AP diameter and female sex were significant predictors of rapid aneurysm growth (Figure-16), (Table-2). The final disposition of patients was as follows: 322 underwent operative repair. This was performed for symptomatic or ruptured AAA in 43 patients (29 men and 14 women) and elective repair in 279 patients. Three hundred and ninety-one patients died, 195 patients were lost or discharged from follow-up, and the remaining 347 patients remained under surveillance at the time of publication.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds Ratio (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.06 (0.90–1.22)</td>
<td>0.420</td>
</tr>
<tr>
<td>Initial AP Diameter</td>
<td>3.83 (3.12–4.55)</td>
<td>0.001</td>
</tr>
<tr>
<td>Female sex</td>
<td>2.04 (1.68–2.40)</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table-2: Factors associated with growth rate of AAAs.
2.1.2 Outcome from abdominal aortic aneurysms in Scotland (Paper-2, appendix 3)

Understanding trends outcome (survival) following treatment for a disease process and associations between these outcomes and factors unique to patients is a very powerful tool in predicting patient needs and areas of improvement. This is often achieved by scrutinizing national registries and databases. This study is an example of the use of datamining in such an endeavour.

2.1.2.1 Data understanding and preparation

Over the past 54 years, National Health Service (NHS) hospitals in Scotland have been obliged to return a standard form for all patients admitted to hospital. These data are stored electronically by the Information and Statistics Division (ISD) of NHS Scotland. The quality of data collected on Scottish Morbidity Record forms (SMR1) in patients with AAA has been shown to match clinical records in over 99% of cases.\(^{169}\)

Anonymised SMR-1 forms for all patients with the primary diagnosis of AAA from 1/1/1991 to 31/12/2006 were obtained. The diagnosis for each in-hospital event was classified using final hospitalization diagnosis codes of the International Classification of Diseases ninth revision (ICD-9). Each patient was assigned a unique identifier, it was possible to follow him or her through the healthcare system and remove duplicate records.

SMR-1 forms record secondary diagnoses and co-morbidities. However, there has been long lasting concern that unlike the primary diagnosis and treatment there are many instances of data imbalance and missing data which have persisted to date.\(^{170}\) Such concerns would have meant that data regarding secondary diagnoses could not be used in data analysis. Since 2008, SMR-1 forms have included calculations of the Charleston comorbidity scores which is
a standardized method of collecting the risk to patients’ wellbeing from their comorbidities. This was not available for the duration of the study. Data obtained from the ISD were cross-referenced against the Scottish Audit of Surgical Mortality database to assure maximum accuracy. Patients in the database were divided by age, by presentation (ruptured versus intact AAA) and by method of repair (open versus endovascular).

2.1.2.2 Data analysis

Stepwise multivariable logistic regression analysis was carried out with in-hospital mortality as the dependent variable, and age, sex, presentation diagnosis and procedure type as independent variables.

2.1.2.3 Findings:

Over the 16 year duration from January 1991 to December 2006, 12,706 patients were admitted to Scottish hospitals with a principal diagnosis of AAA (9779 men and 2927 women). The median age of these patients was 73 (range 17–97) years, [71 (range 17–96) years for men and 75 (range 35–97) years for women (P<0.001)], 8818 patients presented with intact AAA (69.4 %) whilst 3888 presented with ruptured AAA (30.6%). Figure-17 shows the annual rate of admissions and the proportions of ruptured and intact AAAs over the 16 years examined in this study.

Higher proportion of women presenting with AAA to Scottish hospitals (29.5%) presented with ruptured AAA compared with men (27.5%) presenting with the same primary diagnosis ($\chi^2=4.1$, $P=0.043$). The incidence of ruptured AAA remained constant over the 16 years of
the study, (Figure-18). However, there was a relative reduction in the proportion of patients with a ruptured AAA as the rate of elective repair of AAA increased over this time ($\chi^2=17.1$, $P<0.001$).

Figure-17: Annual rate of hospital admissions for abdominal aortic aneurysm in Scotland from 1992 to 2006.
Figure 18: Annual rates of admission to hospital with a primary diagnosis of Abdominal aortic aneurysm stratified by age.

Figure 19: Annual rates of admission to hospital with a primary diagnosis of Ruptured abdominal aortic aneurysm stratified by age.
A total of 8348 patients underwent surgery for AAA; this included 7466 open repairs (89.4%) and 698 endovascular repairs (8.4%). Figure-20 illustrates the estimated yearly rates of open and endovascular AAA repair.

![Figure-20: AAA repair in Scotland by modality (open versus endovascular repair)](image)

Over the 16-year interval, 2579 patients (20.3%) died within 30 days of admission with an AAA. Some 531 (37.7%) of 1410 patients who underwent attempted repair of a ruptured AAA died, compared with 362 (5.2%) of 6938 patients who had repair of an intact AAA. The 30-day mortality rate was significantly higher following open repair of an intact AAA than after endovascular repair (5.6% (351 patients) versus 1.6% (11 patients); $\chi^2 = 19.8$, $P<0.001$).
Early postoperative mortality rates were significantly higher in women than in men (29.9% (875 women) versus 25.8 per cent (2523 men); χ² = 19.3, P <0.001). Logistic regression analysis revealed that age, sex, admission diagnosis (ruptured versus intact AAA) and type of procedure performed (open versus endovascular repair) were independent predictors of 30-day and in-hospital mortality following AAA repair (Table-3).

<table>
<thead>
<tr>
<th></th>
<th>Odds ratio</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (&lt;75 or ≥75 years)</td>
<td>2.52 (2.36, 2.74)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sex</td>
<td>1.63 (1.48, 1.78)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Admission diagnosis (ruptured versus intact)</td>
<td>10.49 (9.53, 11.54)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Operation (endovascular versus open repair)</td>
<td>0.67 (0.58, 0.76)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Values in parentheses are 95 per cent confidence intervals. *Stepwise multivariable logistic regression analysis.

Table-3 Factors associated with death from AAA
2.2 Sequence (Path) Analysis Using a Combination of Deterministic and Intelligent Methods

Sequence (path) analysis involves studying a clinical problem in great detail. The information collected can then be used to develop a bigger picture which can be used for bench-marking of outcomes or designing decision support systems. It involves study of associations between many pertinent variables (often using deterministic data mining models) but also a summative assessment to develop an overall view which can become a decision support system. The paradigm of path analysis is applied to the difficult problem of infra-inguinal vein graft failure. Data sources are departmental registries of vein graft bypasses in the East of Scotland Vascular Network (Royal Infirmary of Edinburgh and Ninewells hospital, Dundee). These 2 distinct registries were used for follow-up of patients who had undergone vein graft bypass surgery for administrative reasons such as generating and regulating follow up scans and recording the results of each investigation as well as assisting in periodic audit of infra-inguinal bypass procedures. The final decision support tool was validated using outcomes of patients in a different hospital in England (James Cook University Hospital in Middlesbrough). All the 3 databases were developed using Microsoft® Access™. The first 2 databases were administrative databases used for audit of infra-inguinal vein graft surveillance with secondary research purposes. The 3rd database was created specifically to validate the decision tree model which is described in the Paper -7, Appendix-8. As Per Caldicott’s guidelines the dataset collected and stored was limited to patient demographics, atherosclerotic risk factors (Smoking, diabetes, hypertension, hyperlipidaemia, chronic kidney disease), Antiplatelet agent and statin therapy, the type of initial bypass operation, conduit used and follow up information including results of consecutive clinical and US examinations.
2.2.1 Business understanding

Consecutive patients, who underwent infra-inguinal bypass procedure using autologous vein, were enrolled into an ultrasound based graft surveillance program. This involved duplex scanning and the measurement of ankle brachial index pressures during each visit.

Ultrasound Examination: Vein graft surveillance protocol was performed before discharge from the hospital and at 6 weeks, 3 months, 6 months, 1 year, 18 month and 2 years after operation. Ultrasound examination began within the inflow artery, progressing down the graft and continuing into the outflow artery. The inflow and outflow vessels were assessed. Peak systolic velocity was measured at sites of stenosis, at multiple sites within the graft and within the outflow vessel. The velocity ratio at the site of stenotic lesions was calculated. Grafts were categorized into 3 groups based on duplex findings during each graft surveillance visit (Table-4)

<table>
<thead>
<tr>
<th>PSV at the site of stenosis</th>
<th>Post stenotic PSV ratio</th>
<th>Drop in ABPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute value (cm/s)</td>
<td>PSV ratio</td>
<td></td>
</tr>
<tr>
<td>Mild stenosis</td>
<td>&lt;200</td>
<td>&lt;2</td>
</tr>
<tr>
<td>Moderate stenosis</td>
<td>200-300</td>
<td>2-3</td>
</tr>
<tr>
<td>Critical stenosis</td>
<td>&gt;300</td>
<td>&gt;3</td>
</tr>
<tr>
<td>Occluded</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

PSV – Peak Systolic Velocity; ABPI – Ankle Brachial Pressure Index; Post stenotic PSV: The peak systolic velocity of blood flow in the graft downstream from the stenotic lesion

Table-4: The velocity criteria identifying different categories of vein graft stenosis identified through duplex surveillance (PSV: peak systolic velocity (cm/s), ABPI: ankle brachial pressure index).169
Both Royal Infirmary of Edinburgh and Ninewells Hospitals utilized the Lothian Surgical Database LSA to document the particulars of the initial operative procedure (in coded format as well as the operation note and outpatient clinic visit letters in freehand text).

2.2.2 Data Understanding and processing

Data collected on the regional vein graft databases was cross referenced and supplemented by information obtained from the hospital administrative, radiological and laboratory databases. Postoperative hypercholestrolemia was defined as the total cholesterol to high density lipoprotein cholesterol (total cholesterol/HDL) ratio of greater than 3.5. Renal failure was defined as the Kidney Dialysis Outcomes Quality Initiative (KDOQI) chronic kidney disease (CKD)\textsuperscript{171} stage 4 or 5 at the time of infrainguinal bypass surgery.

Using a combination of data sources, a data warehouse was created by recording patient demographics, type of operation, conduit and follow up information which were recorded in a computerized database (Microsoft Access and Excel, Redmond, Washington, USA). The data warehouse created was a of simple ‘Data Mart’ structure (Figure-21). Data analysis for each study was performed retrospectively. Figure-21 illustrates the KDD process, chronology of data collection and the sources of information.

For each of the four publications included in this chapter, data analysis and reporting was performed in accordance with the reporting standards of the Society for Vascular Surgery and the North American Chapter of the International Society for Cardiovascular Surgery.\textsuperscript{172}
Figure 21: An overview of the steps involved in data mining (KDD) process and study timelines.
Minimal pre-processing of available data was performed and input variables which were not considered to exhibit sufficient validity for inclusion into the analytical model were excluded.

*Missing data:* Every effort was made to minimise missing data. This was performed through cross referencing of available databases and review of medical records if necessary. Due to the contemporary nature of data (Within 5 years) a single attempt was made to contact patients and his/her general practitioner to complete the dataset.

*Class imbalance:* Due to the standardized nature of data collection and the attempts made to complete missing or incomplete data sets none of the variables used for data analysis (patient demographics, atherosclerotic risk factors, results of vein graft surveillance examinations and treatments) where found to exhibit significant class imbalance.

*Validity of class labels:* Class labels used in each study were kept simple and easily verifiable. Cross referencing was used to assure that each class label included met the criteria used to define that label. Table 2 lists all the available variables for analysis.

*Pre-processing:* Minimal pre-processing was performed. The aim of pre-processing was to make sure that the data format was usable for statistical analysis, define events such as graft occlusion (patency), major amputation and follow up time between the initial procedure and these events (timing of the events).
2.2.3 Data analysis

Patients were stratified into two categories based on the pertinent clinical finding which was being studied such as the findings of the initial surveillance scan or minimum diameter of the long saphenous vein. The groups were compared in terms of stenosis, need for re-intervention, graft patency, and amputation. Patency and amputation rates were determined using Kaplan-Meier analysis. The association of graft patency and limb salvage with risk factors for graft failure as ascertained by the site of the distal anastomosis, minimum VGD, and other atherosclerotic risk factors was analysed using the Cox proportional hazards model. The association between the need for re-intervention due to development of a vein graft stenosis and the above variables was assessed using multivariate logistic regression analysis.

2.2.4 Evaluation

Distribution of patient groups by age, gender, diagnoses of diabetes and chronic kidney disease and smoking history as well as type of bypasses being performed remained constant in the 2 Cohorts of patients being studied.

The full text of the 4 publications included in this work-stream can be found in Appendix (iv-8). The results of the data mining exercises are discussed under 4 headings:

2.2.5 Findings of the early postoperative Duplex scan

2.2.6 Using Small diameter vein grafts for infrainguinal bypass surgery

2.2.7 Decision tree Analysis

2.2.8 Endovascular treatment of Vein graft stenosis
2.2.5 Findings of the early postoperative Duplex scan: (Paper –3, appendix iv)

At the time of the first Ultrasound assessment, 65% had no significant stenosis and 35% of grafts had significant flow abnormalities (stenoses). The distribution of these flow disturbances along the length of the vein graft is shown in Figure-22. Of the grafts with a critical stenoses, (52%) were repaired. A further 38% were not repaired and occluded during subsequent follow-up.

<table>
<thead>
<tr>
<th></th>
<th>Number of grafts</th>
<th>PSV at the site of stenosis (cm/s) Median (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical stenosis</td>
<td>29</td>
<td>355 cm/s (310–508)</td>
</tr>
<tr>
<td>Intermediate stenosis</td>
<td>57</td>
<td>316 cm/s (240–345)</td>
</tr>
<tr>
<td>Mild flow disturbance</td>
<td>42</td>
<td>212 cm/s (181–245)</td>
</tr>
<tr>
<td>Low risk grafts</td>
<td>236</td>
<td>–</td>
</tr>
</tbody>
</table>

Figure-22 the distribution of duplex related flow abnormalities along the length of the graft and their severity based on early postoperative assessment.
Forty-three of grafts with severe stenosis were treated, three by intra-luminal angioplasty, 18 by vein patch, 9 by jump graft, eight by interposition graft and five by revision of the bypass. Eight patients required repeat surgical procedures for recurrent graft stenosis; this was by repeat vein patch angioplasty in three, jump graft repair in four and interposition graft in one patient. Forty-nine limbs with critical or intermediate vein graft stenosis were not treated. No statistically significant difference was observed in graft patency (p=0.19) or amputation rates (p=0.62) rates between grafts with repaired stenosis and grafts without stenosis. Untreated grafts with critical or intermediate stenosis had significantly lower patency (p<0.001) and higher amputation rates (p<0.001) rates, (Figures 23 and 24).
Figure 23: Kaplan-Meier plots of primary assisted patency over time in grafts enrolled into the vein graft surveillance program.

<table>
<thead>
<tr>
<th>Category</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Significant Stenosis</td>
<td>232</td>
<td>201</td>
<td>171</td>
<td>142</td>
<td>134</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild Flow Abnormalities</td>
<td>38</td>
<td>29</td>
<td>14</td>
<td>10</td>
<td>6</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mod-Critical Stenosis not repaired</td>
<td>49</td>
<td>38</td>
<td>27</td>
<td>14</td>
<td>8</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mod-Critical Stenosis repaired</td>
<td>43</td>
<td>41</td>
<td>36</td>
<td>26</td>
<td>19</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log Rank = 82, P < 0.001

Figure 24: Kaplan-Meier plots of limb salvage over time in grafts enrolled into the vein graft surveillance program.

<table>
<thead>
<tr>
<th>Category</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Significant Stenosis</td>
<td>232</td>
<td>211</td>
<td>161</td>
<td>162</td>
<td>154</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild Flow Abnormalities</td>
<td>38</td>
<td>29</td>
<td>14</td>
<td>10</td>
<td>6</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mod-Critical Stenosis not repaired</td>
<td>49</td>
<td>41</td>
<td>31</td>
<td>18</td>
<td>11</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mod-Critical Stenosis repaired</td>
<td>43</td>
<td>41</td>
<td>36</td>
<td>26</td>
<td>19</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log Rank = 48, P < 0.001
A- 2.2.6 Using Small diameter vein grafts for infrainguinal bypass surgery (Paper-4, appendix-6)

Table-5 illustrates patient demographics stratified by diameter of vein graft utilised for the bypass procedure.

<table>
<thead>
<tr>
<th></th>
<th>VGD &lt;3.5 mm</th>
<th>VGD ≥3.5 mm</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, years (range)</td>
<td>72 (39-95)</td>
<td>72 (44-89)</td>
<td>p = 0.85</td>
</tr>
<tr>
<td>Gender (M:F)</td>
<td>93:96</td>
<td>171:67</td>
<td></td>
</tr>
<tr>
<td>Indication (claudication:rest pain:tissue loss)</td>
<td>9:81:49</td>
<td>31:115:84</td>
<td></td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>62 (44.9)</td>
<td>88 (37.0)</td>
<td></td>
</tr>
<tr>
<td>Smoker (%)</td>
<td>74 (53.6)</td>
<td>123 (51.7)</td>
<td></td>
</tr>
<tr>
<td>Hyperlipidemia (%)</td>
<td>31 (22.3)</td>
<td>48 (20.2)</td>
<td></td>
</tr>
<tr>
<td>Renal failure (%)</td>
<td>13 (9.4)</td>
<td>20 (9.2)</td>
<td></td>
</tr>
</tbody>
</table>

Median preoperative internal diameter of vein grafts used was 3.9 mm (range 2.4-6.7). The minimum preoperative internal diameter of the vein graft was <3.5 mm in 36.9% and <=3.5 mm in 63.1%. The distribution of VGD in each group is illustrated in Figure-25.

Hyperlipidemia was defined as high-density/low-density lipoprotein ratio >3.5 or fasting cholesterol level >5 mmol/L, and severe renal failure was defined as estimated glomerular filtration rate <20 mL/min or on permanent renal replacement therapy.

AK, above-knee; BK, below-knee; fem-pop, femoropopliteal.

*Yates’ correction.

Table-5: Patient demographics stratified by the minimum internal diameter of venous conduit used for operation.

Figure-25: Distribution of minimum preoperative VGD in grafts which were <3.5 mm compared with grafts which were >=3.5 mm. Mann-Whitney test (P< 0.0001).
A higher proportion of bypasses performed with small diameter vein grafts required intervention to maintain their patency compared with those with diameter $>$ 3.5 mm (Figure 26). The difference in primary assisted patency rates between grafts with preoperative VGD < 3.5 mm and grafts with preoperative VGD diameter equal or greater than 3.5 mm did not reach statistical significance (log rank = 2.29, $p = 0.13$) (Figure-27). In addition, there was no significant difference in amputation rates between vein grafts which were created using veins with a minimum preoperative diameter < 3.5 mm compared to those with a minimum preoperative diameter equal or greater than 3.5 mm (log rank = 1.69, $p = 0.35$), (Figure-28).

Figure-26: Comparison of the proportion of vein grafts requiring re-intervention to maintain graft patency, stratified by the minimum preoperative internal diameter of the vein graft.
<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>2.87 (1.63-3.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.89 (0.71-1.05)</td>
<td>0.160</td>
</tr>
<tr>
<td>Renal failure</td>
<td>1.16 (0.85-1.30)</td>
<td>0.310</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.83 (1.39-3.20)</td>
<td>0.022</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.39 (0.90-2.33)</td>
<td>0.113</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>1.86 (1.49-2.47)</td>
<td>0.019</td>
</tr>
</tbody>
</table>

Table-6: Factors affecting the need for re-intervention in infrainguinal vein grafts.

![Figure-27](image)

**Figure-27:** primary assisted patency of vein grafts in patients who underwent infrainguinal vein graft bypass, stratified by preoperative minimum internal diameter of the vein graft.
Figure-28: Kaplan-Meier plot of amputation rates over time in patients who underwent infrainguinal vein graft bypass stratified by preoperative minimum internal diameter of the vein graft.

There was no significant association between minimum preoperative VGD and reduced primary assisted patency (Table-7) or higher amputation rates (Table-8).
Table 7: Factors affecting primary assisted patency after infrainguinal bypass assessed using Cox’s multivariate regression analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>1.34 (0.94-2.55)</td>
<td>0.0901</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.40 (1.39-4.13)</td>
<td>0.0017</td>
</tr>
<tr>
<td>Renal failure</td>
<td>2.90 (1.65-5.10)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.46 (1.49-4.06)</td>
<td>0.004</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>0.97 (0.44-2.16)</td>
<td>0.95</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>0.85 (0.47-1.54)</td>
<td>0.2776</td>
</tr>
</tbody>
</table>

Table 8: Factors affecting amputation rates after infrainguinal bypass assessed using Cox’s multivariate regression analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>1.33 (0.78-2.29)</td>
<td>0.295</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.81 (1.62-4.88)</td>
<td>0.002</td>
</tr>
<tr>
<td>Renal failure</td>
<td>3.93 (2.02-7.66)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.29 (1.29-4.05)</td>
<td>0.005</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>0.8339 (0.37-1.87)</td>
<td>0.659</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>0.5876 (0.29-1.19)</td>
<td>0.138</td>
</tr>
</tbody>
</table>
2.2.7 Decision tree Analysis (Paper-5, appendix 7)

A Classification and Regression Tree Model (CART) was used to identify patients who are likely to develop significant vein graft stenosis or occlusion with the aim of identifying patients who would benefit from more intensive duplex ultrasound based follow up. CARTs have been described in the introduction chapter (Page-11, figure-1). The variables used were clinical and duplex ultrasound findings, which were available at the time of first postoperative assessment. The Gini coefficient was used to assess heterogeneity in the distribution of patients from the parent node to children nodes. The Gini coefficient is zero when all observations at a node belong to one level of a dependent variable and is 0.5 when observations are equally distributed in various levels of a dependent variable. The best split on a variable will be the one that minimizes the Gini coefficient. This process continues until a termination criterion is met.\textsuperscript{173}

The level of significance of each variable will be adjusted by Bonferroni’s correction; where the decision tree is first fitted to the data, with the most complex topography and the greatest number of nodes and hence accuracy. Prediction at each node is made based on the weight of observations in each category of the dependent variable and misclassification cost. The CART then undergoes a process of pruning, during which the terminal nodes will be deleted only if its elimination causes a misclassification cost which is significantly lower than the reduction in complexity. This process continues to reach the root node. The optimal CART is selected as the final predictive model. The optimal CART selected was one that prioritized sensitivity over specificity, as its purpose was to relate the decision or an attribute to its consequences whilst maintaining the simplest topography.\textsuperscript{173}
The aim of the model was to repeatedly divide patients into subgroups, each of which consisted of patients with or without severe vein graft stenosis or occlusion. Thereby stratifying patients into high-risk groups who should undergo duplex ultrasound based surveillance and low-risk groups, who do not require intensive follow up. The CART model was then used to describe the possible pathways of different cohorts during the first 2 years of vein graft surveillance.

Figure-29: Tight Proximal graft stenosis identified on Duplex US, with angiographic views of the same stenosis before and after angiographic treatment using drug eluting 4mm x 4cm balloon angioplasty.
The results of Cox regression analysis of potential risk factors associated with infrainguinal vein graft bypass graft stenosis or occlusion are shown in Table-9.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% C.I.)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.03 (0.62-1.36)</td>
<td>0.81</td>
</tr>
<tr>
<td>Gender</td>
<td>0.98 (0.71-1.10)</td>
<td>0.72</td>
</tr>
<tr>
<td>Critical ischaemia</td>
<td>1.41 (0.79-2.60)</td>
<td>0.18</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.86 (1.65-4.97)</td>
<td>0.002</td>
</tr>
<tr>
<td>Smoking (within 3 months of bypass)</td>
<td>2.613 (1.51-4.53)</td>
<td>0.006</td>
</tr>
<tr>
<td>Previous failed ipsilateral bypass</td>
<td>2.51 (1.41-4.32)</td>
<td>0.008</td>
</tr>
<tr>
<td>Non reversed or in situ saphenous vein</td>
<td>0.87 (0.51-1.21)</td>
<td>0.52</td>
</tr>
<tr>
<td>Non single segment saphenous vein</td>
<td>0.79 (0.36-1.80)</td>
<td>0.63</td>
</tr>
<tr>
<td>Infrainguinal distal anastomosis</td>
<td>2.40 (1.31-4.47)</td>
<td>0.01</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.80 (0.36-1.80)</td>
<td>0.6</td>
</tr>
<tr>
<td>Significant flow abnormalities on early post operative Duplex scan</td>
<td>3.22 (1.63-4.69)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table-9: Cox regression analysis of potential factors associated with infra-inguinal vein graft bypass graft revision
Seven hundred and thirty six patients were classified through 5 items into 10 different categories with 2-year stenosis or occlusion rates ranging from 5% to 95%. The CART model stratified 299 grafts (40.8%) as low-risk and 433 (59.2%) as high risk grafts (Figure 30).

Figure 30: Decision tree based on patient profile and first postoperative graft surveillance assessments. The table outlines the Gini coefficient* and the accuracy of prediction** for each node. The Gini coefficient* relates the performance of the node in portioning the patients based on the eventual outcome. The accuracy of prediction** is the accuracy of predicting the eventual outcome based on a single node.
Figure-31 illustrates the final topography of the decision tree and the proportion of patients assigned to each group. The model was able to stratify high-risk grafts that were at risk from significant vein graft stenosis or occlusion with good accuracy. Area under the ROC curve for prediction of graft stenosis or occlusion was 0.88 (95% Confidence Interval (CI): 0.81-0.94 percent). The indicators of accuracy of the decision tree model at correctly classifying the CART model are listed in Table-10. Figures 31 and 32 illustrate the Kaplan-Meier survival curves for primary graft patency and amputation rates for grafts that have been classified as high- and low-risk based on the model. They show that grafts classified as high-risk had significantly lower primary patency rate (Log rank=186, P<0.001) and were associated with higher risk of amputation (Log rank=118, P<0.001) during follow up.

<table>
<thead>
<tr>
<th>Value</th>
<th>Estimated value</th>
<th>95-percent Confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of high risk grafts</td>
<td>0.27</td>
<td>0.25</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.92</td>
<td>0.88</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.52</td>
<td>0.45</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>0.45</td>
<td>0.38</td>
</tr>
<tr>
<td>Negative Predictive value</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>Positive Likelihood ratio</td>
<td>1.55</td>
<td>1.44</td>
</tr>
<tr>
<td>Negative Likelihood ratio</td>
<td>0.6</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table-10 Indicators of accuracy of classification of the CART (decision tree) analysis for prediction of development of vein graft stenosis or occlusion within the first 2 years of follow up.
Figure 31: Kaplan Meier survival curve showing the primary patency rate in the grafts that have been selected to be high-risk versus low-risk by the decision tree.

Figure 32: Kaplan Meier survival curve showing amputation rates in the grafts that have been selected to be high-risk versus low-risk by the decision tree.
A recent study (Paper-7 Appendix-8) has successfully validated the decision tree in a completely new population of patients. The results of this validation study are illustrated in figures-31 to 35.

Figure-33 Decision tree based on patient profile and first postoperative graft surveillance assessments. (Validation population)

Grafts classified as high risk: 154 (61.8%)
Grafts Classified as low risk: 95 (38.2%)
Figure-34 Kaplan Meier survival curve comparing primary assisted patency (y axis) of grafts which were predicted to be high risk (Red) using the decision tree model compared with low risk grafts (Black), (Log Rank=70, P<0.001).

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>10 Months</th>
<th>20 Months</th>
<th>30 Months</th>
<th>40 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>95</td>
<td>83</td>
<td>76</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>High risk</td>
<td>154</td>
<td>79</td>
<td>63</td>
<td>41</td>
<td>28</td>
</tr>
</tbody>
</table>
Figure-35: Kaplan Meier survival curve comparing amputation rates in patients who had undergone infra inguinal bypass grafts which were predicted to be high risk (Red) using the decision tree model compared with amputation rates in low risk grafts (Black), (Log Rank=42, P<0.001)\textsuperscript{244}
2.2.8 Endovascular treatment of Vein graft stenosis (Paper-7– appendix 4)

Seventy-six grafts underwent a total of 99 endovascular procedures (Figures-42 and 43). Median age of the grafts at the time of the first angioplasty was 5 months (range: 7 weeks-27 months). Forty-one primary angioplasties were performed with a standard balloon and 34 with a cutting balloon. Initial technical success was achieved in 60 grafts (78.9%).

Of the grafts in which initial technical success had not been achieved, eight underwent repeat angioplasty with good technical success and another graft was found to have thrombosed the next day. This graft was successfully managed with local catheter thrombolysis with good results. Three grafts occluded prior to attempted surgical repair and another four grafts were surgically repaired. A total of 27 grafts (45%) developed re-stenosis of which 24 underwent a repeat attempt at angioplasty, with technical success in 17 grafts (71%). Repeat angioplasties were performed with cutting balloons. Of the grafts that had developed re-stenosis three occluded prior to the repeat attempt at angioplasty and a further patient was deemed too unfit to undergo further intervention. There was no peri-procedural mortality. Vein graft angioplasty was associated with two puncture site haematomas both of which were treated conservatively; in addition one patient who had undergone cutting balloon angioplasty developed a delayed retroperitoneal haemorrhage from a contralateral puncture site which was managed by open surgery.
Figure-36: Digital subtraction angiograms of a common femoral to below knee popliteal artery bypass graft. The left image shows a proximal graft stenosis (arrowhead), whilst the right image shows the final result after cutting balloon PTA with a 4.0-mm diameter balloon.

Figure-37: Digital subtraction angiograms of a common femoral to above knee popliteal artery bypass graft. The left image shows two distal anastomotic graft stenoses (arrowhead), whilst the right image shows the final result after cutting balloon PTA with a 3.0-mm diameter balloon.
Overall 30-month primary patency, primary assisted patency and secondary patency rates were 73.2%, 82.6% and 84.3%, respectively. No significant difference was observed in graft patency (log rank=1.83, \( P=0.08 \)) or amputation rates (log rank=0.89, \( P=0.32 \)) between the grafts requiring intervention to maintain patency and grafts which did not (Figure-38 and 39).

---

**Figure-38** Kaplan-Meier survival curves of Primary assisted patency.
Tables-12 and 13 list the factors affecting primary assisted patency and amputation rates after infra-inguinal bypass surgery assessed using Coxs’ multivariate regression analysis.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.03 (0.62–1.36)</td>
<td>0.85</td>
</tr>
<tr>
<td>Gender</td>
<td>0.98 (0.71–1.10)</td>
<td>0.61</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.55 (1.49–4.35)</td>
<td>0.006</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.613 (1.51–4.53)</td>
<td>0.006</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.80 (0.36–1.80)</td>
<td>0.592</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.89 (1.19–3.38)</td>
<td>0.040</td>
</tr>
<tr>
<td>Intervention to maintain graft</td>
<td>1.41 (0.79–2.60)</td>
<td>0.18</td>
</tr>
<tr>
<td>patency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recurrent stenosis</td>
<td>3.22 (1.63–4.69)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table-11**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.12 (0.41–1.53)</td>
<td>0.41</td>
</tr>
<tr>
<td>Gender</td>
<td>0.86 (0.51–1.21)</td>
<td>0.52</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.8618 (1.65–4.97)</td>
<td>0.002</td>
</tr>
<tr>
<td>Smoking</td>
<td>4.03 (2.07–7.84)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.93 (0.42–2.057)</td>
<td>0.852</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>2.49 (1.39–4.47)</td>
<td>0.021</td>
</tr>
<tr>
<td>Intervention to maintain graft</td>
<td>1.24 (0.68–1.91)</td>
<td>0.310</td>
</tr>
<tr>
<td>patency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recurrent stenosis</td>
<td>2.51 (1.41–4.32)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

**Table-12**
2.3 Classification using Intelligent Analysis

2.3.1 Assessment of degree of ICA stenosis using an artificial neural network (Paper-8, appendix -9)

Consecutive patients who were investigated for ICA stenosis in a tertiary referral vascular surgery unit (St Vincent’s University Hospital, Dublin) were studied prospectively. All patients underwent carotid duplex ultrasound scanning as well as intra-arterial digital subtraction angiography (Figure-40) for assessment of the degree of extracranial carotid artery stenosis during the same episode of care. Measurement of the degree of ICA stenosis using Intra-arterial digital subtraction angiography was the gold standard against which the ANN and duplex velocity criteria were assessed. Patients data including demographics, co-morbidities, Duplex US cerebral CT results and Intra-arterial digital Subtraction angiography were recorded (by Reza Mofidi) in a Microsoft Access database.

Figure-40: Intra-arterial Digital Subtraction angiography of Left Carotid bifurcation
2.3.2 Neural Network Design

Multi Layered Perceptron MLP models were developed using Neuro-Solutions™ version 4 software, (NeuroDimension, Gainesville, FL). The MLP models which were developed had six input nodes, consisting of entries corresponding to PSV and EDV in the ICA and pre-stenotic CCA, presence of contralateral ICA occlusion, and a bias node (figure-41). The MLP layer had 3 hidden layers. Each node had a sigmoid transfer function of the hyperbolic tangent variety. Learning occurred through back propagation of error a process which generated a learning curve which plotted the reduction in mean squared error of observations in real-time (Figure-42). The input data did not undergo any pre-processing steps prior to input into the neural network apart from randomisation of data sets.

![Neural Network Diagram](image)

<table>
<thead>
<tr>
<th>Input Variables</th>
<th>Output variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICA Peak Systolic Velocity (cm/s)</td>
<td>Continuous</td>
</tr>
<tr>
<td>ICA End Diastolic Velocity (cm/s)</td>
<td>Continuous</td>
</tr>
<tr>
<td>CCA Peak Systolic Velocity (cm/s)</td>
<td>Continuous</td>
</tr>
<tr>
<td>CCA End Diastolic Velocity (cm/s)</td>
<td>Continuous</td>
</tr>
<tr>
<td>Contralateral ICA occlusion</td>
<td>Binary 1 or 0</td>
</tr>
<tr>
<td>Bias node</td>
<td>1</td>
</tr>
<tr>
<td>Degree of ICA stenosis (%)</td>
<td>Degree of ICA stenosis (Bands)</td>
</tr>
</tbody>
</table>

Figure-41 illustrates the graphical representation of a MLP as viewed on the Neuro-Solutions™ software user interface.
Two sets of ANN models were created the first set was tasked with calculating the exact degree of ICA stenosis, whilst the second set was tasked with stratification of degree of ICA stenosis to bands of predefined 10% intervals. The available data were randomly divided into training and validation data sets. The larger data set (60% of total data set) was used to train the model; the other (40% of the total data set) was used to validate the model. These were organised into eight different randomly selected reruns of data and were used to train eight identical MLP models. Four of these were used for prediction of exact degree of ICA stenosis and four to stratify the stenosis according to 10 bands.

2.3.2.1 Training: The input variables used were PSV and EDV in ICA as well as CCA and presence of contralateral ICA occlusion. Sixty percent of the data set was used to train the neural network. The training rule was back-propagation of error.

2.3.2.2 Validation: As described earlier 40% of the available data was not used for training of the neural network model. The mean squared error was used to express the variation between the degrees of predicted value compared to the gold standard which was angiographic measurement of internal carotid artery stenosis.

2.3.2.3 Statistical Analysis: The degree of ICA stenosis predicted by the ANN was compared with the angiographic measurements using linear regression and Bland-Altman analysis.

To compare the duplex velocity criteria with the ANN in correct banding of the degree of ICA stenosis, a measure of the discriminant power (DP) of a test was used:

\[
DP = \frac{1}{\pi} \left\{ \ln[\text{Sen}/(1 - \text{Spec})] + \ln[\text{Spec}/(1 - \text{Sen})] \right\}
\]
Evaluation

Two hundred and eight carotid bifurcations in 104 patients were studied. Eighty-three patients had ipsilateral symptoms prior to presentation (hemispheric, 51; ocular, 27; both symptoms, 5), and 21 patients were asymptomatic. The mean age of subjects was 70.1 years (standard deviation = 10.11); 38 were female and 66 were male.

Duplex Velocity Criteria and the Degree of ICA Stenosis

The overall agreement between the results of duplex ultrasound velocity measurement and angiography in determining the correct banding for the degree of ICA stenosis was 91.8% (191/208), with a sensitivity of 89.7% [95% confidence interval (CI) 82.8-94] and specificity of 93.5% (95% CI 86.5-97) (DP = 1.67). The accuracy of duplex ultrasound criteria at correctly identifying >70% ICA stenosis was 95.2%, with sensitivity of 96.3% (95% CI 90.9-98.6), specificity of 94% (95% CI 87.5-97.3), and DP of 3.3. Carotid duplex examination was also able to identify over 50% ICA stenosis with reasonable accuracy [sensitivity 96.38% (95% CI 89.9-97.5), specificity 91.6% (95% CI 82.8-96.1), and accuracy 94.7% (DP = 3.09)].

Neural Network Analysis

Two hundred and eight data vectors were randomly divided as follows: 124 (59.6%) were used to train the neural network model and 84 (40.4%) were used for model validation. Following completion of training, the ANN model was able to predict the degree of ICA stenosis with a high degree of accuracy (mean squared error 0.012); cross-validation confirmed the validity of the ANN model at predicting the degree of ICA stenosis (Figure-42).
Figure-42: The training curve of the ANN used for banding the degree of ICA stenosis showing the minimization of mean squared error (MSE, y axis) during the training process (x axis = number of epochs). The solid line represents the training process and the broken line, cross-validation.

Good correlation was observed between the degree of ICA stenosis predicted by the ANN and that measured angiographically (R² = 0.9374, p < 0.0001) (Figure-43). Bland-Altman analysis revealed a high degree of agreement between the results of ANN and angiographic measurements (Figure-44). Furthermore, the ANN model was able to accurately band the ICA stenosis according to the predefined 10% intervals with a sensitivity of 97.3% (95% CI 90.7-99.3), a specificity of 97.7% (95% CI 93.6-99.2), and an overall accuracy of 97.5%.
Figure-43: A plot of the angiographic degree of ICA stenosis (x axis) and ANN prediction of ICA stenosis in the validation sample (y axis).

Figure-44: Bland-Altman analysis applied to the comparison of predicted degree of ICA stenosis by ANN and the angiographic degrees of ICA stenosis.
The ANN model was more accurate at correctly banding the degree of ICA stenosis. The DP of ANN was 4.11, while that of duplex velocity criteria alone was 1.67 (p < 0.05). The accuracy of the ANN at correctly identifying >70% ICA stenosis was 98.4%, with sensitivity of 96.4% (95% CI 93.8-99.3), specificity of 98.7% (95% CI 93.4-99.8), and DP of 4.21. This was comparable to duplex velocity criteria (p = nonsignificant). The accuracy of the ANN at correctly identifying >50% ICA stenosis was 97.4%, with a sensitivity of 97.8% (95% CI 93.8-99.3 and) specificity of 98.2% (95% CI 90.4-99.7), which was not significantly better than the duplex velocity criteria. Table 13 illustrates the comparison between duplex velocity criteria and ANN at identification of >50% and >70% stenosis as well as correct banding of ICA stenosis for >50% ICA stenosis. The use of ANN was associated with altered clinical decision in 2.9% of cases (5/208) compared with angiographic findings. This was not significantly different from the duplex velocity criteria, use of which was associated with altered clinical decision in eight carotid bifurcations (3.8%) compared with angiography (p = NS).

<table>
<thead>
<tr>
<th>Modality</th>
<th>Identification of &gt;70% stenosis (p = NS)</th>
<th>Identification of &gt;50% stenosis (p = NS)</th>
<th>Correct 10% banding for stratifying 50-100% stenosis (p &lt; 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>Duplex velocity criteria</td>
<td>95.2</td>
<td>96.3 (90.9-98.6)</td>
<td>94 (87.5-97.3)</td>
</tr>
<tr>
<td>ANN</td>
<td>98.4</td>
<td>96.4 (93.8-99.3)</td>
<td>98.7 (93.4-99.8)</td>
</tr>
</tbody>
</table>

NS, nonsignificant; DOR, diagnostic odds ratio (95% CI in parentheses).
2.4 Prediction of Outcome using Intelligent Analysis

2.4.1 Prediction of Outcome following resectional surgery for Oesophageal and OG junction carcinoma. (Paper–9 appendix-10)

2.4.1.1 Data understanding

Consecutive patients who had undergone resection of carcinoma of the oesophagus and gastro–oesophageal junction in a regional specialist oesophago-gastric surgery unit between January 1995 and August 2004 were included. The data used was National Minimum Core Dataset for Upper Gi Cancer which had been developed by the ISD in Scotland.

2.4.1.2 Neural network design

Multilayer perceptron (MLP), with back propagation circuit were constructed using NeuroSolutions version-4 neural network software (NeuroDimension, Inc., Gainesville, FL, USA). The ANNs were trained through back propagation of error, which is a process by which the error of prediction is minimized by adjusting the weights associated with the synaptic connections in the hidden layers of the ANN. The MLP layer had 4 hidden layers. Each node had a sigmoid transfer function. Learning occurred through back propagation of error a process which is described in the introduction chapter (pages 12-14, figure-2). This process generates a learning curve which plotted the reduction in area under the receiver operator characteristic curve versus the number of observations used in the training process (figure-46).
Two sets of artificial neural network models were created which were used to predict disease free survival at 1 year and at 3 years. The input data did not undergo any pre-processing steps and were randomly assigned into two mutually exclusive datasets. One of these two datasets was used to ‘train’ the ANN (60% of the available dataset), a second to assess the accuracy of predictions of ANN trained using the first dataset (40% of the available dataset). In order to minimize bias based on selection of cases for training and evaluation datasets, panels of 10 identical ANNs were used for each prediction (disease free survival at 1 and 3 years). These ANNs were trained and evaluated independently of each other. The values pertaining to the accuracy of prediction were the average output for the ANN panel (with 95% confidence intervals). The training process was considered complete at the end of eight iterations of training or if the mean squared error of predictions was less than 0.025. The optimum variables used to construct the artificial neural network were selected from 42 potential input variables available using sensitivity analysis. During the process of sensitivity analysis, the network learning is disabled so that the weights associated with the synaptic connections were not affected. The inputs to the network are shifted slightly and the corresponding change in the output is reported as a percentage. The values that cause the highest percentage of change are selected to create the model. Using sensitivity analysis, input variables are categorized according to their ability in influencing outcome. The optimal number of variables for developing the ANN was assessed during backward variable selection, a process whereby the variables with the least predictive value are consecutively removed and the accuracy of prediction of the ANN model is evaluated by receiver operator characteristic analysis, until the ANN with the optimum accuracy of prediction is reached. The predictive power of the UICC TNM classification system in assessing disease free survival at 1 and 3 years was analysed using stepwise linear
discriminant analysis. The variables used to create this model are the T, N and M stage as well as clearance of the resection margins (R₀: clear margins, R₁: microscopically involved margins, R₂: macroscopic residual disease). The accuracy, sensitivity, specificity, positive and negative predictive values as well as likelihood ratios were calculated for the UICC TNM staging system and ANN as predictors of disease free survival. To compare the TNM classification with ANN in predicting disease free survival a measure of the discriminant power (DP) of a test was used.

2.4.1.3 Evaluation

Two hundred and sixteen patients underwent resectional surgery for oesophageal and OG junction carcinoma of whom 155 were male and 61 were female. Median age of the patients was 64 years (range: 30–82). One hundred and sixty-one patients underwent primary surgical resection of the lesion and 55 received neoadjuvant, cisplatin and 5-flourouracil based chemotherapy prior to the definitive surgical procedure. One patient underwent Argon laser ablation of an early adenocarcinoma of OG junction followed by subsequent definitive surgical resection. One-year follow-up data was available in 198 patients and 126 patients had 3-year follow-up data available. Table-14 describes the pathological feature of the lesions based on the examination of the histological specimen. The Kaplan–Meier survival curves based on the UICC classification are shown in figure-45.
Pathological characteristics of the resected specimen (N=216)

<table>
<thead>
<tr>
<th>pN stage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N0</td>
<td>85</td>
</tr>
<tr>
<td>N1</td>
<td>131</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>pM stage (coeliac node involvement)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>209</td>
</tr>
<tr>
<td>M1</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Submucosal lymphatic invasion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>137</td>
</tr>
<tr>
<td>Present</td>
<td>79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vascular invasion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>154</td>
</tr>
<tr>
<td>Present</td>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neural involvement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>168</td>
</tr>
<tr>
<td>Present</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of resection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R₀</td>
<td>169</td>
</tr>
<tr>
<td>R₁</td>
<td>36</td>
</tr>
<tr>
<td>R₂</td>
<td>11</td>
</tr>
</tbody>
</table>

Table-14

Figure-45 Kaplan–Meier survival curves based on the UICC classification, (log-rank test, P<0.01).
Artificial neural network analysis

Table-15 illustrates the sample size and composition of training and evaluation datasets for the panel of neural networks used to predict disease free survival at 1 and 3 years. Using sensitivity analysis 14 input variables were selected to construct the ANN. Table 16 illustrates the input values which were selected through sensitivity analysis to construct the ANN as well as their sensitivity for prediction of disease free survival at 1 and 3 years. Figure-46 shows the change in the accuracy of area under ROC curve in response to the reduction in the number of input variables.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-Year</td>
<td>3-Year</td>
</tr>
<tr>
<td>N</td>
<td>119</td>
<td>73</td>
</tr>
<tr>
<td>Age (range)</td>
<td>64 (41–80)</td>
<td>64.5 (30–82)</td>
</tr>
<tr>
<td>Stage I</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Stage II_A</td>
<td>32</td>
<td>20</td>
</tr>
<tr>
<td>Stage II_B</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Stage III</td>
<td>55</td>
<td>33</td>
</tr>
<tr>
<td>Stage IV</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>R_0</td>
<td>93</td>
<td>55</td>
</tr>
<tr>
<td>R_1</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>R_2</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Survival (%)</td>
<td>85</td>
<td>38</td>
</tr>
</tbody>
</table>

Table-15
The mean squared error (MSE) reached at the completion of training of the ANN for prediction of survival at 1 year was 0.0082, whilst the final training MSE for prediction of 3-year survival was 0.006. Figure-43 shows the learning curve of the artificial neural network used for prediction of survival at 3 years with minimization of mean squared error during training process. ANN was able to accurately predict disease free survival at 1 year (accuracy: 88%, DP=2.3). This was significantly more accurate than the 1-year survival prediction using the UICC TNM classification system (accuracy: 71.6%, DP=1), (P<0.01). The ANN panel was significantly more accurate at predicting survival at 3 years after surgery (accuracy: 91.5%, DP=2.72) than the UICC TNM classification system (accuracy: 74.7%, DP=1) (P<0.05) (Table-16).
### Input variables used to construct the ANN models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data format</th>
<th>Sensitivity (%)</th>
<th>1-Year</th>
<th>3-Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Absolute value</td>
<td>1.67</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>Absolute value</td>
<td>0.91</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>American Society of Anaesthesiology ASA score</td>
<td>I–4</td>
<td>1.31</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Presence of dysphagia</td>
<td>Y/N</td>
<td>1.04</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>Histological type</td>
<td>Adenocarcinoma</td>
<td>0.76</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Squamous cell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td>Well</td>
<td>1.31</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T stage</td>
<td>T1</td>
<td>1.89</td>
<td>2.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of positive nodes</td>
<td>Number</td>
<td>1.87</td>
<td>2.21</td>
<td></td>
</tr>
<tr>
<td>Total nodes retrieved</td>
<td>Number</td>
<td>1.05</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>Coeliac node involvement</td>
<td>Y/N</td>
<td>2.01</td>
<td>1.97</td>
<td></td>
</tr>
<tr>
<td>Margin involvement</td>
<td>Y/N</td>
<td>1.99</td>
<td>1.83</td>
<td></td>
</tr>
<tr>
<td>Lympho-vascular involvement</td>
<td>Y/N</td>
<td>1.32</td>
<td>1.55</td>
<td></td>
</tr>
<tr>
<td>Neural invasion</td>
<td>Y/N</td>
<td>1.13</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant chemotherapy</td>
<td>Y/N</td>
<td>1.08</td>
<td>1.26</td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity of each variable for 1- and 3-year survival. These variables were selected through sensitivity analysis. These variables were selected through sensitivity analysis.

**Table-16**
Figure-46: Changes in the accuracy of the ANN survival at 1-year (red line) and 3-year (black line), plotted against the number of input variables used to create the ANN. With the reduction in the number of variables there is a progressive increase in the accuracy of the ANN (Y-axis) until the optimum number of variables is reached.
Chapter-3  Discussion

Much of practice of general and vascular surgery involves identifying patients with clinical conditions in pre-symptomatic state or after a herald event and planning prophylactic procedures such as angioplasty, stenting or surgical revision where the indication for intervention is preventing a future event such as disabling stroke, rupture of abdominal aortic aneurysm, graft occlusion or limb loss. Individualizing patient care to suit their condition runs counter to an approach which involves following a care pathway with the same steps for each patient once they met predefined criteria such as an AAA greater than 5.5 cm in diameter, or greater than 70% ICA stenosis. The criteria selected for intervention as well as the pathway of care for surveillance are often supported by overwhelming evidence. With regards to vein graft surveillance the situation is further complicated by the fact that the majority of grafts do not develop stenosis and there is no evidence that beyond the early post-operative scan further US based surveillance of vein grafts has any impact on graft patency or amputation rates. With regards to prediction of future outcome the situation is different as the decision to intervene (or not to) can significantly alter the outcome. This narrative set out to examine how data mining tools can help identify and individualize risk benefit analysis in each clinical context.
Appendix-2 (Paper-1) illustrates the challenges involved in analysing longitudinal data collected over several decades. It also illustrates how such data can provide clinically important information once issues regarding missing data and data imbalance are addressed. This publication reported the natural history of AAA in a large cohort of patients over a relatively long duration of time, significantly longer than the duration of follow up of any randomised controlled trial and this has proven to be its lasting legacy. Female sex was found to be an independent risk factor for AAA expansion rate, independent of initial aneurysm size and age on presentation. Another regional community-based AAA screening and surveillance programme has reported similar observations on a total of 274 men and 74 women.176 A similar sized referral-based study also reported that female sex was associated with increased AAA growth177. A number of longitudinal studies have revealed that aneurysms that rupture have significantly higher expansion rates than those which do not rupture in a specified time-frame180-184. Brown et al. reported that ruptured AAAs had significantly higher mean expansion rates (0·84 cm per year) than intact AAAs (0·39 cm per year) but did not report a gender difference.185

RESCAN collaboration which was the seminal publication on the frequency of follow up scans for patients who have AAA below the size that requires early repair. It was based on a meta-analysis of 18 publications worldwide including 15,475 patients under surveillance186. It revealed a significant degree of heterogeneity in growth rates of AAA186. This heterogeneity was thought to be related to the patient population included in each study, i.e. whether patients were identified through screening (these studies were skewed towards
the smaller size and included mainly men) or surveillance of patients with small, moderate sized AAA and whether the study design included a treatment arm for small-moderate sized AAA. A treatment arm would affect the growth rate of AAA as it would reduce the duration of follow up in half of patients each study (those randomised to the treatment arm). Duration of follow up is the denominator in the calculation of growth rate of AAA, the rate of change in AP diameter of AAA being the numerator for this calculation. Study-1 (Appendix-2) was one of the 18 seminal publications which contributed to the RESCAN study (Figure-44).186

Figure-44: Estimated time for which there is a 10% probability of reaching the threshold diameter for surgery (5.5 cm) as a function of baseline diameter and study of origin (in men). Overall results are in black diamonds.186 Paper-1 (Appendix-2) is listed as Edinburgh study.

Brown et al.184 studied a cohort of 476 patients with AAA and reported a fourfold higher rupture rate in women. The RESCAN collaboration confirmed the higher rupture rates of AAA in women compared with men.186 Although we elected not to compare the risk of
rupture in men and women with AAA as we were concerned that at face value this data may have been imbalanced, the data from our study did contribute to the estimation of rupture rates from AAA stratified by size by the RESCAN collaborators\textsuperscript{186}. This was done by knowing which patients had died, and the proportion of patients in whom the diagnosis of ruptured AAA had been confirmed. Using mathematical modelling the RESCAN collaborators were able to estimate the annual risk of rupture in each of the subpopulation of patients (classified by class of interest such as gender). When a large quantity of data is available it is possible to use compensate for missing or imbalanced data through pre-processing or mathematical modelling.\textsuperscript{186}

The author acknowledges that with hindsight their approach on Paper-1, appendix-2 which was to exclude variables with missing data (or those which we suspected of class label imbalance) was too conservative. \textit{“Missing data are a recurring problem that can cause bias or lead to inefficient analyses.”} \textsuperscript{189} The purist clinician’s view is if you don’t have the completed data you should not do the analysis (or exclude the variables with uncertainty) and this was the approach followed by us. Ironically where we to be faced with the same problem today we would employ an imputation technique (data mining) to solve the problem. There are number readily available imputation methods which use assumptions about missing variables or assigning weights to complete datasets or using Bayesian frameworks to address imputation problems. There are a number of software programs which are used routinely to deal with the problem of missing data.\textsuperscript{189}

Data mining exercises in Europe, North America and Australia have provided interesting insights into incidence of rupture, modality of treatment and mortality from AAAs.\textsuperscript{190-198} They have provided complementary evidence for management of AAA to randomised
controlled trials which are discussed in detail in Appendix-1\textsuperscript{199-202}. Screening for AAA with abdominal ultrasonography has been shown to be cost-effective in men and has led to a reduction in aneurysm-related deaths\textsuperscript{199-203}. Universal screening of men between the ages of 65 and 80 years is associated with a significant reduction in the regional incidence of ruptured AAA. There is no evidence that screening for AAA in an unselected population of women is associated with a reduction even in aneurysm related mortality.\textsuperscript{22} Scott and colleagues\textsuperscript{55} reported the prevalence of AAA in women to be 1.3%, with other authors reporting a similar rate of 0.7–1.3% in unselected populations\textsuperscript{204-207}. Scott\textit{et al.} did not demonstrate a difference in rupture rates between the women randomized to screening and control populations of women at 5-and 10-year follow-up\textsuperscript{206}. They concluded that screening for women was neither clinically indicated nor economically viable.\textsuperscript{56} For any screening for AAA to be effective in reducing mortality in women, it will need to be concentrated on women who are at high-risk of having AAA and are also fit to undergo aneurysm repair.\textsuperscript{207} There is some evidence that women with atherosclerotic disease are at significantly higher risk of developing AAA.\textsuperscript{207-213}

The proliferation of endovascular techniques for repair of AAA has had a significant effect on reduction of mortality and morbidity from repair of AAA.\textsuperscript{214,216} Sensi\textit{et al} revealed a significant increase in the number and proportion of patients who underwent endovascular repair of AAA between 2006 and 2011 in Italy (Figure-45).\textsuperscript{214} A recent publication from the Vascular Study Group of New England revealed that the gender differences in mortality from repair of AAA have persisted in the endovascular era (Figure-46).\textsuperscript{215} This difference is largely attributable to age, type of repair, urgency at presentation and comorbidities.\textsuperscript{215}
Women were less likely to undergo EVAR for intact aneurysms. The most likely reason for this is that women less frequently meet the anatomic criteria for EVAR. Chung et al reported that long term, endovascular repair of AAA is more likely to be associated with complications and mortality in women.

A recently published study by Sidloff et al suggested that women undergoing elective AAA repair in the UK continue to have a significantly higher in-hospital mortality rate compared to men after either open repair or EVAR despite the introduction of VSQIP. Sidloff and his co-workers utilized almost identical methodology to Mofidi et al (Appendix-3). Like Mofidi et al. they were unable to explore the reasons for these differences as they existed even after adjusting for the known risk factors for mortality of AAA. It appears therefore that the introduction of endovascular techniques and quality improvement programs have not had the desired effect on in hospital morality for elective or emergency repair of AAA in women. A number of authors have reported significantly higher mortality in women undergoing elective EVAR as well as elective open repair of abdominal aortic aneurysm.

Although aneurysm morphology is difficult to adjust for in studies, there is some evidence that women are more likely to have shorter, more angulated infrarenal necks and small iliac arteries, making EVAR more challenging. Aneurysms with shorter and more angulated infrarenal necks are more likely to be outside of recommendations for use for EVAR devices. These anatomical differences may explain the worse outcomes in women and explain why the exact causes for mortality are so difficult to characterise. A recent systematic review by Ulug et al suggests that smaller proportion of women are eligible for EVAR (or EVAS) compared with men. In addition to higher
operative mortality a significantly larger proportion of women are not offered intervention due to complexity of their lesion and risk benefit analysis of intervention, which means the true difference in aneurysm related mortality between men and women may be higher than what is observed by examining hospital statistics or vascular registries.  

Introduction of VSQIP has resulted in the development of standardised pre-assessment processes to assess patients’ fitness for EVAR and open repair of AAA. These processes which include cardiopulmonary exercise testing collect objective and reliable measures of patients’ cardiopulmonary reserve for vascular intervention. Examination of this data may identify subtle differences between men and women which may have not been obvious using standard data mining techniques.

Without understanding the chain of causality it is difficult to suggest specific strategies to improve mortality in women undergoing AAA repair. There are two on-going studies examining the value of screening for AAA in women. The first is the National Institute for Health Research trial of AAA screening of women (Female Aneurysm Screening Study), the other is the (Screening Women for abdominal aortic Aneurysms; SWAN study) which is examining the potential clinical benefits and cost-effectiveness of AAA screening in women. As these studies will collect follow up data on women with AAA they will be able to provide valuable prospectively collected data on AAA in women from which the reasons for the difference in mortality could be effectively addressed. This will take many years to accomplish.

Around the same time as the study described in Appendix-2, several similar data mining exercises were performed in order to answer pertinent questions regarding the
management of AAAs. Filipovic and Goldacre used Hospital Episode Statistics (HES) data from NHS in England and Wales (analogous to SMR-1 data collected by NHS in Scotland.) to examine trends in mortality (Case Fatality Rates: CFR) and hospital admission for AAA for elective abdominal aortic aneurysm repair in England and Wales. They observed that CFR following AAA repair in England and Wales was higher than expected. This finding became the main driver for VSQIP program to improve survival and outcomes following vascular surgery.

Holt et al examined the relationship between annual number of patients with AAA treated in each hospital (case volume) and outcome using similar data mining methodology and HES data collected between 2000 and 2005 in England. They repeated this data mining exercise for the fiscal years 2005 to 2007 using deterministic statistics. They found a close correlation between volume clinical activity (number of patients treated with primary diagnosis of intact AAA) and CFR. Long-term follow up revealed that the survival advantage form being treated in a high volume centre was sustained (Figures-47,48). It would have been impractical to any other research methodology to examine real life practice in a meaningful way. These findings were confirmed by a worldwide systematic review and became the main driver for centralizing vascular surgery services and contributed to the development of vascular surgery as a speciality in the United Kingdom.

Path analysis and sustainability of infra-inguinal vein grafts

Vein graft surveillance has remained controversial topic in vascular surgery for 20 years. whilst it is true that a significant proportion of vein grafts develop graft threatening stenotic lesions within the first year of the initial procedure, there is little evidence outside
observational and case control studies which are performed by enthusiasts of vein graft surveillance to suggest that duplex surveillance is associated with improved graft patency and improved limb salvage.\textsuperscript{242-248} Duplex surveillance is resource intensive and difficult to justify on the basis of cost, unless a large number of vein grafts (limbs) are being saved. Publications in Appendix 3 to 7 confirm the finding that many grafts go through vein graft surveillance cycle without exhibiting any abnormality on Duplex US.\textsuperscript{9,112} The nature of the problem is that patients / grafts under surveillance exhibit significant differences in risks of failure. Risk stratification can be used to focus ultrasound based surveillance on grafts with highest risk of failure in order to improve its efficiency.

A significant proportion of vein graft bypasses develop flow disturbance on duplex scanning at the time of the early post-operative surveillance examination. Several authors reported similar finding and suggested that an early duplex US examination is the most important assessment of the graft surveillance\textsuperscript{11,243-248} Apart from denoting an early threat to patency (if these lesions cause severe stenosis), early abnormalities also predict the natural history of the graft and outlook for the limb in the medium term. In the present study, vein grafts without flow anomalies at 6 weeks ran a benign course. The vast majority of grafts that occluded or required intervention exhibited significant anomalies by the time of the 6-week scan. It should prove possible to use this finding to select vein grafts at particular risk for duplex surveillance, thereby reducing the cost of duplex surveillance.\textsuperscript{99}

The fate of intermediate stenoses has been a matter for debate. This is a topic was not addressed by the VGST. VGST resolved the issue by having a very low threshold peak systolic velocity for defining clinically significant graft stenosis. This would remove intermediate
stenoses as a category and means their natural history would not be addressed. Mofidi et al.\textsuperscript{24}, reported that several intermediate stenoses and flow abnormalities did not progress or significantly improved during follow-up and therefore require surveillance rather than immediate intervention.\textsuperscript{244,248-251} Vesti et al. studied intermediate stenoses situated within the body of the graft that derived from valve cusps.\textsuperscript{250} They reported that over half of these lesions regressed without intervention over the follow-up period.\textsuperscript{250}

Autogenous vein grafts are the preferred conduits for infra-inguinal arterial reconstruction.\textsuperscript{252,253} In the absence of autogenous vein, the use of prosthetic grafts, whilst necessary, is associated with a higher incidence of graft occlusion, which can lead to loss of the affected limb.\textsuperscript{254} However, a commonly encountered problem is veins of a small diameter identified, which may be considered unsuitable for infra-inguinal arterial reconstruction. Previous authors have reported that the incidence of small-calibre vein during vein mapping is relatively high and, under these circumstances, contralateral long saphenous vein, short saphenous veins, or arm veins are sought as conduits for infra-inguinal vein graft bypass surgery.\textsuperscript{255}

Previous studies with in situ vein grafts have shown that there is a significant increase in diameter of the vein in the early postoperative scan.\textsuperscript{254} Fillinger et al.\textsuperscript{255}, reported that veins of small diameter respond more to shear stress, resulting in a greater increase in size than veins with large diameter. Conversely, large-diameter veins with a low initial shear stress showed a decrease in diameter over time.\textsuperscript{256} A similar but competing process of adaptation to arterial circulation is responsible for neointimal hyperplasia, which can result in the development of vein graft stenosis.\textsuperscript{106} Other gross and histological features of vein graft,
such as reduced vein compliance, smooth muscle hyperplasia, and inflammatory infiltrates, have been associated with the development of vein graft stenosis.\textsuperscript{98,99} Westerband et al. reported a close association between intimal and subintimal angiogenesis and thickness of intima in human stenotic vein grafts.\textsuperscript{244} All of these findings serve to remind that the quality of the venous conduit is important in predicting outcome of infra-inguinal bypass surgery.

A significant proportion of infra-inguinal vein graft bypasses develop graft stenosis, which threatens graft patency.\textsuperscript{248} Varty et al. reported that preoperative preoperative vein graft diameter is an independent risk factor for development of vein graft stenosis. In this study there was a significantly higher re-intervention rate for vein graft stenosis in bypasses which were formed using small-calibre vein grafts\textsuperscript{106}. Despite the fact that VGD was an independent risk factor for re-intervention, there was no significant difference in primary assisted patency or amputation rates between grafts created using vein with a minimum preoperative diameter $<3.5$ mm and those with larger diameter veins. This highlights the importance of vein graft surveillance and timely repair of vein graft stenosis in this population of patients.

Development of vein graft stenosis or occlusion is multifactorial in nature and the implications of vein graft occlusion vary depending on what was the initial indication for the bypass procedure. This would mean that a screening test to identify high risk grafts would also need to take account of multiple factors.

Classification and regression trees are a novel method of identifying vein grafts which are at risk of developing stenosis or occlusion. These grafts can be offered intensive surveillance
whilst grafts deemed to be low risk can be followed up clinically. There are well established standards for the development of clinical decision rules.\textsuperscript{255-261} It is very important that new predictive model such as the one discussed in appendix-6 (study-5) is validated using a new population of patients (Appendix-7, study6)\textsuperscript{262}. There are limitations to how much is extrapolated from any predictive model including the decision tree. In this and all recent studies on infra-inguinal vein graft surveillance, stenoses thought to threaten graft patency have undergone prophylactic revision. Therefore, the natural history of vein graft stenosis cannot accurately be determined, as all subsequent studies have a selected sub-group of patients and these selected groups may not be fully representative of the population.

Surgical revision of failing infra-inguinal vein grafts has been the gold standard in the management of vein graft stenosis. Techniques of vein patch angioplasty for short stenoses and interposition or jump graft repair for longer lesions have been associated with excellent medium and long term patency rates.\textsuperscript{263} Surgical repair of vein graft stenosis is an invasive procedure requiring general or regional anaesthesia in a population of patients with significant co-morbidity. It is also partly dependent on the availability of an appropriate length of an additional venous conduit. Revised vein grafts are themselves at risk of development of vein graft stenosis.\textsuperscript{263} Percutaneous transluminal angioplasty is an attractive option for the management of vein graft stenosis. Some years ago Berkowitz et al. reported on the results angioplasty as the primary treatment modality in 81% of failing infra-inguinal reversed vein grafts.\textsuperscript{264} They reported a 5-year primary assisted patency rate of 61%. Lesions in the proximal graft and proximal anastomosis taken as a group had significantly better patency than the mid-graft and distal anastomotic lesions.\textsuperscript{264} Subsequent authors have
reported similar results following vein graft angioplasty. Alexander et al. reported a significantly lower primary assisted patency after percutaneous transluminal angioplasty (PTA) for the treatment of vein graft stenosis, with failure rates of 31% at 6 months, 55% at 1 year, and 63% at 2 years.

A significant proportion of grafts, which had undergone angioplasty with a satisfactory technical result, develop re-stenosis, requiring secondary procedures. In the majority of patients, the secondary procedure was endovascular with a technical success rate approaching that of angioplasty for primary vein graft stenosis. Carlson et al. reported their experience of PTA of infra-inguinal vein graft stenoses in 36 patients. Their initial technical success rate was 91% and overall graft patency rate was 78% at 24 months.

Since Mofidi et al, the interest in endovascular treatment of vein graft stenosis has increased as the nuances related to this mode of treatment has become more understood. Risk of re-stenosis has led to the need for regular surveillance and repeat procedures, on the other hand unlike surgical repair endovascular procedures can be repeated without adding to the complexity of the procedure. Due to the fibrotic nature of vein graft stenosis some of these lesions do not respond to standard balloon angioplasty. In these circumstances, cutting balloon angioplasty has been used with some success. Short covered endovascular stents have been used with success in order to treat this population of patients with sustainable results.
Critical vein graft stenosis is a complex clinical problem. These series of publications have provided valuable and new insight into identification and treatment of vein graft stenosis. Ten years ago following the publication of VGST trial surveillance of infra-inguinal vein grafts was considered not clinically valuable or cost effective.\textsuperscript{101} Despite this vein graft surveillance has continued and there has been a paradigm shift in the management this condition using endovascular techniques.\textsuperscript{264-270} The recent insight gained has led to reappraisal of the processes involved in haemodynamic adaptation of vein grafts\textsuperscript{271} and how that relates to the treatment of graft stenosis.\textsuperscript{272}

**Artificial Neural Networks and assessment of degree of internal carotid stenosis**

Benefit from carotid endarterectomy in patients who have had recent ipsilateral neurological or ocular symptoms depends not only on the degree of carotid stenosis, but also on several other clinical characteristics such as delay to surgery after the presenting event, in addition to patient characteristics such as age, gender and co-morbidities such as diabetes, hypertension and ischaemic heart disease. Decision aids such as the Oxford Stroke Risk calculators utilize this information to individualise the risk-benefit calculus.\textsuperscript{128-130,273,274} Figure-55 illustrates the interplay between timing of surgery (one of these risk factors) and the degree of ICA stenosis in risk reduction afforded by carotid endarterectomy and highlights why knowing the exact degree of stenosis in a timely way is an important part of this risk reduction strategy.\textsuperscript{275}

Duplex ultrasound scanning is the premier modality of for assessment of degree of ICA stenosis\textsuperscript{276} owing to its non-invasive nature and widespread availability it has been used as a
screening tool for assessment of presence and degree of ICA stenosis. Carotid endarterectomy can be safely performed without preoperative angiography.\textsuperscript{133,277-278} Proliferation of duplex velocity criteria has resulted in lack of uniformity in the measurement of ICA stenosis using duplex imaging. This is in direct contrast to the standardized nature of angiographic assessment of ICA stenosis during the randomised controlled trials, which form the evidence-base for carotid endarterectomy.\textsuperscript{269,270}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure55.png}
\caption{Absolute risk reduction afforded by carotid endarterectomy in patient suffering from symptomatic ICA stenosis classified by degree of ipsilateral ICA stenosis and symptom latency (from Rothwell et al., Lancet 2004; 363(9413):915-24).\textsuperscript{128}}
\end{figure}

Initially the randomised controlled trials which form the evidence base for the management of symptomatic and asymptomatic carotid atherosclerosis identified different threshold values for ICA stenosis as representing the indication to perform carotid endarterectomy...
relating to different way the measurements are performed. Several previous authors have changed their velocity criteria in order to comply with these studies. Although Intersociety recommendations and reporting standards have been put forward on both sides of the Atlantic to standardise interpretation of Doppler waveforms by vascular technologists, differences between exist and remain clinically significant. Knowing the exact degree of stenosis in conjunction with preoperative clinical risk factors will influence decision making with regard to surgery, especially in patients with moderate ICA stenosis in whom the benefits from intervention may be marginal.

ANNs remain novel approach in the assessment of degree of ICA stenosis. Accurate measurement of the degree of ICA stenosis using velocity of blood flow contains an inherent degree of variability. This variability increases with progressively lower degrees of stenosis in the ICA. It would be impossible to account for this variability through conventional statistical models. ANNs separate background noise (variability) from information embedded in the data set. ANNs can be built into duplex devices. In this way, after selection of the appropriate sample volume by the operator, the machine could automatically calculate the duplex parameters for predicting the probability of threshold ICA stenosis and measure the exact degree of ICA stenosis. ANNs have an acceptable degree of accuracy in calculating the exact degree of ICA stenosis and represent a modest but significant improvement in classification of ICA stenosis into 10% bands. Carotid duplex criteria have been proven to be excellent tools in identification of threshold ICA stenosis at the clinically significant 70% level used for identification of surgical candidates amongst patients with symptomatic disease and 60% stenosis screening for asymptomatic carotid artery stenosis. However, with increasing use of decision support tools which require the
exact degree of ICA stenosis the information demanded from any modality assessing the degree if ICA stenosis has increased beyond just thresholding the degree of stenosis.

Other approaches such as multivariate regression have been used in order to assess the degree of ICA stenosis\textsuperscript{283,139} albeit with a moderate degree of correlation with angiographic controls whilst novel, these do not offer an improvement over duplex velocity criteria\textsuperscript{162} or an ANN.

Direct measurement of the residual lumen using a combination of β-mode ultrasound and colour flow imaging components of duplex investigation has been used to measure ICA stenosis. Despite the advances in β-mode ultrasonography has failed to gain widespread use.\textsuperscript{284,285} B-mode ultrasound and angiography measure the degree of vessel stenosis in different ways: duplex ultrasound assesses stenosis by percentage area reduction, whereas angiography measures the percentage reduction in luminal diameter. Despite close correlation, the two values are not interchangeable.

Previous authors have suggested that duplex velocity profiles may overestimate the degree of ICA stenosis in the presence of contralateral occlusion.\textsuperscript{286,287} This is difficult to adjust for when duplex velocity criteria are used alone. However, such an adjustment is possible in an ANN model in the form of another input variable (node). Study of Doppler wave form and the shape and pattern of the Doppler wave from within the sample volume (Spectral Analysis) can be used to analyse the degree of ICA stenosis. This provides subjective confirmation of presence of stenosis by a trained observer. ANNs can be used to objectively
analyse the Doppler wave from and accurately classify presence of ICA stenosis and occlusion.\textsuperscript{288-291}

Muller et al have reported a novel way of addressing the problem of assessing the degree of ICA stenosis by using combined modalities of Duplex US and CT angiography and developing a semi-automated method. This model combines two non-invasive modalities but has the downside of requiring intravenous contrast and exposure of the patient to ionising radiation. It also remains to be validated in an external population of patients.\textsuperscript{292} In absence of a single imaging modality approach for measurement of degree of ICA stenosis, multi-modality strategy can be utilized. A paper by Tholen et al examined the cost effectiveness of this approach using a decision tree model and like Muller they suggested that Duplex US (as a screening test) followed by CT angiogram and MRA as secondary modalities was the most cost effective approach.\textsuperscript{293}

This publication is unique as it high-lights the supportive role ANNs can play in predictive data-mining. Almost everybody who studies in the field of data mining is familiar with the use of ANNs as the primary data analysis model for data-mining\textsuperscript{294}. A PubMed\textsuperscript{®} search using the term “Artificial Neural Network” and “Data Mining” as Boolean operators yields 1817 publications. Rarely, if ever have ANNs been used to provide accurate data for use in another risk prediction model. In this case ANNs provide valuable information about the degree of ICA stenosis which is a key determinant for the Oxford Stroke Risk Calculator\textsuperscript{128}. Although artificial neural networks (ANNs) have been successfully applied in a wide range of machine learning applications, they are often regarded as black boxes which provide the answers without allowing the user to examine the processes by which ANN reaches the
answer. In this paper ANN plays a concise symbolic role which is easily explained and verifiable. It plays this role with versatility by producing results that are useful to the clinicians involved (degree of stenosis was provided in 10% bands, as exact degree of stenosis and as greater than 50% or 70% stenosis by the ANN model) in a verifiable way. This is why this publication is as clinically significant today as it has been in the last 10 years.

**Artificial Neural Networks and Prediction of outcome from oesophageal cancer**

The use of artificial neural network remains a relatively novel approach for prediction of outcome in patients with carcinomas of oesophagus and OG junction almost ten years after the publication of this study. As this study illustrates, intelligent data mining methods such as artificial neural networks provide an opportunity to improve prediction of outcomes following surgical treatment of cancers including cancers of oesophagus and oesophago-gastric junction. The implication being that targeted administration of adjuvant therapy such as radio or chemotherapy to patients with highest levels of risk would reduce the risk of recurrence in the high-risk group. Whilst the low risk patients would undergo surgery with curative intent. There are many instances where randomised controlled trials have proven the efficacy of neo-adjuvant chemo or radiotherapy or the combination of the two in order to improve the efficacy of the eventual surgical treatment of solid cancers.

Superior accuracy of ANN models is partly related to superiority of analytical model and partly to the increased number of input variables used to develop the ANN. Most clinical analytical models such as TNM staging are heavily weighted towards the histopathological
interpretation of the anatomical extent of disease. The ANN model in this study selected the appropriate input variables using sensitivity analysis. This by itself may be responsible for improved accuracy. A large-scale study by examining cancer specific 5-year survival in breast and colorectal carcinoma found that the ANN models constructed using TNM staging system alone significantly improved the predictive accuracy of TNM staging system.\textsuperscript{64, 295}

Despite the fact that ANNs are well suited to outcome prediction, to date their routine clinical use has been limited to the management of prostatic carcinoma.\textsuperscript{296} Some of the resistance to the use of ANNs in the clinical setting relates to the fact that their introduction as predictive models involves addition of new nomenclature and methodology, with which the clinicians are not familiar. In addition, clinical use of ANNs is tempered by the innate suspicion of a computer based system which has a relatively transparent process but cannot explain its reasoning in individual cases. There is a lot less reluctance in accepting the result of ANN analysis in dealing with problems such as genomics study gene expression in at risk populations\textsuperscript{157, 297}, structural and functional relationships between genomes or single nucleotide polymorphisms\textsuperscript{298} that due to their complexity defy conventional statistics or image analysis where data analysis is a simulation of a cognitive process performed by a trained observer.\textsuperscript{154, 155} In these situations, the approach towards ANN is pragmatic and result orientated and based on adequate validation of the model without attempts at understanding the reasoning on case-by-case basis.\textsuperscript{299, 300}
This is not to say that using ANN models for outcome prediction in patients with invasive cancer is not without its pitfalls. A complex ANN with a large number of input variables can be very accurate at identifying patterns within the training dataset. This can result in overfitting of the model, whereby the ANN identifies patterns peculiar to the training dataset that do not represent the larger population of patients. This is overcome by limiting the iterations of training and using adequately sized samples of patients for training and validation of the ANN. Model instability is another potential source of inaccuracy, where identical independently trained ANNs can have varying predictions for individual patients.

There are a variety of ways to minimise these risks, in this study, a panel of 10 ANNs trained on randomly selected subsections of data. The training and evaluation samples were randomly selected and in order to be representative of the patient population. A large number of potential input variables are often available in presence of limited number of training and evaluation. This is a potential source of model instability. In this study, variable selection was performed through the well established process of sensitivity analysis, which limited the number of input variables and selected the optimum combination of variables, which was associated with the most accurate predictions. Additional searches for prognostic variables for oesophageal carcinoma have identified cellular and histological markers of prognostic significance, this is a continuous process. These variables could easily be incorporated into the design of an ANN in order to improve its accuracy.

Any prediction model is as good as the quality of data used to develop and validate it. Missing and imbalanced or misclassified data would defeat the purpose of any prediction model. This is the main reason why the ISD of NHS in Scotland introduced the mandatory
National Minimum Core Dataset for Cancer in order to streamline the flow of information and standardize reporting practices. The use of this data supplanted by data from other hospital information systems enhanced the accuracy of ANN models.

The use of artificial neural networks as a predictive tool in vascular surgery has been limited. A number of years ago Turton et al used a simple 4 variable multi-layered perceptron to predict mortality rate from ruptured abdominal aortic aneurysm. They reported that ANN model was accurate at predicting in-hospital mortality in this cohort of patients. More recently Wise and his co-workers developed a similar ANN to predict mortality from RAAA. They compared the ANN model (Simple MLP model with 4 input variables) with multiple logistic regression and Glasgow Aneurysm Score and reported that ANN was more accurate than the commonly used Glasgow Aneurysm score or the logistic regression model. Although the two publications are 15 years apart, they both utilize the simplest ANN topography and the minimum number of input variables and, yet they report a high degree of accuracy.

An interesting use of neural networks in vascular surgery was their use in identifying the risks of re-intervention and mortality after Endovascular repair of abdominal aortic aneurysm. Karthikesalingam et. al. developed a relatively complex Bayesian Neural Network (19 input variables) using a large cohort of patients with a high degree of accuracy. Although the variables that Karthikesalingam utilized for developing their network included generic morphological information such as AAA diameter, sac volume, iliac diameter, aneurysm length, thrombus within the sac, neck length, angulation and diameter obtained by a trained observer, these values can be obtained automatically from the planning CT scan. The implication being that this investigation in addition to
being used for planning the EVAR procedure can provide assessment of risk of re-intervention or death in the first 5 years following endovascular repair of AAA.\textsuperscript{312}

Bayesian Networks differ from standard ANNs in that they map relationships between events and input variables in terms of probability of one event occurring in conjunction with another event as well as their mutual contribution to the final output variable. i.e. in a Bayesian network each node represents an event and edges represent probabilities of that event occurring (or being true). In standard ANN model although an individual node contributes to the predictive model by itself it does not have a discernible output. Standard Bayesian networks (such as the one used by Karthikesalingam) are unsupervised i.e. they do not require prior training. Bayesian networks are in more widespread use than standard multi-layered perceptions. For example, if your email account uses a spam filter the chances are that the basis for the classifier used in that filter is a Bayesian network.

Artificial Neural Networks are an effective tool for predicting outcomes in patients with oesophageal carcinoma and have had some tentative uses in predicting outcomes in vascular surgery. In management of cancer can incorporate multiple tumour-related and patient-related variables which are not currently part of the NHS minimal cancer dataset.\textsuperscript{313-315} In vascular surgery they can include direct imaging data through automated analysis of CT scans and Duplex images\textsuperscript{37} The information provided by these models may improve our ability to select appropriate and effective treatments.
Concluding Remarks

It was almost 20 years ago when Leroy Hood questioned the reactive way medicine approached health and disease and introduced the concept of precision medicine (P4-Medicine), healthcare which was predictive, preventative, personalized and participatory and focused on the needs of the individual patient rather than a cohort of patients with a similar condition(s) and accounts for the diversity of comorbidities and clinical needs.316 Such an approach starts from genomic and proteomic bases of ill health and builds a biological cross disciplinary culture in order to study the complex biological systems associated with health and disease.317 Although P4 medicine is some years away from implementation, biological systems theory on which P4 depends has been responsible for major advances in healthcare from the human genome project318 to control of Ebola and SARS where mathematical modelling developed clear strategies for success in preventing Pandemics.319

In healthcare as in business the concept of “Big Data” which describes comprehensive and systematic exploration of data derived from large cohorts of individuals has become not only a buzzword but as an established investigative paradigm320-323 With the advent of mobile computing and low cost of storage, big data will become more available and beyond the ability of individuals to comprehend, whilst advances in computational tools have provided the means by which this data can be analysed.324 Enthusiasts would like to believe that these complex models can replace clinicians who as individuals are prone to heuristics and cognitive biases in their decision making process. On the other hand, by removing human intervention modelling artefacts can develop and grow and analyse data without
regards to context. An integrated approach resulting in creation of decision support tools which are used by healthcare practitioners is a marriage of these two approaches which helps clinicians in areas of uncertainty. The Oxford stroke risk calculator is an example of this.133

Over the last 3 decades new data formats with increasing level of complexity have entered the field of data mining. These include two and 3-dimensional images, data clusters, text or multi-relational and multi-dimensional data. Many data mining methods are designed for collections of objects well-represented in rigid tabular formats. Traditional data mining tools tend to struggle to deal with complexities of life as it relates to multi-dimensionality of the data source. This data may be in the form of massive sets of unstructured information or as is more and more the case, objects which have natural representations such as protein structure or dynamics of protein interaction networks.

As illustrated by a number of publications discussed in this narrative, significant quantities of high-quality data are available from a number of sources. This mandates the need for merging data from multiple sources in as efficient manner as possible, while maximizing the accuracy of the result. This is described as the merge and purge problem. A number of paradigms have been developed in order to address this these include the use of human interaction if the size of data source permits, data clustering and the Sorted Neighbourhood method which are automated, but expensive.
Entropy-based data mining is an emerging methodology for use in this type of problems where complexity, noise and missing data introduce uncertainty and risk of modelling artefacts. In 1949 Claude Shannon introduced the information theory in which the concept of entropy relates to as a measure for the uncertainty in a message. The aim of this type of data mining is to discover knowledge that is not only accurate but also comprehensible for the users. This type of data mining is used for analysis of biomedical data such as electroencephalography (EEG), hormone secretion data, foetal heart rate monitoring, assessment of complexity measures of DNA sequences and hand writing recognition.

With rapid advances in the fields of bioinformatics and computational biology, healthcare and biomedical sciences are becoming data intensive disciplines where the clinicians are confronted with increased volume of diverse, multi-dimensional data needing analysis. In addition to being increasingly complex and diverse, the data available for analysis, may be weakly structured, available from different databases or contain noise. It may be imbalanced and missing elements. Nonetheless we rely on these data in order to assess the performance of clinicians, create safe pathways of care and make meaningful clinical decisions and even formulate healthcare policy. Our reliance on this level of evidence is likely to grow.

This narrative has highlighted opportunities and pitfalls that knowledge discovery offers for healthcare practitioners. This includes quantitative risk assessment of outcomes but is not
limited to it.\textsuperscript{339-341} The discipline of systems medicine is going to lead to a revolutionary transformation in the field of healthcare delivery. A transformation which will lead to increased understanding of health and disease and will have implications regarding how healthcare is delivered and will require considerable changes to how medical education and training are provided as well as how clinicians engage with patients in order to make healthcare a more effective process.
References:


12- Obenshain MK. Application of data mining techniques to healthcare data. *Infection Control* 2004; 25(08):690-5.


21- Chawla NV. Data mining for imbalanced datasets: An overview. In Data mining and knowledge discovery handbook 2009 (pp. 875-886). Springer US.


32- Mofidi R, Patil PV, Suttie SA, Ogston S, Parks RW. The Value of Procalcitonin at Predicting the Severity of Acute Pancreatitis and Development of Infected Pancreatic Necrosis. Surgery 2009; Volume 146(1); 72-81.


45- Weiss, S, Indurkhya N. Predictive Data Mining. *Morgan Kaufmann (Elsevier)Publishers 1997*


51- Müller E, Günemann S, Färber I, Seidl T. Discovering multiple clustering solutions: Grouping objects in different views of the data. In *Data Engineering (ICDE), 2012 IEEE 28th International Conference on 2012 (pp. 1207-1210).* IEEE.


The Evidence for Management of Abdominal Aortic Aneurysms: Lessons Learned from Randomised Controlled Trials

Reza Mofidi¹ and Stuart A. Suttie²

¹James Cook University Hospital, Middlesbrough
²University of Dundee, Dundee, United Kingdom

1. Introduction

Abdominal aortic aneurysm (AAA) is a common life threatening condition in the western world. In England and Wales alone, over 2500 patients present to hospital with rupture of AAA annually, of whom over two thirds die of their condition¹. The best treatment for AAA is elective repair of pre-symptomatic abdominal aortic aneurysms. Such a therapeutic strategy depends on effective identification of patients with AAA and the subgroup of patients in whom there is a real risk of aneurysm rupture. As the vast majority of patients with AAAs are asymptomatic, timely identification of AAA may be achieved through targeted screening of the at risk populations. Over the last two decades longitudinal studies of patients with smaller AAAs have provided insights into the timing of AAA repair and the need for and frequency of ultrasound surveillance if an expectant management strategy is followed. This chapter discusses the available evidence for screening for AAA as well as all the other measures which have helped to optimise therapeutic strategies in the management of patients with AAA throughout the patients’ journey from the initial diagnosis to the eventual repair of AAA.

2. Targeted screening for AAA

In the past 40 years with the advent and generalised use of abdominal ultrasonography there has been an accurate, cheap and non invasive tool for the diagnosis of abdominal aortic aneurysms. Abdominal ultrasonography has been found to be an accurate and reproducible modality in measuring the dimensions of AAA. This has led to the concept of its use for screening of at risk populations. In the last 20 years there have been four population based randomised controlled trials which have assessed the value of targeted screening in reducing mortality from abdominal aortic aneurysms in the unselected elderly male population²⁻⁵. These trials which have been undertaken in Chichester (England)², England (MASS trial)³, Viborg County (Denmark)⁴ and the city of Perth and suburbs (Western Australia)⁵ have together recruited over 120,000 subjects. All of these studies have reported on long term (over 10 years) follow up. Using the predefined criteria set by the US Preventative Screening Task Force USPSTF⁶ the MASS trial has been classified as good with
the other three trials classified as fair i.e. not meeting all the criteria but judged to have no fatal flaws.

The Chichester trial was the first to assess the value of screening for AAA in the at risk population. It was also unique as it included women as well as men. It identified all men and women aged between 65 and 80 years of age from 9 general practices in the catchment area of St Richard’s hospital in Chichester between 1988 and 1991. The subjects were randomised to undergo a single screening ultrasound (US) or a control group who were followed up. AAA rupture rates, aneurysm related mortality, and overall mortality was compared between the two groups. Upon identification of AAA the therapeutic strategy for AAAs with maximum diameters between 30-44mm was once yearly surveillance US, AAAs between 44 and 59mm underwent 3 monthly ultrasound scans, whilst aneurysms greater than 60mm in diameter were considered for repair. Overall 6040 men were randomised, the authors reported a significant reduction in aneurysm related mortality which has been maintained over 15 years. However, to date this study has demonstrated no difference in the all cause mortality between the two groups. The Chichester trial has been criticized for its relative small size, a relatively high aneurysm diameter threshold for repair and including 75-80 year old patients in whom the benefits of screening are marginal. In addition 27% of subjects who were invited for screening refused to participate thereby diluting the benefits of screening. Despite these criticisms the Chichester study remains a landmark as it demonstrated the feasibility of US screening for AAA and its potential value and remains a blue print for other aneurysm screening studies. This study identified a low but nonetheless troubling rate of AAA rupture in patients who had a non aneurysmal aorta on the first screening study. A population based screening study in Gloucester demonstrated that 2.2% of men aged 65-73 years have a maximal aortic diameter of 2.5 to 2.9 mm and suggested that this group of patients should undergo repeat US scanning at 5 yearly intervals.

The second RCT to study the value of population based screening for AAA was carried out in Viborg County of Denmark. In 1994 all men aged between 65 and 74 were randomised to either undergo a single screening US or the control group. In all 12639 patients were randomised. This study reported a 66-percent reduction in the aneurysm related mortality which has been maintained over 14-years. In addition they reported a 2-percent reduction in overall mortality after long term follow-up which did not reach significance. The Western Australia population based screening was a study of similar design. It randomised 41000 men between the ages of 65 and 85 years to a single US screening and a control groups. They reported no difference in aneurysm outcomes in the full study population but when the analysis was restricted to 65-74 year old men they reported a significant reduction in aneurysm related mortality after 5 years of follow-up. Long term follow-up results of this study have not been published as a separate publication to date, however in a reply to a correspondence by Lederle, Norman and Lindholt did report a surprisingly high, 3-percent reduction in overall mortality in the restricted (65-74 year old) patient population after 10 years of follow-up from the Western Australia trial which was statistically significant.

The MASS trial which was a population based screening RCT for men aged between 65 and 74 years of age included 4 screening centres in the United Kingdom. This study randomised 67770 patients again to single screening ultrasound or a control group and was designed to study cost effectiveness of screening in addition to reductions aneurysm related and overall mortality. This study reported a 48-percent relative risk reduction in aneurysm related
The Evidence for Management of Abdominal Aortic Aneurysms: Lessons Learned from Randomised Controlled Trials

Mortality as a result of screening. This benefit was present at 4 years\textsuperscript{14} and was maintained at 10 years (Figure-1)\textsuperscript{3}. There was a reduced AAA rupture rate in the patients who were invited for screening. Most of these ruptures occurred in patients who were excluded from the potential benefits of screening, such as patients who refused or did not attend screening, patients who were lost to follow-up and those who either refused or deemed not fit for surgery\textsuperscript{3}. The MASS trial also reported a small rate of AAA rupture in patients who did not have an AAA on the screening scan, this rate was reported as 3 per 10,000 person years after 10 years of follow up\textsuperscript{3}.

![Cumulative deaths related to abdominal aortic aneurysm](image)

Fig. 1. Cumulative deaths related to abdominal aortic aneurysm, by time since randomisation (MASS Trial)\textsuperscript{3}. From: Thompson SG, Ashton HA, Gao L, Scott RA and Multicentre Aneurysm Screening Study Group, Screening men for abdominal aortic aneurysm: 10 year mortality and cost effectiveness results from the randomised Multicentre Aneurysm Screening Study, BMJ 2009; 338: b2307.

In addition to the above RCTs a number of systematic reviews and meta-analyses have attempted to assess the value of population based screening in the medium and long term. Cosford and Leng in a Cochrane systematic review reported that there was significant evidence of reduction in aneurysm related mortality from AAA in men aged 65 to 80 years who undergo population based ultrasound screening, but no significant reduction in all cause mortality\textsuperscript{16}. This review was based on the 3-5 year follow up data from the above RCTs. Subsequent to this Norman and Lindholt published a meta-analysis which showed that population based AAA screening after 7-15 years of follow up resulted in a reduction of both AAA and all cause mortality\textsuperscript{17}. Their findings were contested as the reported 3-percent all cause mortality reduction was larger than what was expected by an approximately 50-percent reduction in aneurysm related mortality, bearing in mind that the mortality from AAA in the patient population is reported to be between 1.1 to 3-percent\textsuperscript{18}. 
Takagi et al. conducted a further meta-analysis of US screening in the male population over the age of 65 years using long term 10 to 15 year follow up data from the RCTs. They reported an absolute risk reduction in aneurysm related mortality of 4 per 1000 subjects screened (Figure-2). They also revealed a strong trend towards a significant reduction in all cause mortality. The latter finding was surprising for the reasons mentioned already. The authors hypothesized that screening may coincide with the asymptomatic at risk population for cardiovascular disease coming in contact with health care professionals and becoming aware of smoking risk, their blood pressure etc. The resultant reduction in cardiovascular risk factors may be in part responsible for additional reduction in all cause mortality. Such a hypothesis opens the door to the possibility of risk factor alteration and institution of secondary prevention measures such as commencement of anti-platelet agents and statin therapy during screening programmes thereby increasing the value of the screening.

Fig. 2. Forrest Plot of illustrating the reduction in aneurysm related mortality (A) and the trend towards a reduction in overall mortality (B) as a result of population based screening of men between the ages of 65 and 80 years after 10 years of follow up. From: Takagi H, Goto SN, Matsui M, Manabe H, Umemoto T. A further meta-analysis of population-based screening for abdominal aortic aneurysm. J Vasc Surg. 2010; 52(4):1103-8.

Cost effectiveness of a population based screening programme is calculated by measuring the costs of ultrasound screening as well as the extra procedures and surveillance that is required for the screen identified AAA and subtracting them from the costs of treating ruptured AAA. It is expressed in cost per life year gained. As the survival advantage in terms of life year gained continues to increase with time, the cost effectiveness of screening continues to improve. A comprehensive analysis of costs of screening was performed by the MASS trial participants. They calculated the cost per life year gained to be £41,000 after 4 years, £14,000 after 7 years and £7600 after 10 years. Using the estimated life span of men aged 65 years the cost per life year gained is estimated to be in the region of £2300, which is well below the guideline figure of £25,000 which is considered acceptable for the adaptation of new medical technologies and interventions in the National Health Service of the United Kingdom.

Lindholt et al. also performed a comprehensive cost analysis of population based AAA screening using data obtained from the Viborg trial. They reported cost per Quality
Adjusted Life Year (QALY) gained as a result of screening to be €179 albeit with relatively wide 95% confidence intervals (€-4083 to €4682). Both of these values for costs of screening are much lower than the cost analysis carried out by the USPSTF using primarily economic modelling in 2003 and suggest that population based AAA screening in men is more cost effective than the initial assessments suggested. The role of screening for AAA in women remains controversial. To date there is no evidence that screening for AAA in an unselected population of women is associated with a reduction even in aneurysm-related mortality. Scott and colleagues conducted the only RCT (Chichester trial) which studied the value of screening in women over the age of 65 in an unselected population (n=9342). They reported the prevalence of AAA in women to be 1.3 percent, with other authors reporting a similar rate of 0.7–1.3 percent in unselected populations. Scott et al. did not demonstrate a difference in rupture rates between the women randomized to screening and control populations of women at 5- and 10-year follow-up. They concluded that screening for women was neither clinically indicated nor economically viable. This study was limited by high rate of non attendance of women for AAA screening which ranged between 27 and 42-percent depending on patients age. They screened an unselected population of women without consideration of risk factors for aneurysm disease and fitness for repair; consequently a large proportion of women who were found to have an AAA did not undergo aneurysm repair. The UK Small Aneurysm Trial revealed that female sex was an independent risk factor for AAA rupture; the rupture rate in women was three times higher than that in men, despite a smaller initial AP diameter. Furthermore, mean AP diameter preceding rupture was significantly lower in women than men. A number of other authors have reported a higher growth and rupture rate of AAA in women. A Finnish community-based follow-up study reported that the aortic diameter was less than 5.5 cm in 24 per cent of women with a ruptured AAA, compared with only 5 per cent of men. In light of these findings the 6 cm cut off value for repair of AAA in Chichester trial may have been too large to prevent aneurysm rupture in a proportion of screened women thereby reducing the value of screening in women.

For screening to be effective in reducing aneurysm-related mortality in women, it will need to be limited to high-risk women who are fit to undergo aneurysm repair. There is increasing evidence that women with atherosclerotic disease are at significantly higher risk of developing AAA. Derubertis and colleagues reported that the prevalence of AAA in women with multiple (more than three) atherosclerotic risk factors was 6.4 per cent. When these findings are considered in conjunction with the increased growth rates of AAA and higher aneurysm rupture rate in women, screening in women with multiple risk factors for AAA may become clinically and economically viable.

3. Optimum therapeutic strategy for small AAAs
Abdominal aortic aneurysms are treated in order to prevent rupture and the associated mortality. Aneurysm treatment has its own associated morbidity and mortality. Open surgical repair is an invasive procedure which is tolerated poorly by the subgroup of patients with multiple medical co-morbidities. Even endovascular repair cannot be accomplished without an obligatory complication rate as a result of the initial deployment of the stent graft, in addition to which a proportion of patients require secondary procedures necessary to address complications such as endoleaks, device migration and stent thrombosis requiring long term close surveillance. A small proportion of patients...
who have undergone endovascular repair (EVAR) succumb to rupture. Therefore the natural history of the AAA needs to be balanced against the risk associated with treatment. Aneurysm diameter is one variable which has been consistently associated with the risk of rupture and has therefore been used to stratify patients into risk categories which decides whether US based surveillance or intervention is required to repair the aneurysm. In patients who are entered into surveillance programmes the maximum diameter of the aneurysm is used to decide on the frequency of scanning. In case of aneurysms greater than 5.5 cm there is consensus that risk of rupture mandates repair if the patient is fit to undergo the procedure. In the case of aneurysms less than 4.0 cm in diameter, most clinicians agree on a watchful waiting approach. The evidence for the optimum therapeutic strategy in the mid-sized aortic aneurysms (maximum diameter between 4.0 to 5.5 cm in diameter) has been strengthened by a number of randomised controlled trials in the last 20 years which have consolidated the modern management of AAA.

The UK small aneurysm trial (UKSAT) was a multicentre RCT which randomised 1090 patients, who were diagnosed as having an AAA with maximum AP diameter of 4.0 to 5.5 cm and were deemed fit to undergo an open repair of AAA to either immediate open repair or 3 monthly ultrasound surveillance. They reported the rupture rate of these AAA in the surveillance group to be in the 1-percent per year. They did not find any significant difference in aneurysm related or all cause mortality between the two groups after a follow up period of 7 years (Figure-3). During the follow up period over two thirds of patients who were randomised to surveillance had undergone repair of their aneurysms based on clinical grounds. Long term follow up data from the small aneurysm trial has confirmed the initial findings of the UKSAT.

![Figure 3. Kaplan-Meier survival curves comparing survival of patients with small abdominal aortic aneurysms randomised to ultrasound surveillance and early surgery from UK small aneurysm trial.](image.png)

A number of years after the publication of the UKSAT, the Veterans Affairs Cooperative Study group published the Aneurysm Detection and Management ADAM study\textsuperscript{39}. This study involved screening of 126,196 veterans aged between 50 and 79 years of age for AAA with a single abdominal US. Those with AAA measuring 4.0 to 5.4 cm in diameter were offered entry to the trial. In all, 1136 subjects were randomly assigned to undergo early elective repair or ultrasound surveillance. Annualized rupture rate in the surveillance arm of the study was 0.6-percent, with no difference in aneurysm related and overall mortality between the two arms of the study\textsuperscript{39}. In this study as in UKSAT the majority of patients in the surveillance arm of the study had undergone elective repair after 8 years of follow up based on clinical grounds (symptomatic aneurysm, growth to greater than 5.5cm in diameter or rapid expansion by greater than 1 cm per year)\textsuperscript{39}. Completion of these two landmark trials which utilised open elective repair coincided with the advent and generalised use of endovascular repair as a primary modality treatment of AAA. This resulted in some authors questioning the validity of these landmark trials in the era of endovascular repair and suggested that as endovascular repair can be performed with significantly lower peri-procedural morbidity and mortality a policy of surveillance for smaller AAAs should be examined against endovascular repair.

To date two randomised controlled trials (PIVOTAL\textsuperscript{40} and CAESAR\textsuperscript{41}) have been conducted to compare early endovascular repair of small AAAs with ultrasound surveillance. The prerequisite for both studies was that the patients which were randomised had AAAs which were anatomically suitable for endovascular repair. The PIVOTAL trial which was published in 2010, randomised 728 patients with AAAs measuring 40 to 50 mm in diameter to ultrasound based surveillance or early endovascular repair\textsuperscript{40}. The mean duration of follow up was 20 months (+/-12 months) they found no difference in all cause or aneurysm related mortality between the two groups\textsuperscript{40}. At the end of the relatively short follow up duration almost one third of patients who were in the surveillance group had undergone an aneurysm repair based on clinical grounds\textsuperscript{40}. The other study of a similar design was the CAESAR trial which randomised 360 patients with AAAs measuring between 40 and 54 mm to early endovascular repair or a watchful waiting strategy. After 54 months of follow up there was no significant difference in rupture rates, aneurysm related and overall mortality between the two groups (Figure-4). This study revealed that the probability of the patients in the surveillance arm of the study requiring delayed repair based on clinical grounds during the duration of follow up was 60-percent\textsuperscript{41}. In addition they reported that 16.4-percent of aneurysms which upon entry into the trial were suitable for endovascular repair will be no longer suitable for EVAR after 36 months\textsuperscript{41}. A constant finding in these trials has been that a significant proportion of AAAs under ultrasonographic surveillance come to require repair within the duration of the study\textsuperscript{26,39}. This, taken together with the low but present annual risk of rupture has lead to differing interpretations of the results of these trials with some authors still advocating in favour of early repair of small AAA using the justification that a policy of early EVAR is as safe as a policy of US Surveillance\textsuperscript{42}. To date there is no objective data to recommend either open or endovascular repair of smaller AAAs over a policy of watchful waiting and US surveillance. A policy of early EVAR is associated with a risk of early and delayed complications and a need for secondary procedures, thus mandating the need for close surveillance in patients who undergo early EVAR. It is therefore unlikely that there will be an economic justification for early endovascular repair.
Aneurysmal Disease of the Thoracic and Abdominal Aorta

Fig. 4. Kaplan–Meier estimates of survival at 54 months from time of randomisation in EVAR versus Surveillance groups. $P = 0.6$. Numbers at risk are shown. CAESAR trial. From:Cao P; DeRango P, Verzini F, Parlani G et al. Comparison of surveillance vs Aortic Endografting for Small Aneurysm Repair (CAESAR) trial: results of a randomised controlled trial. Eur J Vasc Endovasc Surg. 2011; 41(1): 13-25.

4. Open versus endovascular repair of AAA

Ever since its inception, EVAR has offered the promise of reducing the perioperative morbidity and mortality which has been associated with open elective repair. By the end of last century, data from EVAR registries such as RETA and EUROSTAR suggested that endovascular repair, although safe was associated with an immediate complication rate in addition to events such as endoleak and device migration which mandate lifelong surveillance and in a group of patients re-intervention. As with any new or emerging technology or intervention the case for primacy of EVAR over open repair in terms of perioperative mortality rate, post operative complications and cost effectiveness needs to be made using good quality evidence. A number of trials with a similar design have been commissioned in order to compare the outcomes following EVAR and open repair of AAA in patients who are anatomically suitable to undergo endovascular repair and fit to undergo open repair. These include the Dutch Randomised Endovascular Aneurysm Management (DREAM) trial, EVAR-1 Trial (United Kingdom), ACE trial (France) and Open Versus Endovascular Repair (OVER) of abdominal aortic aneurysms trial (United States). The DREAM trial which was the first to report its results enrolled 351 patients between November 2000 and December 2003 from 24 centres in the Netherlands and 4 centres in Belgium. This study focused on short term combined mortality and morbidity outcomes.
reported a significantly lower operative mortality and severe complication rates in the EVAR group compared to the patients who had been randomised to open repair. At 2 years follow up aneurysm related mortality following EVAR was still significantly lower than open repair (2.1% versus 5.7%) however after 2 years of follow up there was no significant difference in the overall survival rates or freedom from moderate to severe complications between the two groups. The conclusions drawn from this trial was that there was a significant reduction in early morbidity and mortality following EVAR compared to open aneurysm repair but this difference is not sustained past 2 years\textsuperscript{45,46}.

EVAR-1 trial was a multicentre RCT which was conducted in 37 hospitals in the UK. It randomised 1252 patients with large AAA to either open or endovascular repair. Unlike the DREAM trial, EVAR-1 was designed to perform a comparison of long term survival, graft durability, quality of life and hospital costs associated with open repair and EVAR in addition to comparing short term mortality and morbidity between the two groups\textsuperscript{47}. They reported a significantly lower in perioperative morbidity and mortality following EVAR. Four years after randomisation, all cause mortality was similar between the two groups, although there was a persistent reduction in aneurysm related mortality in the EVAR group,(Figure-5)\textsuperscript{47}. After 12 months there was no difference in quality of life scores between the two groups with a greater number of complications and re-interventions at 4 years in the EVAR arm of the study. The hospital costs of EVAR were 25-percent higher than open repair\textsuperscript{47}.

Fig. 5. EVAR-1 Kaplan-Meier survival curves comparing aneurysm related and overall mortality between patients who have been randomised to open elective and endovascular (EVAR) repair of AAA (EVAR-1 trial)\textsuperscript{47}. From: EVAR trial participants. Endovascular aneurysm repair versus open repair in patients with abdominal aortic aneurysm (EVAR trial 1): randomised controlled trial. Lancet 2005; 365(9478): 2179-86.

The OVER trial is a RCT which included 42 Veterans Affairs medical centres in the United States. It randomised 881 patients who had AAA with a greater than 50 mm in maximal diameter, an iliac aneurysm greater than 30mm in diameter or rapid sac expansion, to elective open repair or EVAR. The preliminary results from this study indicated that the
EVAR group had significantly lower 30-day mortality as well as all cause mortality\textsuperscript{49}. After a mean follow up of 1.8 years the complication rate was not significantly different between the two groups nor was the secondary reintervention rate. As in the DREAM trial, the reintervention following EVAR was mainly due to a device related complications whereas the commonest reason for reintervention following open repair was for incisional hernia\textsuperscript{46,49}. Early results from the ACE trial suggest similar early mortality benefit following endovascular repair which is lost after medium term follow up\textsuperscript{48}.

Some subgroups of patients such as those who have significant co-morbidities such as cardiovascular or respiratory disease, octogenarians and women with AAA, require an individualised approach and revised criteria for the management of AAA. From its inception EVAR has provided the promise of repairing AAA in patients in whom open repair poses a high risk. Therefore armed with the knowledge that smaller AAAs are best managed by a policy of watchful waiting, EVAR appeared to be an ideal modality for the management of patients with larger AAAs which are anatomically suitable for endovascular repair, have a reasonable predicted longevity but are unfit to undergo open repair. The EVAR-2 trial was designed to answer this question. EVAR-2 trial was a randomised controlled trial of 338 patients who had an AAA with a maximum diameter of greater than 5.5cm and their aneurysm morphology was anatomically suitable for EVAR, but were medically unsuitable to undergo open repair. Primary endpoint was all-cause mortality, with secondary endpoints of aneurysm-related mortality, health-related quality of life, postoperative complications, and hospital costs\textsuperscript{50}.

The 30-day operative mortality in the EVAR group was 9.0-percent and the no intervention group had an annual rupture rate of 9.0-percent per year. Aneurysm related mortality in the patient population was 13-percent and all cause mortality after 4 years of follow up was 64-percent\textsuperscript{50}.

![Fig. 6. Kaplan-Meier curves comparing aneurysm related and overall mortality between patients who have been randomised to EVAR and no intervention group (EVAR-2 trial)\textsuperscript{30}.

There was no significant difference in all-cause mortality between the EVAR group and the no intervention group (hazard ratio 1.21, 95% CI 0.87–1.69). There was no difference in aneurysm-related mortality (Figure-6)⁵⁰. A policy of early endovascular repair was significantly more expensive than expectant management and was associated with a higher complication and reintervention rate. There was no difference in quality of life scores between the two arms of the study⁵⁰. Therefore the conclusion drawn by the authors was that this population of patients are best served by conservative treatment. Clearly the design of such a study provides one difficulty and that is the definition of not fit for open AAA repair is subject to clinical opinion and may be related to factors that do not affect patient’s longevity. The other group of patients are those with one organ morbidity such as respiratory disease or borderline medical fitness, who have a large AAA and favourable anatomy for endovascular repair. Therefore clinical judgement is exercised in the application of results of EVAR-2 trial.

5. Medical treatment of patients with AAA

In addition to risk of growth and rupture, patients with AAA are at risk from other cardiovascular events by the virtue of their age, medical co-morbidities and male preponderance of AAA. Medical management of patients with known AAA follows two parallel but different aims, reducing cardiovascular event rates perioperatively and during follow up in addition to aneurysm specific therapy which is aimed at slowing aneurysm growth and reducing the risk of rupture⁵¹-⁵³. Hyperlipidaemia, a known modifiable risk factor in the development of cardio-vascular disease, can be treated with the use of drugs such as the statins (3-hydroxyl-3-methylglutaryl coenzyme A reductase inhibitors). Patients with AAA are known to be at high risk of cardio-vascular disease as well as increased risk of cardio-vascular complications following AAA repair⁵⁴. Statin therapy has been associated with improved survival due to decreased risk of cardio-vascular complications, in both open and endovascular repair⁵⁴-⁵⁸. Although the primary mechanism of statins is in reducing low density lipoproteins and total cholesterol levels along with increasing levels of high density lipoproteins, other protective non lipid mechanism may be at work. These so called pleiotropic effects describe a diversity of cellular events which have an effect on several components of the arterial wall, including: endothelial cells; smooth muscle cells; platelet function, monocytes and macrophages, which together help to modify the inflammatory process in the vessel wall. Statins have been shown to be beneficial in the secondary prevention of coronary heart disease even in those patients with normal lipid profiles⁵⁹-⁶⁰.

Matrix Metallo Proteinase-9 (MMP-9) expression is closely linked to aneurysm formation in animal models. In vitro experiments have shown that addition of Cerivstatins to human organ cultures from AAA reduces tissue levels of both total and active MMP-9 in a concentration dependent manner. Evans et al reported significantly reduced MMP-9 levels in excised tissue obtained from the aneurysm sac at the time of the aneurysm repair in patients who had been started on statins 3-weeks preoperatively compared with controls⁵⁹. Schouten et al monitored 150 patients with small AAAs for 12 months and reported a reduction in the aneurysm expansion rate in patients receiving statin therapy⁶⁰. In an observational study of 130 patients under surveillance, Sukhija reported no aneurysm expansion in 75 patients who were on statin therapy over a 2 year follow up period⁶¹. Schlosser et al in an analysis of the results of a large observational cohort study which
involved 5057 patients with vascular disease (Second Manifestation of ARTerial disease (SMART) study) and included 230 patients with small AAA revealed an independent association between statin therapy and reduced aneurysm growth rate. This reduced growth and rupture rates were independent of serum lipid values\(^{62,63}\).

Over the years there has been some interest in \(\beta\)-blockers, both to slow the growth rate of AAA and to reduce perioperative morbidity form cardiovascular events. The benefit was postulated partly due to their haemodynamic properties and partly due to the effect of \(\beta\)-blockers on matrix proteins. In a trial reported by Lindholt and colleagues the use of Propranolol did not reduce the rate of expansion of AAA, admittedly in the treatment arm of the study the compliance was poor with only 22-percent continuing on Propranolol by 2-years\(^{64}\). Another trial which was carried out in Canada came to a similar conclusion owing to poor patient compliance in the treatment arm of the study\(^{65}\).

In the last 15 years there has been significant interest in using peri-operative \(\beta\)-blockade as a means of increasing myocardial oxygen delivery thereby reducing the risk of perioperative myocardial infarction and death. Mangano et al randomised 200 patients who were undergoing major elective non-cardiac surgery to either receive Atenolol or placebo. This was started before the induction of anaesthesia. Patients with evidence of congestive cardiac failure, systolic blood pressure of less than 100mmHg or pulse rate of less than 55 beats/minute, 3\(^{rd}\) degree heart block or broncho-spasm were excluded. This treatment was continued for 6 months postoperatively. They reported a significant reduction in cardiovascular event rate and death from cardiac causes\(^{66}\).

Poldermans and colleagues performed a similar study in patients undergoing elective aneurysm or infrainguinal arterial reconstruction. They screened 1351 patients for cardiac disease using Dobutamine stress testing, 173 patients had a positive test of whom 59 were randomised to receive Bisoprolol and 53 placebo\(^{67}\). They also reported a significant reduction in non fatal cardiac events as well as cardiac death. In these patients \(\beta\)-blockade was started at least a week in advance of the operation and they were screened for bradycardia and hypotension preoperatively\(^{67}\).

POISE was a large international randomised controlled trial of the use of extended release Metoprolol in patients undergoing non-cardiac surgery, the study randomised 8351 patients to either receive Metoprolol or placebo which was started 2-4 hrs before surgery and continued for 30 days. They reported a significantly reduced risk of myocardial infarction in the Metoprolol group but at the expense of higher mortality and stroke rate in the treatment arm of the study\(^{68}\). Similarly, Yang et al randomised such patients undergoing major vascular surgery, not already \(\beta\)-blocked, to dose adjusted Metoprolol or placebo 2 hours prior to surgery and until discharge or maximum of 5 post-operative days, and found no protective effects of \(\beta\)-blockade in terms of 30 day myocardial infarction and death rates\(^{69}\). \(\beta\)-blockade did result in significantly more episodes of bradycardia and hypotension. In light of these findings the American Heart Association guidelines regarding perioperative \(\beta\)-blocker therapy in patients undergoing non cardiac surgery have been altered to be more cautious and circumspective (Table-\()\(^{70}\).

In a large observational study, Hackham et al have shown that the use of Angiotensin Converting Enzyme Inhibitor (ACE\(_\text{i}\)) therapy taken 3-12 months prior to data analysis significantly reduced the risk of rupture from AAA, independently of blood pressure\(^{71}\). This data was obtained from a large administrative database of 3379 patients with ruptured and 11947 with non ruptured AAA. Other anti-hypertensive medications had no such effect \(^{71}\). Interestingly, patients who had stopped ACE\(_\text{i}\) therapy prior to admission were more likely
to present with ruptured AAA. The effect of ACE inhibitor on expansion of AAA is still equivocal, with some studies demonstrating no protective effect of ACE inhibitor therapy. Thompson et al. in a recent observational study of 1269 patients with small AAA who were followed up for a mean of 3.4 years, reported a significant reduction in aneurysm growth rate as a result of ACE inhibitor therapy. The follow up data from UK small aneurysm trial does not support the above finding.

Infection with Chlamydia pneumonia has been postulated as a risk factor for AAA expansion, as the organism has been isolated from atherosclerotic plaque and the walls of AAA. Three small trials have aimed to elucidate the effect of the antibiotics Doxycycline and Roxithromycin in AAA growth, two of which have shown reduced aortic expansion associated with treatment, whilst another one by Baxter and colleagues showed no effect of doxycycline on aortic diameter. These three trials were limited by their small numbers. In addition administration of Doxycycline has been shown to suppress MMP-9 in both human and animal studies, suggesting that the reduction in aneurysm expansion rate with administration Doxycycline may be mediated through a mechanism which is independent from treatment of Chlamydia pneumonia infection.

To date there is no conclusive evidence that any medical therapy is associated with a reduction in aneurysm growth or risk of rupture. However diagnosis of AAA provides a forum for instituting appropriate secondary prevention therapies, which will reduce morbidity and mortality in the peri-operative period as well reduce long term cardiovascular risk. There is some evidence that instituting some of these treatments such as statin therapy, ACE inhibitors may well have an effect on aneurysm growth and rupture rates.

6. References


Influence of sex on expansion rate of abdominal aortic aneurysms


Department of Vascular Surgery, Royal Infirmary of Edinburgh, Old Dalkeith Road, Edinburgh EH16 4SA, UK
Correspondence to: Mr R. T. A. Chalmers. (e-mail: rod.chalmers@luht.scot.nhs.uk)

Background: The UK Small Aneurysm Trial suggested that female sex is an independent risk factor for rupture of abdominal aortic aneurysm (AAA). This study assessed the effect of sex on the growth rate of AAA.

Methods: Between January 1985 and August 2005 all patients who were referred to the Royal Infirmary of Edinburgh with an AAA who were not considered for early aneurysm repair were assessed by serial abdominal ultrasonography. Maximum anteroposterior and transverse diameters of the AAAs were measured.

Results: A total of 1255 patients (824 men and 431 women) were followed up for a median of 30 (range 6–185) months. A median of six examinations (range 2–37) was performed for each patient. Median diameter on initial examination was 41 (range 25–83) mm. Median growth rate overall was 2.79 (range −4.80–37.02) mm per year. Median growth rate of AAA was significantly greater in women than men (3.67 (range −1.2–37.02) versus 2.03 (range −4.80–21.00) mm per year; \( P < 0.01 \)). Weighted linear regression analysis revealed that large initial anteroposterior AAA diameter and female sex were significant predictors of faster aneurysm growth rate (\( P < 0.001 \) and \( P = 0.006 \) respectively).

Conclusion: The growth rate of AAA was significantly greater in women than in men. This may have implications for the frequency of follow-up and timing of repair of AAA in women.


Paper accepted 18 October 2006
Published online 29 January 2007 in Wiley InterScience (www.bjs.co.uk). DOI: 10.1002/bjs.5573

Introduction

Ruptured abdominal aortic aneurysm (AAA) is a catastrophic event; many patients die without reaching appropriate medical help and only 50 per cent of those undergoing surgical repair survive beyond 30 days\(^1,2\). This is in contrast to the elective repair of AAA, which can be achieved with a mortality rate of less than 5 per cent\(^3\).

Longitudinal observational studies suggest that 25–41 per cent of AAAs larger than 5 cm will rupture within 5 years\(^4–6\). The management of AAAs that measure 4–5.5 cm in maximum anteroposterior (AP) diameter is guided by the UK Small Aneurysm Trial\(^7\), which revealed that early elective surgery conferred no long-term survival benefit over regular surveillance. A follow-up publication from the participants of the UK Small Aneurysm Trial found that the risk of AAA rupture was independently associated with aneurysm size, current smoking, mean blood pressure and female sex\(^8\). The risk of aneurysm rupture in women was three times higher than in men.

Growth rate is one of the factors used to identify patients who are at potential risk of aneurysm rupture. The mean growth rate depends on the population studied. Reported expansion rates vary widely, being four times higher in referral-based studies compared with community-based screening programmes. The growth rate of an AAA is associated with its initial size. Other factors such as age, cigarette smoking and severe cardiac disease have also been associated with rapid expansion of AAA. More recently a community-based screening programme has suggested that female sex may be an independent risk factor for rapid growth of AAA\(^9\). The aim of this study was to assess the influence of sex on growth rate in a large series of referral-based infrarenal AAAs.
Patients and methods

Between January 1985 and August 2005 all patients who were referred to the Royal Infirmary of Edinburgh (a supraregional vascular surgery service covering a population of 1-2 million) with AAA and not considered for early aneurysm repair were assessed serially by abdominal ultrasonography. This included all patients with small AAA and patients who either refused or, owing to a range of co-morbidities, were not considered to be candidates for elective aneurysm repair. Maximum AP and transverse diameters of AAAs were measured. The maximum AP diameter was used to assess changes in aneurysm size. All patients with aneurysms smaller than 4 cm in AP diameter underwent yearly ultrasonography and clinical follow-up, whereas patients with an AAA with a maximum AP diameter greater than 4 cm had ultrasonography every 6 months.

As post-mortem examination concerning the cause of death was unavailable in most cases, and it is believed that knowledge of the presence of AAA may have biased the recorded cause of death, the risk of rupture of AAA was not calculated in this study.

The data were stored in an Access® database (Microsoft, Redmond, Washington, USA). Calculation and organization of the data were performed in Excel™ spreadsheets (Microsoft) Statistical analysis was performed using SPSS™ version 12 software (SPSS, Chicago, Illinois, USA). The association between the initial diameter of an AAA and its growth rate was assumed to be linear and was modelled using linear regression analysis. The growth rate of AAA was considered to be the dependent variable, and age, sex and initial aneurysm size the independent variables. Although the use of linear regression analysis to assess expansion of AAAs may be subject to upward bias, as this would affect aneurysm growth rates in men and women equally, it is unlikely to influence comparisons based on sex.

Results

A total of 1255 patients with an AAA were not considered for early repair and had regular follow-up ultrasonography; there were 824 men (65.7 per cent) and 431 women (34.3 per cent). The median age of the patients was 72 (range 41–94) years, 71 (range 41–94) years for men and 74 (range 51–89) years for women (P = 0.031). These patients were followed for a median of 30 (range 6–185) months. The median number of examinations per patient was six (range 2–37).

Of the patients under surveillance, 322 underwent operative repair of the AAA (233 men and 89 women). This was performed for symptomatic or ruptured AAA in 43 patients (29 men and 14 women). The remaining 279 patients underwent elective repair of AAA (204 men and 75 women). Three hundred and ninety-one patients died during follow-up, 195 patients were lost or discharged from follow-up, and the remaining 347 patients remain under surveillance.

Median AP diameter on initial AAA examination was 41 (range 25–83) mm. It was 43 (range 25–83) mm in men and 39 (range 25–72) mm in women (P < 0.001). Fig. 1 illustrates the distribution of AP diameters of AAA in men and women on initial presentation.

The median overall growth rate was 2.79 (range 4.80–37.02) mm per year. A significant correlation was observed between the growth rate of AAAs and their initial size (R² = 0.22, P < 0.001) (Fig. 2). The median growth rate of AAA was significantly higher in women than in men (3.67 (range 1.20–37.02) versus 2.03 (range 4.80–21.00) mm per year; P < 0.01) (Fig. 3). Fig. 4 illustrates the expansion rates of AAA according to initial AP diameter. Weighted linear regression analysis revealed that initial AP diameter (odds ratio (OR) 3.83 (95 per cent confidence interval (c.i.) 3.12 to 4.55); P < 0.001) and female sex (OR 2.04 (95 per cent c.i. 1.68 to 2.40); P = 0.006) were significant predictors of rapid aneurysm growth. Age at initial presentation was not significantly associated with increased aneurysm growth (OR 1.06 (95 per cent c.i. 0.90 to 1.22); P = 0.420) (Table 1).

Discussion

In the present study female sex was found to be an independent risk factor for AAA expansion rate. Median growth rate of AAA in women was almost twice that in men. This was independent of initial aneurysm size

![Fig. 1](https://example.com/f1.png) Distribution of anteroposterior (AP) diameter of abdominal aortic aneurysm on initial presentation in men and women
Aneurysm expansion rate (mm/year)

Fig. 2 Correlation between expansion rate and initial maximum anteroposterior (AP) diameter of the abdominal aortic aneurysm $R^2 = 0.22$, ($P < 0.001$)

Fig. 3 Box and whisker plot illustrating the difference in growth rates of abdominal aortic aneurysms between men and women. The thick black lines represent median values, boxes the interquartile range, whiskers the range and open circles the outlying values.

and age on presentation. A regional community-based AAA surveillance programme has recently reported similar observations on a total of 274 men and 74 women. Being a community-based study the AAAs identified were heavily weighted towards small size (less than 4 cm in diameter). A recent similar sized referral-based study also reported that female sex was associated with increased AAA growth.

The evidence for timing of intervention in patients with asymptomatic AAA derives mainly from the UK Small Aneurysm Trial and the Veterans Administration aneurysm trial, both of which suggest that ultrasonographic surveillance is appropriate for AAAs less than 5.5 cm in AP diameter. The use of expansion rate as an indication for intervention makes sense but has never been validated. A few large-scale longitudinal studies have revealed that aneurysms that rupture have significantly greater expansion rates than non-ruptured AAAs. Brown et al. reported that ruptured AAAs had significantly higher mean expansion rates (0.84 cm per year) than non-ruptured AAAs (0.39 cm per year). Lederle et al. described similar findings, although in a population consisting predominantly of men.

AAA growth has been shown more closely to fit an exponential curve. Yet, in many cases AAA growth...
follows a staccato or discontinuous pattern with periods during which the aneurysm size remains unchanged\(^\text{15}\). Growth is closely associated with initial AAA size, which is probably the main reason behind the higher growth rates in referral-based studies than screening-based studies. Other factors associated with aneurysm growth include smoking and increasing age. A few studies have reported that AAAs appear to expand faster in current smokers\(^\text{16–18}\). Analysis of the patients from the Small Aneurysm Trial by Brady \textit{et al}.\(^\text{10}\) showed that smoking increased AAA growth rates by 15–20 per cent. Although data on smoking history were not available in the present study, there is some evidence that women with an AAA are more likely to be current smokers than men\(^\text{19}\). Although a difference in smoking patterns has not been reported in a Scottish population, it is possible that the result of this study could have been confounded by higher smoking rates in the women.

Brady \textit{et al}.\(^\text{10}\) did not find any evidence to link hypertension or hyperlipidaemia with AAA growth. Although they did not find any significant difference in growth rates between men and women\(^\text{10}\), analysis of the same cohort of patients revealed that hypertension and sex were independent risk factors for AAA rupture\(^\text{15}\). In a recent study of 150 consecutive patients with AAA, Schouten \textit{et al}.\(^\text{20}\) reported that statin use might be associated with significantly slower aneurysm growth. These findings suggest that improvement in cardiovascular risk factors of patients enrolled in surveillance programmes, in addition to improving cardiac survival, may lead to slower AAA growth and a reduced risk of rupture.

All guidelines for management of AAA have included both men and women, and offered identical thresholds for intervention. The UK Small Aneurysm Trial revealed that female sex was an independent risk factor for AAA rupture; the rupture rate in women was three times higher than that in men, despite a smaller initial AP diameter. Furthermore, mean AP diameter preceding rupture was significantly lower in women than men\(^\text{8}\). Brown \textit{et al}.\(^\text{12}\) studied a cohort of 476 patients with AAA and reported a fourfold higher rupture rate in women. A Finnish community-based follow-up study reported that the aortic diameter was less than 5.5 cm in 24 per cent of women with a ruptured AAA, compared with only 5 per cent of men\(^\text{21}\).

Although surgical decisions regarding elective repair of AAA remain complex, this study adds weight to the view that AAA is a more rapidly progressive condition in women. This observation suggests that women with an AAA should have more frequent follow-up and be offered elective intervention at a reduced AP diameter than men.

### References


Background: This study assessed the impact of sex, presentation and treatment on outcome from abdominal aortic aneurysm (AAA) in Scotland.

Methods: All patients admitted from January 1991 to December 2006 with a primary diagnosis of AAA were identified. Patients were stratified by age, sex, admission diagnosis (ruptured versus intact) and procedure performed (endovascular versus open repair). Multivariable logistic regression analysis was used to determine predictors of mortality.

Results: Some 9779 men and 2927 women were admitted with a principal diagnosis of AAA. Women were significantly older than men (median (range) age 75 (35–97) versus 71 (17–96) years; P < 0·001). A higher proportion of women presented with a ruptured AAA (29·5 versus 27·5 per cent; P = 0·043). Age (odds ratio (OR) 2·52 (95 per cent confidence interval 2·36 to 2·74); P < 0·001), female sex (OR 1·63 (1·48 to 1·78); P < 0·001) and admission diagnosis (OR 10·49 (9·53 to 11·54); P < 0·001) were independent predictors of early death, whereas endovascular repair predicted survival (OR 0·67 (0·58 to 0·76); P < 0·001).

Conclusion: Women presenting with an AAA were older and more likely to be admitted with a ruptured aneurysm. Female sex was an independent risk factor for death from AAA.

Introduction

The possibility of sex differences in outcome following abdominal aortic aneurysm (AAA) repair has been a matter of debate for the past decade. It has been suggested that female sex not only adversely affects selection for AAA repair but is also associated with significantly poorer outcome following surgical intervention.

The past two decades have witnessed major advances in understanding of the natural history of AAA, preoperative assessment and the appropriate timing of intervention. In addition, the introduction of endovascular aneurysm repair holds the promise of reducing the peri-procedural morbidity and mortality of treatment. The impact of these advances on national practices and utilization of services has yet to be analysed in Scotland. Furthermore, no information is yet available on whether heightened awareness of AAA in primary care has resulted in the desired outcome of reducing aneurysm rupture rates.

The aim of the present study was to examine the trends in open and endovascular AAA repair, and outcomes over the past 16 years in Scotland, and to explore the possible impact of sex on outcome from AAA.

Methods

Over the past 46 years, National Health Service (NHS) hospitals in Scotland have been obliged to return a standard form (Scottish Morbidity Record form) for all patients admitted to hospital. Patients’ demographics, co-morbidities, surgical and other interventional procedures, and discharge diagnoses are recorded on these forms in coded format. These data are stored electronically by the Information and Statistics Division (ISD) of NHS Scotland and are used to conduct nationwide audits of patient care, and to steer national and regional healthcare policy. The quality of data collected on Scottish Morbidity Record forms (SMR1) in patients with AAA has been audited.
previously and shown to match clinical records in over 99 per cent of cases.

Data regarding admissions for all patients with the primary diagnosis of AAA from 1 January 1991 to 31 December 2006 were obtained from the ISD of NHS Scotland. The diagnosis for each in-hospital event was classified using final hospitalization diagnosis codes of the International Classification of Diseases ninth revision (ICD-9). These were 441.3/441.4 for ruptured AAA or 441.4 for intact AAA. The primary procedure performed during the inpatient event was recorded from the ICD-9 primary procedure codes: 38.34 (aorta resection and anastomosis), 38.36 (abdominal vessel resection with anastomosis), 38.44 (resection of abdominal aorta with replacement), 38.64 (excision of aorta), 39.25 (aorta-iliacl or femoral bypass) and 39.52 (aneurysm repair not otherwise specified) for open repair of AAA, and 39.71 for endovascular AAA repair.

All transfers between hospitals and wards were scrutinized using an anonymized unique database identifier, so only the record of the pertinent admission for an AAA was included for each patient. As each patient was assigned a unique identifier, it was possible to follow him or her through the healthcare system and remove duplicate records.

Discharge data were grouped into the following outcome categories: discharge to home, transfer to a short-term care facility, transfer to a long-term care facility and death. In-hospital death was defined as death recorded during the in-hospital event, within 30 days of admission to hospital or within 7 days of discharge. Data obtained from the ISD were cross-referenced against the Scottish Audit of Surgical Mortality database to assure maximum accuracy. Patients in the database were divided by age (less than 75 years versus 75 years or more), by presentation (ruptured versus intact AAA) and by method of repair (open versus endovascular).

Statistical analysis
The data were stored in a Microsoft® Access database (Microsoft, Redmond, Washington, USA). Organization of the data and calculations were performed in Microsoft® Excel spreadsheets. Statistical analysis was performed using SPSS® version 10 software (SPSS, Chicago, Illinois, USA).

To avoid distributional assumptions, values were stated as median (range) and non-parametric statistical tests (Fr2, Mann–Whitney, Kruskal–Wallis tests) were used. Stepwise multivariable logistic regression analysis was carried out with in-hospital mortality as the dependent variable, and age, sex, presentation diagnosis and procedure type as independent variables. P < 0.050 was considered significant.

Results
Over the 16-year interval, 12 706 patients were admitted with a principal diagnosis of AAA, 9779 of whom were men and 2927 women. The median age of all the patients was 73 (range 17–97) years, 71 (range 17–96) years for men and 75 (range 35–97) years for women (P < 0.001).

The diagnosis was intact AAA in 8818 patients (69.4 per cent) and ruptured AAA in 3888 patients (30.6 per cent). Fig. 1 shows the overall annual rate of admissions for AAA in Scotland as well as the proportions with the primary diagnosis of intact and ruptured AAA. A total of 2694 men (27.5 per cent) had a ruptured AAA, whereas 7085 had an intact AAA. In contrast, 863 women (29.5 per cent) had a ruptured AAA and 2064 had an intact AAA (Fr2 = 4.1, 1 d.f., P = 0.043). The absolute rate of ruptured AAA remained relatively constant over the 16 years of the study (Fig. 1). However, there was a relative reduction in the proportion of patients with a ruptured AAA in the last 5 years compared with the first 5 years (Fr2 = 1431 (28.2 per cent) of 5068 versus 1296 (32.3 per cent) of 4017 respectively; Fr2 = 17.1, 1 d.f., P < 0.001). Figs 2 and 3 illustrate the trends in age-specific rate of admission for intact and ruptured AAA.

A total of 8348 patients underwent surgery for AAA; this included 7466 open repairs (89.4 per cent) and 698 endovascular repairs (8.4 per cent). An additional 184 patients (2.2 per cent) underwent laparotomy without AAA repair. Of patients who underwent open AAA repair, 4441 had a tube graft, 2278 an aortoiliac repair and 508 an aortofemoral bypass. In 239 procedures the type of open repair was not specified. Fig. 4 illustrates the estimated yearly rates of open and endovascular AAA repair.

No surgical procedure was documented for a further 4358 patients. The recorded diagnosis for this group was ruptured AAA in 2254 patients and intact AAA in 2104.

Over the 16-year interval, 2579 patients (20.3 per cent) died within 30 days of admission with an AAA. Some 531 (37.7 per cent) of 1410 patients who underwent attempted repair of a ruptured AAA died, compared with 362 (5.2 per cent) of 6938 patients who had repair of an intact AAA. The 30-day mortality rate was significantly higher following open repair of an intact AAA than after endovascular repair (5.6 per cent (351 patients) versus 1.6 per cent (11 patients); Fr2 = 19.8, 1 d.f., P < 0.001).

Early postoperative mortality rates were significantly higher in women than in men (29.9 per cent (875 women) versus 25.8 per cent (2523 men); Fr2 = 19.3, 1 d.f., P < 0.001). Logistic regression analysis revealed that age, sex, admission diagnosis (ruptured versus intact AAA) and type of procedure performed (open versus endovascular repair)
Fig. 1 Annual rate of hospital admission for abdominal aortic aneurysm in Scotland from 1992 to 2006

Fig. 2 Annual rate of hospital admission with a primary diagnosis of intact abdominal aortic aneurysm stratified by age

Fig. 3 Annual rate of ruptured abdominal aortic aneurysm stratified by age

Discussion

Until recently, the incidence of AAA has risen steadily over the last two and a half decades in Scotland, in line with similar reported increases in Europe, North America, and Australia. There has been a parallel rise in the number of hospital admissions and mortality rates from this condition. Best and co-workers reported a significant increase in mortality from AAA between 1981 and 2000, although they noted that the mortality from AAA reached a plateau in the mid-1990s. This is in contrast to the findings of Semmens and colleagues, who reported a steady reduction in the rates of emergency surgery and mortality from AAA between 1985 and 1994 in Western Australia. McPhee and co-workers found that...
that AAAs may grow faster in women. In the UK, the incidence of AAA is common than in men. Longitudinal studies have shown a more sinister disease process in women, although it is less prevalent compared with 1992–1995. Evans and colleagues did not find any significant difference in mortality rates between men and women who underwent surgical repair of a ruptured AAA, although this has not been a universal observation.

Screening for AAA with abdominal ultrasonography has been shown to be cost-effective and has led to a reduction in aneurysm-related deaths in men. Universal screening of men between the ages of 65 and 80 years is associated with a significant reduction in the regional incidence of ruptured AAA. However, the role of screening for AAA in women remains controversial. To date there is no evidence that screening for AAA in an unselected population of women is associated with a reduction even in aneurysm-related mortality. Scott and colleagues reported the prevalence of AAA in women to be 1.3 per cent, with other authors reporting a similar rate of 0.7–1.3 per cent in unselected populations. Scott et al. did not demonstrate a difference in rupture rates between the women randomized to screening and control populations of women at 5- and 10-year follow-up. They concluded that screening for women was neither clinically indicated nor economically viable. However, they screened an unselected population of women without consideration of risk factors for aneurysm disease and fitness for repair; consequently a large proportion of women who were found to have an AAA did not undergo aneurysm repair.

For screening to be effective in reducing aneurysm-related mortality in women, it will need to be limited to high-risk women who are fit to undergo aneurysm repair. There is increasing evidence that women with atherosclerotic disease are at significantly higher risk of developing AAA. Derubertis and colleagues reported that the prevalence of AAA in women with multiple (more than three) atherosclerotic risk factors was 6% compared with 2% in men. One factor that the authors have not considered is the increased growth rates of AAA and higher aneurysm rupture rate in women, screening in women with multiple risk factors for AAA may become clinically and economically viable.

Acknowledgements

The authors thank Dr J. A. Young for his help with data collection.

References


Significance of the Early Postoperative Duplex Result in Infrainguinal Vein Bypass Surveillance

R. Mofidi,* J. Kelman, O. Berry, S. Bennett, J.A. Murie and A.R.W. Dawson

Department of Vascular Surgery, The Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK

**Background.** Duplex surveillance of infrainguinal vein grafts may not be efficient.

**Methods.** Consecutive patients who had received infrainguinal vein grafts were enrolled in a duplex surveillance program. A first scan at 6 weeks after surgery categorized grafts into four groups: (a) low risk grafts, (b) mild flow disturbance, (c) intermediate stenosis and (d) critical stenosis. Disease progression was assessed over time.

**Results.** Of 364 grafts followed-up for a median of 23 months, 236 (65%) had no flow abnormality at 6-weeks, and had a 40-month cumulative patency rate of 82%. The remaining 128 (35%) grafts had a flow disturbance. Of 29 critical stenoses, 15 were repaired, 11 occluded and three did not change. Of 57 intermediate lesions, 32 progressed to critical, nine occluded, two were repaired, and 14 did not change or improved. Of 42 mild lesions, 16 progressed to a higher grade, four occluded, and 22 did not change or improved. There was no significant difference in graft patency between grafts with repaired stenoses and those without stenoses, but grafts with untreated critical stenoses were associated with lower patency ($p < 0.001$).

**Conclusions.** A duplex scan 6 weeks after operation can predict those patients who require continuing duplex surveillance.

© 2007 European Society for Vascular Surgery. Published by Elsevier Ltd. All rights reserved.

Keywords: Infrainguinal bypass; Vein Graft; Duplex; Surveillance.

**Introduction**

Infrainguinal bypass using autogenous vein is an established treatment for critical ischaemia of the leg. Vein grafts are prone to develop stenoses, which may precipitate failure of the bypass. Stenosis may develop from technical error, vein valve leaflets, pre-existing vein abnormality, and myointimal hyperplasia. Evidence of most of these problems may be recognisable by duplex ultrasound scanning, a technique acknowledged for its accuracy in identifying and grading stenotic lesions that threaten graft patency. Duplex scanning has been widely used for graft surveillance, the rationale being that correction of stenotic lesions is likely to improve graft patency and limb salvage rates.

The wisdom of duplex vein graft surveillance has been recently questioned. The Vein Graft Surveillance (VGST) trial has revealed no significant difference in graft patency or limb salvage in the medium term between patients who were followed up clinically and those who underwent vein graft surveillance. This study, however, recruited patients at 6 weeks after operation, following the first postoperative surveillance scan. It is possible that some grafts with significant early stenosis may have been treated early, effectively excluding them from further follow-up in the trial. Most duplex surveillance protocols are not initiated until 3 months after surgery and, as a consequence, relatively little is known about the incidence and nature of early vein graft stenosis. However, a significant number of bypasses contain stenotic lesions by 6 weeks after operation.

This study assesses the nature of early flow disturbance and stenosis after infrainguinal bypass using autologous vein.

**Patients and Methods**

Consecutive patients, who had an infrainguinal bypass procedure using autologous vein between 1st January 2000 and 31st December 2005, were enrolled...
in a graft surveillance program. This involved duplex scanning at 6 weeks, 3 months, 6 months and 1 year after operation, after which surveillance continued by clinical examination and measurement of ankle-brachial pressure index at intervals of 6 months. Extra duplex scanning outside the surveillance program was performed on clinical grounds.

Ultrasound examination used angle of insonation as close to 60° as possible and began within the inflow artery, progressing down the graft and continuing into the outflow artery. The inflow and outflow vessels were assessed for quality of flow based on velocity, waveforms and colour flow characteristics. Peak systolic velocity was measured at sites of stenosis, at multiple sites within the graft and within the outflow vessel. The velocity ratio at the site of stenotic lesions was calculated. Grafts were categorized into four groups based on duplex findings at the first scan: (a) low risk grafts, (b) mild flow disturbances, (c) intermediate stenosis and (d) critical stenosis. Table 1 shows the velocity criteria identifying different categories of vein graft stenosis identified through duplex surveillance (PSV: Peak Systolic Velocity (cm/s), ABPI: Ankle: Brachial Pressure Index).

<table>
<thead>
<tr>
<th>Indication</th>
<th>Number</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermittent claudication</td>
<td>64</td>
<td>17.6</td>
</tr>
<tr>
<td>Critical ischaemia</td>
<td>279</td>
<td>76.6</td>
</tr>
<tr>
<td>Popliteal artery aneurysm (asymptomatic)</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Popliteal artery aneurysm (symptomatic)</td>
<td>12</td>
<td>3.3</td>
</tr>
<tr>
<td>Popliteal artery aneurysm (ruptured)</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>364</td>
<td>100</td>
</tr>
</tbody>
</table>

The initial patient group of 371 comprised 238 men and 133 women who had 385 bypass procedures in total. During the first 6 weeks after operation 21 grafts occluded and were excluded from further study, leaving 364 vein grafts in 352 patients (225 men and 127 women). The median (range) age of the patients was 71 (37–88) years; 148 (41%) were current smokers, 158 (43%) had diabetes and 32 (9%) had chronic renal failure. The indications for surgery are shown in Table 2. Three hundred and fifty eight bypasses were reversed vein grafts and six were in-situ bypasses. Three hundred and forty one bypasses had their origin at the femoral artery in the groin (93.7%), nine from the external iliac artery (2.5%) and 14 bypasses (3.8%) had their proximal anastomosis from the superficial femoral artery. The distal anastomoses were to the above knee popliteal (139, 38%), the below knee popliteal (154, 42%), and the tibial arteries (71, 20%). Overall 40-month primary patency, primary assisted patency and secondary patency rates were 73, 79 and 80 per cent, respectively.

The 364 bypasses underwent surveillance and the median (range) follow-up was 23 (2–60) months. The first postoperative vein graft surveillance scan was performed at a median of 6 weeks (range 4–9 weeks). At the time of the first duplex scan, 236 grafts (65%) had no significant stenosis; these grafts ran a benign course with a 40-month cumulative patency rate of 82 per cent and a limb salvage rate of 93 per cent.

The first postoperative scan identified 128 grafts with a significant flow abnormality (Table 3). The distribution of these flow disturbances along the length of the vein graft is shown in Fig. 1. Of the 29 grafts with a critical stenoses, 15 (52%) were repaired. A further 11 (38%) grafts with critical stenoses were not repaired and occluded during subsequent follow-up; in six a clinical decision was made not to intervene, three had been scheduled for repair but occlusion supervened, and two patients refused surgery. A final three (10%) grafts with critical stenosis were not repaired but did not change during follow-up.

Of 57 grafts with an intermediate stenosis at 6 weeks, 32 (56%) had lesions that progressed to a critical stenosis; nine (16%) lesions did not progress but the graft occluded during follow-up. Two grafts were repaired and 14 (25%) did not change or
appeared to improve during follow-up. Of the 42 grafts with mild flow abnormalities at 6 weeks, 16 (38%) had lesions that progressed to a more severe category, four (10%) occluded and 22 (52%) did not change or appeared to improve.

Over the whole study period 92 (25%) of the 364 grafts developed a critical or intermediate stenosis. Only 11 of the 92 limbs had any recurrence of symptoms, but 26 had a reduction in ankle-brachial pressure index. Forty-three of these grafts were treated, three by intra-luminal angioplasty, 18 by vein patch, 9 by jump graft, eight by interposition graft and five by revision of the bypass. Eight patients required repeat surgical procedures for recurrent graft stenosis; this was by repeat vein patch angioplasty in three, jump graft repair in four and interposition graft in one patient. Forty-nine limbs with critical or intermediate vein graft stenosis were not treated.

No statistically significant difference was observed in graft patency ($p = 0.19$) or amputation rates ($p = 0.62$) rates between grafts with repaired stenosis and grafts without stenosis. Untreated grafts with critical or intermediate stenosis had significantly lower patency ($p < 0.001$) and higher amputation rates ($p < 0.001$) rates (Figs. 2 and 3).

**Discussion**

Duplex surveillance of infrapopliteal vein grafts remains a controversial issue. While it is true that a significant proportion of vein grafts develop graft threatening stenotic lesions within the first year of the initial procedure, there is little evidence outside observational and case control studies to suggest that duplex surveillance is associated with improved graft patency and improved limb salvage.$^{11-14}$ Duplex surveillance is resource intensive and difficult to justify on the basis of cost, unless a large number of vein grafts, and by extension, limbs are being saved.

A recent multi-centre randomized controlled trial, the VGST trial, showed no benefit from duplex scanning in terms of graft patency, limb salvage or quality-of-life, despite significantly increased costs.$^8$ This study recruited patients at the time of the first postoperative scan, which was performed at a median of 6 weeks after surgery. The reasoning behind this was to exclude grafts occluding within 6 weeks, as affected patients would not benefit from surveillance. The definition of a graft at risk used in the VGST trial was doubling of peak systolic velocity at the site of the stenosis or a post stenotic PSV less than 0.45 m/s.$^8$ Such a definition would have selected grafts with mild and intermediate lesions as well as those grafts with critical stenoses.

In the present study five per cent of grafts failed within the first 6 weeks—a much smaller proportion than the 33 per cent that contained a significant flow disturbance on duplex scanning at 6 weeks after operation. Therefore the early postoperative duplex scan is of more value in identifying grafts at risk of early failure than excluding grafts that have already occluded by the time of the patient’s first postoperative follow-up visit. Ferris et al. reported a relatively high incidence of vein graft stenosis in the first 6 weeks and suggested that an early scan is the most important scan for graft surveillance. It would identify grafts that are likely to occlude by 3 months, when the first surveillance scan is usually performed.$^{10}$ Other authors have reported a similar incidence of vein graft stenosis within the first 6 weeks of operation.$^{15-18}$

Apart from denoting an early threat to patency, early abnormalities also predict the natural history of the graft and outlook for the limb in the medium term. In the present study, vein grafts without flow anomalies at 6 weeks ran a benign course. The vast
majority of grafts that occluded or required intervention exhibited significant anomalies by the time of the 6-week scan. It should prove possible to use this finding to select vein grafts at particular risk for duplex surveillance, thereby reducing the cost of duplex surveillance.8 This is in agreement with the findings of the VGST trial which was in reality a comparison of intensive duplex based vein graft surveillance versus a single early postoperative duplex scan performed at six weeks followed by clinical follow-up.8 Although Davies et al. did not elaborate on the management of the patients with ‘at risk grafts’ which were identified by pre-randomization duplex scan, it seems improbable that these grafts would have been randomized without receiving appropriate surgical or endovascular correction. Since no mechanism existed in the design of VGST trial, for re-including grafts with early stenosis in the study after corrective treatment, it is probable that some or all of these grafts were excluded from randomization.

In the present study, several intermediate stenoses and flow abnormalities did not progress or resolved during follow-up. This finding has been reported previously.13,17,18 Versti et al. studied 17 vein graft intermediate stenoses, situated within the body of the graft that derived from valve cusps. They reported that over half of these lesions regressed without intervention over the follow-up period.18 Most patients with graft stenoses did not experience any recurrence of symptoms throughout follow-up, and only 27 per cent of all critical or intermediate stenoses were associated with a significant drop in ankle-brachial pressure index. This makes it improbable that a useful number of such lesions could be detected by clinical follow-up alone. This is in accordance with the findings of others.10–19

This study is a real life review of vein graft surveillance rather than a trial of one follow-up strategy versus the other. There are two steps to vein graft surveillance. The first step is the identification of patients who have or are likely to develop vein graft stenosis. In that regard the early postoperative duplex scan can play a significant role. The second step is repairing the stenoses that are identified, whilst accepting that this may not be possible in every case. The Kaplan-Meier survival curves are therefore useful in order to illustrate the natural history of vein grafts with no stenosis, vein grafts with significant stenosis which is repaired and vein grafts with significant stenosis which for one reason...
or another are not intervened on rather than a direct comparison between the groups. Factors such as quality of graft and run off, the site and number of stenoses play an important role in selecting patients for repair of vein graft stenosis.

In summary, flow abnormalities detected at 6 weeks after operation can predict the natural history of a vein graft and such abnormalities can be used to select grafts for continued duplex surveillance. Sub-critical stenosis does not necessarily progress to a higher grade and may actually regress with time. For grafts without any flow abnormality at 6 weeks, the yield from continuing with duplex surveillance is likely to be low and probably little better than what is achievable by simple clinical follow-up.

Acknowledgements

The Authors would like to acknowledge Mr. RTA Chalmers, Mr. SCA Fraser and Mr. Z Raza for including their patients in this study, and the significant contribution of our vascular technologists, Ms Sarah-Jane Carmichael, Ms Karen Gallagher & Ms Maggie Glass who performed the vein graft surveillance scans.

References


Accepted 11 April 2007
Available online 22 May 2007
The Value of Vein Graft Surveillance in Bypasses Performed with Small-Diameter Vein Grafts

R. Mofidi, S. Pandanaboyana, M.M. Flett, J. Nagy, G.D. Griffiths, and P.A. Stonebridge, on behalf of the East of Scotland Vascular Network Dundee, United Kingdom

We assessed the impact of preoperative diameter of the venous conduit on reintervention rate and outcome following infrainguinal vein graft bypass. Consecutive infrainguinal vein bypasses between January 2001 and December 2006 were reviewed. All patients underwent preoperative measurement of vein graft diameter (VGD). Grafts were classified into those with VGD <3.5 mm and those with VGD ≥3.5 mm. All patients were enrolled in a duplex surveillance program. The association between VGD and reintervention rate was assessed. Graft patency and amputation rates were compared. There were 377 bypasses followed up for a median of 23 months (range 8-67). VGD was <3.5 mm in 139 grafts (36.9%) and ≥3.5 mm in 238 grafts (63.1%). A higher proportion of smaller vein grafts (32.3%) required reintervention to maintain graft patency compared with larger conduits (20.2%) (χ² = 7.7, p < 0.001). VGD (odds ratio [OR] = 2.87, 95% confidence interval [CI] 1.63-3.81; p < 0.001), smoking (OR = 1.83, 95% CI 1.39-3.20; p = 0.02), and type of bypass (OR = 1.86, 95% CI 1.49-2.47; p = 0.02) were variables associated with higher reintervention rate. There was no difference in graft patency (p = 0.13) or amputation rates (p = 0.35) between the two groups. Use of smaller vein grafts was associated with a higher reintervention rate. Provided that these grafts are surveyed and where necessary repaired, the use of smaller vein grafts is successful and expands the availability of autogenous conduit for infrainguinal arterial reconstruction.

INTRODUCTION

Infrainguinal bypass using autogenous vein is well established in the management of disabling intermittent claudication or critical ischemia of the leg. These vein grafts are prone to develop stenosis, which may precipitate failure of the bypass. Stenosis may develop from technical error, vein valve leaflets, preexisting vein abnormality, and myointimal hyperplasia.

Several factors have been found to influence the outcome of infrainguinal vein bypass. While many of these factors are endogenous to the patients and beyond the direct control of the surgeon, the size and quality of the venous conduit available for bypass have been considered important factors affecting the outcome of infrainguinal bypass. Although the use of smaller veins does not invariably result in graft failure, a number of authors have suggested that grafts with a diameter <3.5 mm are at increased risk of failure. Under these circumstances, the use of composite vein or arm vein grafts rather than uninterrupted small-caliber long saphenous veins is advocated. The aim of this study was to assess the significance of minimum vein diameter measured at preoperative vein mapping on early and medium-term outcomes of infrainguinal vein bypass and to assess the extent of postoperative increase of vein diameter.

METHODS

Consecutive patients who underwent infrainguinal bypass using autologous vein between January 1,
2001, and December 31, 2006, were identified from a prospectively collected database. Data from the database were supplemented by retrospective chart review and cross-referenced against the hospital laboratory databases.

All patients underwent preoperative vein mapping, using duplex ultrasound (Sequoia; Acuson, Redwood, CA) with a 5 MHz linear transducer, by two experienced vascular technologists. Ultrasound examination used an angle of insonation as close to 60° as possible and examined the long and short saphenous veins bilaterally; upper limb veins were examined if the latter two were not available or were of dubious quality. The venous conduit was examined along its length with the patients in the standing position. The maximum internal diameter of the vein was recorded at the proximal and distal thigh, at the knee, and at the ankle. Patients were stratified into two categories based on the minimum diameter of the long saphenous vein in the standing position: <3.5 and ≥3.5 mm. The appropriate length of ipsilateral long saphenous vein was harvested and examined intraoperatively for distensibility and general quality. If this was found unsuitable, the contralateral long saphenous vein was used. Short saphenous and arm veins were used in patients who had previous harvesting of long saphenous vein or in whom long saphenous vein had been found to be unusable as a conduit for bypass surgery by preoperative vein mapping.

Thereafter, all patients were enrolled in a graft surveillance program. This involved duplex scanning and measurement of ankle brachial pressure index (ABPI) before discharge from hospital and at 6 weeks, 3 months, 6 months, 1 year, and 2 years after operation. Extra duplex scanning outside the surveillance program was performed on clinical grounds.

Duplex scanning protocol during vein graft surveillance included assessment of the vein graft along its full length together with examination of inflow and outflow vessels for quality of flow based on velocity, waveforms, and color flow characteristics. Peak systolic velocity (PSV) was measured at sites of stenosis and at multiple sites within the graft and within the outflow vessel. The velocity ratio at the site of stenotic lesions was calculated. Data regarding any postoperative complications and further surgical and radiological interventions were recorded. The postoperative complications, 30-month cumulative patency rates, and amputation rates were calculated in relation to the minimum preoperative vein graft diameter (VGD).

Critical vein graft stenosis was defined as PSV >300 cm/sec or a 3.5-fold increase in PSV at the site of stenosis in comparison to prestenotic PSV, a PSV <40 cm/sec downstream from a stenotic lesion, or a drop of >0.15 in the postoperative ABPI measurements.

Patient demographics, type of operation, conduit, and follow-up information were recorded in a computerized database (Access and Excel; Microsoft, Redmond, WA). Data analysis was performed retrospectively. Statistical analysis was performed using the Statistical Package for Social Sciences, version 12 (SPSS, Inc., Chicago, IL). Results were analyzed and reported in accordance with the reporting standards of the Society for Vascular Surgery and the North American Chapter of the International Society for Cardiovascular Surgery. The groups were compared in terms of stenosis, need for reintervention, graft patency, and amputation. Patency and limb salvage were determined using Kaplan-Meier analysis. The association of graft patency and limb salvage with risk factors for graft failure as ascertained by the site of the distal anastomosis, minimum VGD, and other atherosclerotic risk factors was analyzed using the Cox proportional hazards model. The association between the need for reintervention due to development of a vein graft stenosis and the above variables was assessed using multivariate logistic regression analysis. Differences between groups were assessed using the log rank test. p < 0.05 was considered significant.

RESULTS

There were 377 infrainguinal vein graft bypasses performed in 351 patients during the study period, of whom 238 (67.7%) were men and 113 (32.3%) were women. The median age of the patients was 73 years (range 37-92). One hundred and ninety-seven patients (56.1%) were current smokers, 150 (42.7%) had diabetes, and 33 (9.4%) had chronic renal failure. The indication for surgery was critical limb-threatening ischemia in 330 cases (87.5%) and disabling intermittent claudication in 48 cases (12.5%). Three hundred and forty-six bypasses (92.1%) had their origin at the femoral artery in 92.1% had their origin at the femoral artery in 125 bypasses (33.2%), the below-knee popliteal in 135 (35.8%), and the crural arteries in 117 (31.0%). Three hundred and twenty-six bypasses were reversed vein grafts and 51 were in situ vein graft bypasses. We utilized uninterrupted ipsilateral
or contralateral long saphenous vein as the venous conduit in 321 bypasses (85.1%). In 29 bypasses (7.7%) composite long saphenous vein graft was used, while in 18 (4.8%) short saphenous venous conduit was used; in nine patients (2.4%) upper limb veins were used to perform the bypass procedure.

Median preoperative internal diameter of vein grafts used was 3.9 mm (range 2.4-6.7). The minimum preoperative internal diameter of the vein graft was <3.5 mm in 139 vein grafts (36.9%) and ≥3.5 mm in 238 grafts (63.1%). The distribution of VGD in each group is illustrated in Figure 1. Table I denotes the demographics of patients with preoperative VGD <3.5 mm compared to those with preoperative VGD ≥3.5 mm.

Median duration of follow-up was 23 months (range 8-67). Eighty-four patients (24%) died during the follow-up. Overall Kaplan-Meier estimates at 40 months of the proportions with primary, primary assisted, and secondary patency rates were 73%, 79%, and 83%, respectively.

Eight grafts required reintervention during the early postoperative period for critical flow abnormalities, which were identified in the early postoperative period. Seventy of these were in vein grafts which were <3.5 mm in diameter. Eighty-seven grafts developed critical vein graft stenosis (22.5%) during the follow-up period. Twelve patients (13.8%) developed recurrent symptoms, and 27 patients (31%) had a significant drop in postoperative ABPI measurements. The remainder of patients remained asymptomatic with no clinical evidence of vein graft stenosis.

Of the 87 patients who developed critical vein graft stenosis, nine grafts occluded prior to scheduled repair, another one graft occluded following diagnostic angiography which revealed that the stenotic lesion identified was not amenable to angioplasty, and one graft was not intervened upon. Thirteen grafts underwent open repair, and a further 63 grafts underwent endovascular repair by a mixture of standard and cutting balloon angioplasty.

A significantly higher proportion of vein grafts which were created using veins with preoperative VGD <3.5 mm required reintervention to maintain graft patency compared with vein grafts which were ≥3.5 mm in diameter ($\chi^2 = 7.7, p < 0.001$) (Fig. 2). Preoperative VGD (odds ratio [OR] = 2.87, 95% confidence interval [CI] 1.63-3.81; $p < 0.001$), smoking (OR = 1.83, 95% CI 1.39-3.20; $p = 0.02$), and type of bypass (OR = 1.86, 95% CI 1.49-2.47; $p = 0.02$) were the only variables associated with greater risk of requiring reintervention in order to maintain graft patency (Table II).

The difference in primary assisted patency rates between grafts with preoperative VGD <3.5 mm and grafts with preoperative VGD ≥3.5 mm did not reach statistical significance (log rank = 2.29, $p = 0.13$) (Fig. 3). In addition, there was no significant difference in amputation rates between vein grafts which were created using veins with a minimum preoperative diameter <3.5 mm compared to those with a minimum preoperative diameter ≥3.5 mm (log rank = 1.69, $p = 0.35$) (Fig. 4). There was no significant association between minimum preoperative VGD and reduced primary assisted patency (Table III) or higher amputation rates (Table IV).

**DISCUSSION**

Autogenous vein grafts are the preferred conduits for infrainguinal arterial reconstruction. In the absence of autogenous vein, the use of prosthetic grafts, while necessary, is associated with a higher incidence of graft occlusion, which can lead to loss of the affected limb. However, a commonly encountered problem is veins of a small diameter identified during preoperative vein mapping, which may be considered unsuitable for infrainguinal arterial reconstruction. Previous authors have reported that the incidence of small-caliber vein during vein mapping is relatively high and, under these circumstances, contralateral long saphenous vein, short saphenous veins, or arm veins are sought as conduits for infrainguinal vein graft bypass surgery.

The premier modality used for preoperative vein mapping is B-mode Doppler ultrasound. Preoperative measurements are often made while the patient
is in the standing position in order to increase its diameter. However as these measurements are made while the vein is in the low-pressure venous system, they may not reflect the true distended diameter of a vein after it is placed in the arterial system.11 Previous studies with in situ vein grafts have shown that there is a significant increase in diameter of the vein in the early postoperative scan.11 Interestingly, the percentage increase in vein diameter was more profound with small-diameter veins compared to large-diameter veins.11

Fillinger et al.,15 investigating vein adaptation to the hemodynamic environment of infrainguinal grafts, reported that veins of small diameter respond more to shear stress, resulting in a greater increase in size than veins with large diameter. Interestingly, large-diameter veins with a low initial shear stress showed a decrease in diameter over time.15 However, a similar but competing process of adaptation to arterial circulation is responsible for neointimal hyperplasia, which can result in the development of vein graft stenosis.16 Other gross and histological features of vein graft, such as reduced vein compliance, smooth muscle hyperplasia, and inflammatory infiltrates, have been associated with the development of vein graft stenosis.17-19 In addition, Westerband et al.20 reported a close association between intimal and subintimal angiogenesis and thickness of intima in human stenotic vein grafts.

In this study the use of venous conduit with minimum preoperative diameter <3.5 mm was associated with satisfactory 40-month graft patency and limb salvage rates.21,22 The use of vein with a smaller preoperative diameter meant that most (85%) infrainguinal bypass procedures were performed with uninterrupted long saphenous vein, considerably reducing operative time as well as the need for spliced grafts. The alternative to this approach

Table I. Demographics of patients with preoperative VGD <3.5 mm compared to those with VGD ≥3.5 mm

<table>
<thead>
<tr>
<th>Variable</th>
<th>VGD &lt;3.5 mm (n = 139)</th>
<th>VGD ≥3.5 mm (n = 238)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, years (range)</td>
<td>72 (39-95)</td>
<td>72 (44-89)</td>
<td>p = 0.85</td>
</tr>
<tr>
<td>Gender (M:F)</td>
<td>93:46</td>
<td>171:67</td>
<td></td>
</tr>
<tr>
<td>Indication (claudication:rest pain:tissue loss)</td>
<td>9:81:49</td>
<td>31:115:84</td>
<td></td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>62 (44.9)</td>
<td>88 (37.0)</td>
<td></td>
</tr>
<tr>
<td>Smoker (%)</td>
<td>74 (53.6)</td>
<td>123 (51.7)</td>
<td></td>
</tr>
<tr>
<td>Hyperlipidemia (%)</td>
<td>31 (22.3)</td>
<td>48 (20.2)</td>
<td></td>
</tr>
<tr>
<td>Renal failure (%)</td>
<td>13 (9.4)</td>
<td>20 (9.2)</td>
<td></td>
</tr>
<tr>
<td>Type of bypass (AK fem-pop: BK fem-pop: crural)</td>
<td>77:88:73</td>
<td>44:47:48</td>
<td></td>
</tr>
</tbody>
</table>

Hyperlipidemia was defined as high-density/low-density lipoprotein ratio >3.5 or fasting cholesterol level >5 mmol/L, and severe renal failure was defined as estimated glomerular filtration rate <20 mL/min or on permanent renal replacement therapy.

AK, above-knee; BK, below-knee; fem-pop, femoropopliteal.

*Yates’ correction.

Fig. 2. Comparison of the proportion of vein grafts requiring reintervention to maintain graft patency, stratified by the minimum preoperative internal diameter of the vein graft.

Table II. Factors affecting reintervention rate after infrainguinal bypass assessed using multivariate logistic regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>2.87 (1.63-3.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.89 (0.71-1.05)</td>
<td>0.160</td>
</tr>
<tr>
<td>Renal failure</td>
<td>1.16 (0.85-1.30)</td>
<td>0.310</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.83 (1.39-2.47)</td>
<td>0.022</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.39 (0.90-2.33)</td>
<td>0.113</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>1.39 (0.90-2.33)</td>
<td>0.113</td>
</tr>
</tbody>
</table>
is the use of spliced segments of long saphenous vein or short saphenous, cephalic, or basilic veins instead of smaller-caliber uninterrupted long saphenous veins. When following the latter strategy, the length of conduit required often cannot be obtained by a single segment of vein, and two or more segments have to be used. This raises concerns that such venovenostomy may be a site of intimal hyperplasia. Despite their complexity, patency rates achieved with arm veins are similar to those achieved using uninterrupted long saphenous vein.23

A significant proportion of infrainguinal vein graft bypasses develop graft stenosis, which threatens graft patency.24 Varty et al.8 reported that preoperative VGD is an independent risk factor for development of vein graft stenosis. In this study there was a significantly higher reintervention rate for vein graft stenosis in bypasses which were formed using small-caliber vein grafts. However, despite the fact that VGD was an independent risk factor for reintervention, there was no significant difference in primary assisted patency or amputation rates between grafts created using vein with a minimum preoperative diameter $<3.5$ mm and those with a diameter $\geq 3.5$ mm. This highlights the added importance of vein graft surveillance and timely repair of vein graft stenosis in this population of patients.

Vein graft surveillance has been a controversial topic. Although there is ample evidence that a significant proportion of vein grafts do develop stenosis and that the natural history of vein graft stenosis is graft occlusion,12-20 to date there has been no definitive evidence to suggest that duplex surveillance is

Table III. Factors affecting primary assisted patency after infrainguinal bypass assessed using Cox’s multivariate regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>1.34 (0.94-2.55)</td>
<td>0.0901</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.40 (1.39-4.13)</td>
<td>0.0017</td>
</tr>
<tr>
<td>Renal failure</td>
<td>2.90 (1.65-5.10)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.46 (1.49-4.06)</td>
<td>0.004</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>0.97 (0.44-2.16)</td>
<td>0.95</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>0.85 (0.47-1.54)</td>
<td>0.2776</td>
</tr>
</tbody>
</table>

Table IV. Factors affecting amputation rates after infrainguinal bypass assessed using Cox’s multivariate regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein graft diameter</td>
<td>1.33 (0.78-2.29)</td>
<td>0.295</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.81 (1.62-4.88)</td>
<td>0.002</td>
</tr>
<tr>
<td>Renal failure</td>
<td>3.93 (2.02-7.66)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.29 (1.29-4.05)</td>
<td>0.005</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>0.8339 (0.37-1.87)</td>
<td>0.659</td>
</tr>
<tr>
<td>Type of bypass</td>
<td>0.5876 (0.29-1.19)</td>
<td>0.138</td>
</tr>
</tbody>
</table>
superior to clinical follow-up in improving vein graft patency or reducing amputation rates. Ihilberg et al.\textsuperscript{25} studied 185 vein grafts, which were randomized to duplex or clinical follow-up. They did not find any significant difference between the two groups in graft patency or limb salvage at 1 year. In a study of similar size, Lundell et al.\textsuperscript{26} reported a significant improvement in graft patency in patients who had been recruited into duplex surveillance after 3 years of follow-up. More recently, the Vein Graft Surveillance Randomized Trial (VGST) reported no significant improvement in outcome following duplex surveillance of infrainguinal vein grafts.\textsuperscript{27} All patients who were recruited into the VGST underwent a duplex scan upon recruitment, which was 6 weeks following the bypass procedure. There is increasing evidence that many patients who develop vein graft stenosis have significant evidence of vein graft stenosis by 6 weeks.\textsuperscript{28,29} Conversely, the outcome in vein grafts that do not have evidence of vein graft stenosis in the early postoperative scan is relatively good, and these patients may not benefit from intensive surveillance.\textsuperscript{29} The natural history of vein grafts which develop critical stenosis, especially those occurring early, is graft failure, which is associated with a high rate of limb loss.\textsuperscript{29}

While VGD is one of the factors associated with increased risk of vein graft stenosis, other factors, such as continued smoking, serum fibrinogen levels,\textsuperscript{30,31} and serum C-reactive protein levels,\textsuperscript{30,31} have been associated with the development of progressive neointimal hyperplasia and vein graft failure. In this study, in addition to VGD, continued smoking and distal anastomosis of the bypass were associated with increased incidence of vein graft stenosis. In addition, history of diabetes, smoking, and renal failure were associated with reduced graft patency and increased risk of amputation. While these are not novel observations, they do serve as a reminder of the importance of risk factor modification in the maintenance of graft patency and limb salvage in patients who have undergone infrainguinal vein graft bypass. Vein graft surveillance visits provide an ideal opportunity for secondary prevention of atherosclerotic risk factors, which will have a direct impact on graft patency and limb salvage rates.

The results from this study have shown that small-caliber veins can be used safely for infrainguinal arterial reconstruction, albeit with a significantly higher rate of reintervention in order to maintain graft patency in comparison to larger-caliber veins. Provided that these grafts are entered into a vein graft surveillance program and repaired when necessary, infrainguinal bypass using smaller vein grafts is successful and expands the availability of autogenous conduit for infrainguinal arterial reconstruction.

The authors acknowledge the significant contribution of the vascular technologists Ms. Rosie Levison, Ms. Rose Ross, and Mr. Scott Steedman, who performed the preoperative vein mapping and vein graft surveillance scans.

REFERENCES

Balloon Angioplasty as the Primary Treatment for Failing Infra-inguinal Vein Grafts

R. Mofidi*,1, M. Flett1, J. Nagy1, R. Ross1, G.D. Griffiths1, S. Chakraverty1, P.A. Stonebridge1

Department of Vascular Surgery and Radiology, Ninewells Hospital, Dundee, DD1 9SY, United Kingdom

Submitted 11 September 2008; accepted 28 October 2008
Available online 28 November 2008

KEYWORDS
Vein graft; Stenosis; Angioplasty; Surveillance

Abstract  Background: We sought to evaluate the role of balloon angioplasty as the primary modality in the management of vein graft stenoses.

Methods: Patients who underwent infraringuinal vein graft bypass from January 2002 to December 2007 were enrolled into a surveillance program. Grafts which developed critical stenoses were identified and underwent urgent angiography with a view to angioplasty of the stenotic lesion. Lesions which were deemed unsuitable for angioplasty underwent urgent surgical repair.

Results: Four hundred and eleven grafts were followed up for a median of 19 months (range: 2–61). Ninety-six grafts (22.6%) developed critical stenosis. Twelve grafts occluded prior to repair and one was not intervened upon electively. Eight grafts underwent primary surgical repair. Seventy-six grafts underwent 99 endovascular procedures. Technical success was achieved in 60 grafts (78.9%). Of the grafts in which technical success had not been achieved, eight underwent repeat angioplasty and three were surgically repaired. Twenty-four grafts underwent repeat angioplasty for re-stenosis with a technical success rate of 71%. No difference was observed in graft patency (P = 0.08) or amputation rates (P = 0.32) between the grafts requiring intervention to maintain patency, and grafts which did not. Smoking [OR: 2.61 (95% CI: 1.51–4.53), (P = 0.006)], diabetes [OR: 2.55 (95% CI: 1.49–4.35), (P = 0.006)], renal failure [OR: 1.89 (95% CI: 1.19–3.38), (P = 0.040)] and recurrent stenosis [OR: 3.22 (95% CI: 1.63–4.69), (P < 0.001)] were risk factors for graft occlusion.

Conclusions: Balloon angioplasty of failing infraringuinal vein bypass grafts is safe and can be performed with an acceptable medium term patency rate, albeit with a significant risk of re-stenosis which can be successfully treated in most patients using repeat endovascular intervention.

© 2008 Published by Elsevier Ltd on behalf of European Society for Vascular Surgery.
Introduction

Maintaining patency of infrainguinal vein bypass grafts has been a challenging task for vascular surgeons for decades. These grafts are prone to the development of vein graft stenosis, which may precipitate failure of the bypass. Features associated with the development of vein graft stenosis are recognisable by duplex ultrasound scanning. Duplex scanning has been widely used for graft surveillance, the rationale being that correction of stenotic lesions is likely to improve graft patency and reduce the risk of amputation. Considerable uncertainty still exists about the optimal management of these threatened bypass grafts once they are identified. Traditional operative techniques such as patch angioplasty, interposition graft or replacement graft remain the gold standard for the management of vein graft stenosis. However these techniques require the availability of an additional autogenous vein and carry the risk of morbidity and mortality associated with re-operation. In addition, jump and interposition vein grafts used to treat vein graft stenosis may by themselves be at risk of development of vein graft stenosis. Increased experience with the application of angioplasty in the treatment of stenoses in other venous conduits such as haemodialysis arterio-venous fistulae has led to the introduction of this technique for the management of threatened bypass grafts. However the efficacy of this treatment modality has been questioned and to date no consensus exists about its primary use in the management of vein graft stenosis.

The aim of this study was to assess the efficacy and safety of balloon angioplasty as the primary modality in the management of vein graft stenoses, which are identified as a result of duplex vein graft surveillance.

Methods

Consecutive patients, who underwent infrainguinal bypass procedure using an autologous vein between 1st January 2002 and 31st December 2007 were studied. Infrainguinal vein bypass grafts were performed with a variety of techniques including reversed and in situ bypass grafts. The ipsilateral long saphenous vein was the conduit of choice for bypass. If this was found unsuitable the contralateral long saphenous vein was used. Short saphenous and arm veins were used in patients who had previous harvesting of the long saphenous vein or in whom the long saphenous vein had been found to be unusable.

Thereafter all patients were enrolled in a graft surveillance program. This involved duplex scanning and the measurement of ankle brachial index pressures before discharge from the hospital and at 6 weeks, 3 months, 6 months, 1 year and 2 years after operation. Duplex examination was performed by two experienced vascular technologists using an Acuson Sequoia duplex ultrasound system (Acuson, Redwood, California, USA) using a 5 MHz linear transducer. The duplex scanning protocol during vein graft surveillance included the assessment of inflow and outflow vessels for quality of flow based on velocity, waveforms and colour flow characteristics. Peak systolic velocity was measured at the sites of stenosis, at multiple sites within the graft and within the outflow vessel. The velocity ratio at the site of stenotic lesions was calculated. Any bypass grafts with sub critical flow abnormalities found on duplex scanning were studied more frequently at 3-month intervals, whilst all grafts with critical vein graft stenosis were referred for urgent angiography with a view to primary endovascular treatment. Table 1 shows the duplex parameters used to classify vein graft stenosis and indications for treatment.

Antegrade ipsilateral puncture was the preferred angiographic technique. Mixtures of standard and cutting balloon techniques were used. The size and the type of balloon (cutting versus standard balloon) used for angioplasty were based on the angiographic appearance and the site of the stenotic lesion within the vein graft. All interventions were evaluated with completion angiograms. Technical success was defined as a less than 20% stenosis remaining on completion angiogram (in keeping with the PTA guideline of the Society of Cardiovascular and Interventional Radiology). Vascular laboratory evaluations, including duplex scans, ankle brachial indices, were obtained prior to discharge. Thereafter all patients re-entered the surveillance programme at the starting point. Patients with an unsuccessful endovascular procedure were admitted for urgent surgical correction of vein graft stenosis whilst those with residual stenosis were studied more frequently: at 2 weeks, 6 weeks and then, at 3-month intervals. This process was repeated for all patients who developed recurrent vein graft stenosis. Postoperative hypercholestrolemia was defined as the total cholesterol to high density lipoprotein cholesterol (total cholesterol/HDL) ratio of greater than 3.5. Renal failure was defined as the Kidney Dialysis Outcomes Quality Initiative (KDOQI) chronic kidney disease (CKD) stage 4 or 5 at the time of infringuinal bypass surgery.

Patient demographics, type of operation, conduit and follow up information were recorded in a computerized database (Microsoft Access and Excel, Redmond, Washington, USA). Data analysis was performed retrospectively. Results were analysed and reported in accordance with the reporting standards of the Society for Vascular Surgery and the North American Chapter of the International Society for Cardiovascular Surgery. Statistical analysis was performed using Statistical Package for Social Sciences version 12 SPSS (SPSS, Chicago, Illinois, USA) statistical software. Patency and limb salvage were determined using Kaplan–Meier analysis. The difference in primary assisted patency and amputation rates between patients who underwent angiography with a view to intervention for vein graft stenosis and patients who did not develop vein graft stenosis was assessed using the log rank test. P < 0.05 was considered significant.

Results

The initial patient group of 389 patients comprised 263 men (67.6%) and 126 women (32.4%) who underwent 411 bypass procedures. These grafts were followed up for a median of 19 months (range: 2–61). The median (range) age of the patients was 72 years (range: 38–87); 219 (56.3%) were current smokers, 168 (43.2%) had diabetes and 35 (9%) had...
chronic renal failure. Three hundred and three patients (77.9%) were on statin therapy. Ninety-eight patients (25%) had hypercholesterolemia on follow up [median post-operative TC/HDL ratio: 3.1 (range: 0.8–4.3)]. The indications for surgery are shown in Table 2. Three hundred and fifty six bypasses were reversed vein grafts and 55 were in situ vein graft bypasses. Three hundred and fifty five bypasses (86.6%) had their origin at the femoral artery in the groin, 16 from the external iliac artery (3.9%) and 40 bypasses (9.7%) had their proximal anastomosis from the superficial femoral artery. The distal anastomoses were to the above knee popliteal (132, 34%), the below knee popliteal (148, 35%), and the tibial arteries (131, 31%).

During the follow up period, 272 grafts did not develop significant vein graft stenosis and therefore did not require intervention in order to maintain graft patency, whilst 96 grafts (22.6%) developed critical vein graft stenosis. A further 43 grafts (10.4%) occluded prior to the early post-operative (6 week) follow up scan. The distribution of these stenoses along the length of the vein graft is shown in Fig. 1. Of these 12 grafts occluded prior to angiography and another graft occluded following diagnostic angiography, which revealed that the stenotic lesion identified, was not amenable to angioplasty. One graft was found to be stenotic for a significant proportion of its length and was electively not intervened upon. Eight grafts underwent open repair after diagnostic angiography, which identified lesions not amenable to angioplasty. This was due to inability to traverse the stenosis angiographically in three cases, an unfavourable proximal anastomotic lesion in two cases and external compression in a further three.

Seventy-six grafts underwent a total of 99 endovascular procedures (Figs. 2 and 3). Median age of the grafts at the time of the first angioplasty was 5 months (range: 7 weeks–27 months). Forty-one primary angioplasties were performed with a standard balloon and 34 with a cutting balloon. Initial technical success was achieved in 60 grafts (78.9%). Of the grafts in which initial technical success had not been achieved, eight underwent repeat angioplasty with good technical success and another graft was found to have thrombosed the next day. This graft was successfully managed with local catheter thrombolysis with good results. Three grafts occluded prior to attempted surgical repair and another four grafts were surgically repaired. A total of 27 grafts (45%) developed re-stenosis of which 24 underwent a repeat attempt at angioplasty, with technical success in 17 grafts (71%). Repeat angioplasties were performed with cutting balloons. Of the grafts that had developed re-stenosis three occurred prior to the repeat attempt at angioplasty and a further patient was deemed too unfit to undergo further intervention. There was no peri-procedural mortality. Vein graft angioplasty was associated with two puncture site haematomas both of which were treated conservatively; in addition one patient who had undergone cutting balloon angioplasty developed a delayed retroperitoneal hemorrhage from a contralateral puncture site which was managed by open surgery.

### Table 1

<table>
<thead>
<tr>
<th>Indication</th>
<th>Pressure at site of stenosis</th>
<th>Post stenotic Pressure</th>
<th>Drop in ABPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical stenosis (urgent treatment)</td>
<td>&gt;300</td>
<td>3.5</td>
<td>&gt;0.15</td>
</tr>
<tr>
<td>Mild flow disturbance (intensive surveillance)</td>
<td>200–300</td>
<td>2</td>
<td>&gt;0.15</td>
</tr>
<tr>
<td>Low risk grafts (standard surveillance)</td>
<td>&lt;200</td>
<td>&lt;2</td>
<td>&lt;0.15</td>
</tr>
</tbody>
</table>

Overall 30-month primary patency, primary assisted patency and secondary patency rates were 73.2%, 82.6% and 84.3%, respectively. No significant difference was observed in graft patency (log rank = 1.83, P = 0.08) or amputation rates (log rank = 0.89, P = 0.32) between the

### Table 2

<table>
<thead>
<tr>
<th>Indication</th>
<th>Number</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermittent claudication</td>
<td>39</td>
<td>9.5</td>
</tr>
<tr>
<td>Critical ischaemia</td>
<td>340</td>
<td>82.7</td>
</tr>
<tr>
<td>Popliteal artery</td>
<td>18</td>
<td>4.4</td>
</tr>
<tr>
<td>Popliteal artery (asymptomatic)</td>
<td>14</td>
<td>3.4</td>
</tr>
<tr>
<td>Total</td>
<td>411</td>
<td>100</td>
</tr>
</tbody>
</table>
grafts requiring intervention to maintain patency and grafts which did not (Figs. 4 and 5). Smoking [odds ratio: 2.61 (95% CI: 1.51–4.53), \(P = 0.006\)], diabetes [odds ratio: 2.55 (95% CI: 1.49–4.35), \(P = 0.006\)] and persistent hyperlipidemia [odds ratio: 1.89 (95% CI: 1.19–3.38), \(P = 0.04\)] were associated with reduced graft patency whilst the need for intervention in order to maintain graft patency was not [odds ratio: 1.41 (95% CI: 0.79–2.60), \(P = 0.180\)]. However recurrent graft stenosis was an independent risk factor for vein graft failure [odds ratio: 3.22 (95% CI: 1.63–4.69), \(P < 0.001\)] (Table 3). Similarly, continued smoking [odds ratio: 4.03 (95% CI: 2.07–7.84), \(P < 0.0001\)], diabetes [odds ratio: 2.86 (95% CI: 1.65–4.97), \(P = 0.0002\)] and persistent hyperlipidemia [odds ratio: 2.49 (95% CI: 1.39–4.47), \(P = 0.021\)] were associated with an increased risk of amputation, whereas the need for intervention in order to maintain graft patency was not [odds ratio: 1.24 (95% CI: 0.68–1.91), \(P = 0.32\)]. Recurrent vein graft stenosis was a risk factor for amputation of the ipsilateral limb [odds ratio: 2.51 (95% CI: 1.41–4.32), \(P = 0.002\)] (Table 4).

<table>
<thead>
<tr>
<th>Follow up (months)</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grafts requiring no intervention</td>
<td>272</td>
<td>206</td>
<td>160</td>
<td>85</td>
<td>49</td>
</tr>
<tr>
<td>Grafts requiring intervention to maintain patency</td>
<td>96</td>
<td>75</td>
<td>49</td>
<td>33</td>
<td>18</td>
</tr>
</tbody>
</table>

In addition vein graft surveillance identified two pseudoaneurysms associated with the proximal anastomosis, which were successfully treated with jump graft repair and three grafts with threatened inflow, which were treated with balloon angioplasty with successful technical results.

**Discussion**

The optimal treatment of a threatened vein graft bypass has been a source of debate for several decades. The controversy remains about the definition of a threatened
infrainguinal vein bypass graft and also the clinical significance of the stenotic lesions once they are identified through duplex vein graft surveillance. The authors offer graft surveillance to all patients who have undergone infrainguinal vein graft bypass procedures and this study relates to the efficacy of angioplasty for the management of these lesions rather than a discussion regarding the efficacy of duplex vein graft surveillance.

Surgical revision of failing infrainguinal vein grafts has been the gold standard in the management of vein graft stenosis. Techniques of vein patch angioplasty for short stenoses and interposition or jump graft repair for longer lesions have been associated with excellent medium and long term patency rates. Repair of critical vein graft stenosis has been shown to return patency and amputation rates of the affected vein grafts to levels approaching that of vein grafts with no significant stenosis. On the other hand surgical repair of vein graft stenosis is an invasive procedure requiring general or regional anaesthesia in a population of patients with significant co-morbidity. It is also partly dependent on the availability of an appropriate length of an additional venous conduit. In addition revised vein grafts are themselves at risk of development of vein graft stenosis.

Some years ago Berkowitz et al. reported on the results of selective use of angioplasty for the management of vein graft stenosis. They performed angioplasty as the primary treatment modality in 81% of failing infrainguinal reversed vein grafts. The rest were repaired by short jump grafts. They reported a 5-year primary assisted patency rate of 61%. Lesions in the proximal graft, proximal anastomosis, and distal graft taken as a group had significantly better patency than the mid-graft and distal anastomotic lesions. Subsequent authors have reported similar results following vein graft angioplasty.

In this study a significant proportion of grafts, which had undergone angioplasty with a satisfactory technical result, developed re-stenosis, requiring secondary procedures. In the majority of patients the secondary procedure was endovascular with a technical success rate approaching that of angioplasty for primary vein graft stenosis. Carlson et al. reported their experience of PTA of infrainguinal vein graft stenoses in 36 patients. Their initial technical success rate was 91% and overall graft patency rate was 78% at 24 months. However, nine bypasses (25%) required further attempts at angioplasty for recurrent vein graft stenosis and seven grafts required secondary procedures. This high-lights the importance of continued graft surveillance and maintenance intervention in patients in whom successful endovascular repair of vein graft stenosis has been achieved. After successful treatment of vein graft stenosis our practice was to re-enter the graft at the starting point of vein graft surveillance. This resulted in most re-stenoses being identified and treated in a timely manner.
A number of authors have reported that, notwithstanding the higher re-stenosis rates and the need for re-intervention, the results of endovascular treatment for vein graft stenosis are equivalent to open surgical repair.\textsuperscript{18,32,33} Although most of these observations were performed in small studies with heterogeneous patient groups which were not randomized, they were potentially subject to type II error. Avino et al.\textsuperscript{34} reviewed a series of 144 infrainguinal vein graft stenoses from a population of 528 grafts, which were under surveillance over a period of 6 years.\textsuperscript{26} Seventy-seven of the failing grafts were treated with open surgical repair and 67 were treated using PTA. They reported similar stenosis free patency rates of 63\% for both groups\textsuperscript{26} and no significant difference in patency rates between interventions for primary versus recurrent vein graft stenoses or between infra-genicular and supra-genicular bypasses.\textsuperscript{26} However this has not been a universal observation; Berceli et al.\textsuperscript{34} reported that open surgical revascularization imparts an improved graft survival over endovascular interventions. They also observed no significant difference in the hospital length of stay or global quality of life between the two groups.\textsuperscript{34} One might postulate these outcomes would be potential benefits of endovascular intervention.\textsuperscript{34} In this study renal failure was a risk factor for vein graft failure.

The use of cutting balloon angioplasty is a relatively new development for the management of vein graft stenosis.

A number of authors have reported that, notwithstanding the higher re-stenosis rates and the need for re-intervention, the results of endovascular treatment for vein graft stenosis are equivalent to open surgical repair.\textsuperscript{18,32,33} Although most of these observations were performed in small studies with heterogeneous patient groups which were not randomized, they were potentially subject to type II error. Avino et al.\textsuperscript{34} reviewed a series of 144 infrainguinal vein graft stenoses from a population of 528 grafts, which were under surveillance over a period of 6 years.\textsuperscript{26} Seventy-seven of the failing grafts were treated with open surgical repair and 67 were treated using PTA. They reported similar stenosis free patency rates of 63\% for both groups\textsuperscript{26} and no significant difference in patency rates between interventions for primary versus recurrent vein graft stenoses or between infra-genicular and supra-genicular bypasses.\textsuperscript{26} However this has not been a universal observation; Berceli et al.\textsuperscript{34} reported that open surgical revascularization imparts an improved graft survival over endovascular interventions. They also observed no significant difference in the hospital length of stay or global quality of life between the two groups.\textsuperscript{34} One might postulate these outcomes would be potential benefits of endovascular intervention.\textsuperscript{34} In this study renal failure was a risk factor for vein graft failure.

The use of cutting balloon angioplasty is a relatively new development for the management of vein graft stenosis.

### Table 3
Factors affecting primary assisted patency after infrainguinal bypass assessed using Cox’s multivariate regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.03 (0.62–1.36)</td>
<td>0.85</td>
</tr>
<tr>
<td>Gender</td>
<td>0.98 (0.71–1.10)</td>
<td>0.61</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.55 (1.49–4.35)</td>
<td>0.006</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.613 (1.51–4.53)</td>
<td>0.006</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.80 (0.36–1.80)</td>
<td>0.592</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.89 (1.19–3.38)</td>
<td>0.040</td>
</tr>
<tr>
<td>Intervention to maintain graft patency</td>
<td>1.41 (0.79–2.60)</td>
<td>0.18</td>
</tr>
<tr>
<td>Recurrent stenosis</td>
<td>3.22 (1.63–4.69)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

### Table 4
Factors affecting amputation rates after infrainguinal bypass assessed using Cox’s multivariate regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.12 (0.41–1.53)</td>
<td>0.41</td>
</tr>
<tr>
<td>Gender</td>
<td>0.86 (0.51–1.21)</td>
<td>0.52</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.8618 (1.65–4.97)</td>
<td>0.002</td>
</tr>
<tr>
<td>Smoking</td>
<td>4.03 (2.07–7.84)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.93 (0.42–2.057)</td>
<td>0.852</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>2.49 (1.39–4.47)</td>
<td>0.021</td>
</tr>
<tr>
<td>Intervention to maintain graft patency</td>
<td>1.24 (0.68–1.91)</td>
<td>0.310</td>
</tr>
<tr>
<td>Recurrent stenosis</td>
<td>2.51 (1.41–4.32)</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Engelke et al. reviewed their preliminary experience in the use of cutting balloon PTA for the treatment of infrainguinal vein graft stenosis. They reported 18-month primary assisted and secondary patency rates of 67% and 83% respectively. A larger study by Garvin et al. appears to confirm that cutting balloon angioplasty is associated with a high technical success rate but they reported relatively low term patency and higher complication rates. They therefore cautioned against its widespread use for the treatment of vein graft stenosis. In this study in 16 grafts (20.6%) at the time of initial attempt at angioplasty satisfactory technical success was not achieved. These grafts are at a significant risk of early graft occlusion. Therefore consideration should be given to immediate or early surgical or endovascular re-intervention. Our practice was to perform duplex reassessment of the stenotic lesion following the angioplasty attempt. Grafts, which had a persistent severe stenosis, underwent urgent surgical repair. Whilst grafts with intermediate residual stenosis underwent early duplex follow up with a repeat attempt at angioplasty if clinically indicated. Experience with other arterialized venous conduits suggests that immediate attempts at repeat angioplasty of venous conduits with cutting balloons are associated with a risk of graft rupture. Multivariate analysis offers some interesting insight into the factors that may influence outcomes after endovascular intervention. Not unexpectedly, smoking and diabetes were significant risk factors for graft occlusion and amputation. In addition hypercholesterolemia during follow up was associated with increased risk of graft occlusion and amputation. The authors used this as a surrogate marker of compliance with statin therapy. Berceli et al. reported that statin therapy was associated with improved graft durability after both endovascular and surgical revision. In addition they found that the diagnosis of hypertension was more frequent in patients receiving statin therapy. Berceli et al. reports that the estimated glomerular filtration rate predicts the predictor of survival, leg salvage and amputation free survival in patients with critical lower limb ischaemia. Critical vein graft stenosis is a complex clinical problem. Provided that grafts are surveyed closely, primary endovascular treatment of failing infrainguinal vein grafts appears to be safe and is associated with acceptable early and medium term patency rates.

Funding
N/A.

Conflict of Interest
None declared.

Ethical Approval
N/A.

References
18. Sanchez LA, Suggs WD, Marin ML, Panetta TF, Wengertner KR, Verith FJ. Is percutaneous balloon angioplasty appropriate in the...
Development of a Decision Tree to Streamline Infrainguinal Vein Graft Surveillance

Olivia M.B. McBride, Reza Mofidi, Gareth D. Griffiths, A. Raymond Dawson, Roderick T.A. Chalmers, and Peter A. Stonebridge, Edinburgh and Dundee, United Kingdom

Background: Duplex ultrasound (DU) remains the gold standard for identification and grading of infrainguinal vein graft stenosis. However, DU-based graft surveillance remains controversial. The aim of this study was to develop a decision tree to identify high-risk grafts which would benefit from DU-based surveillance.

Methods: Consecutive patients undergoing infrainguinal vein graft bypass were enrolled in a DU surveillance program. An early postoperative DU was performed at a median of 6 weeks (range 4–9). Based on the findings of this scan and 4 established risk factors for graft failure (diabetes, smoking, infragenicular distal anastomosis, revision bypass surgery), a classification and regression tree (CART) was created to stratify grafts into grafts which are at high and low risk of developing severe stenosis or occlusion. The accuracy of the CART model was evaluated using area under receiver operator characteristic curve (ROC).

Results: Of 796 vein graft bypasses performed (760 patients), 64 grafts were occluded by the first surveillance visit and 732 vein grafts were entered into surveillance program. The CART model stratified 299 grafts (40.8%) as low-risk and 433 (59.2%) as high-risk grafts. One hundred twenty-six (17.2%) developed critical vein graft stenosis. Overall, 30-month primary patency, primary-assisted and secondary patency rates were 76.2%, 83.6%, and 85.3%, respectively. The area under ROC curve for the CART model was 0.88 (95% confidence interval 0.81–0.94). Primary graft patency rates were higher in low-risk versus high-risk grafts (log rank 186, P < 0.0001). Amputation rates were significantly higher in the high-risk grafts compared with low-risk ones (log rank 118, P < 0.0001).

Conclusion: A clinical decision rule based on readily available clinical data and the findings of significant flow abnormalities on an early postoperative DU scan successfully identifies grafts at high risk of failure and will contribute to safely improving the efficacy of infrainguinal vein graft surveillance services.

INTRODUCTION

Infrainguinal bypass using autologous vein is an established treatment for critical lower limb ischemia and disabling intermittent claudication. Despite advances in surgical technique and improvements in preventative medical therapy for atherosclerosis, these grafts are prone to developing stenoses, which may precipitate failure of the bypass. Over the past 30 years, little has changed, with 25–40% of infrainguinal bypasses demonstrating a significant flow disturbance within the graft by 6 weeks after operation and these are often asymptomatic.
Recognition of risk factors for vein graft failure is essential for risk-factor modification and identifying individuals who would benefit from more intensive surveillance protocols. It is evident that graft failure is multifactorial and factors are often viewed along temporal lines with respect to the postoperative period; the early failures are largely attributable to technical issues and operative technique and this group present acutely requiring intervention. In the early to intermediate postoperative period vascular remodeling occurs leading to the development of myointimal hyperplasia and late graft failure is commonly attributed to progression of atherosclerosis.\(^{10}\) The contribution of patient-related factors to graft failure is complex. There is no doubt that smoking plays a significant role and that diabetes and renal failure are associated with increased mortality and amputation rates.\(^{11-15}\)

Evidence of most of these abnormalities are recognizable by duplex ultrasound scanning, a modality acknowledged for its accuracy in identifying and grading stenotic lesions that threaten graft patency.\(^{7}\) In addition to clinical examination and measurement of ankle brachial pressure index (ABPI), duplex ultrasound scanning has been widely used for graft surveillance. The aim of this is to identify grafts at risk of failure, as intervention in a patent but failing graft results in improved long-term patency and limb salvage rates compared with rescue of an occluded graft.\(^{16-19}\)

Surveillance programs are resource intensive and put a considerable strain on the workload of a vascular unit. The necessity of repetitive testing over long periods of time is expensive and time consuming. In the modern climate of increasing cost-awareness, these observations highlight the need for an improvement in the rationalization of vein graft surveillance.

Recursive partitioning is a multivariable analysis model; it creates a decision tree, which strives to correctly classify members of the population based on several dichotomous variables.\(^{20}\) Decision (Classification) trees are often used as a clinical decision support tool, where they map clinical and therapeutic decisions and the consequences of these decisions in a graph or tree-like manner, from which a therapeutic algorithm can be created. Such a model is ideally suited to the problem of identifying grafts that are at the highest risk of developing graft stenosis and failure.

The aim of this study was to create and validate a simple clinical decision support tool to stratify vein grafts into the appropriate surveillance program, based on the level of risk.

**MATERIALS AND METHODS**

Consecutive patients who underwent infrainguinal bypass procedures using autologous conduit in 2 tertiary vascular surgical units in the United Kingdom for 5 years were enrolled into vein graft surveillance programs.\(^{21}\) Approval to develop the database was granted by the Caldicott Guardian in both centers.

Infrainguinal vein bypass grafts were performed with a variety of techniques including reversed and in situ bypass grafts, with the ipsilateral long saphenous vein being the conduit of choice. If unsuitable, the contralateral long saphenous vein was used, followed by the short saphenous and arm veins, if required. Patients were then enrolled into a vein graft surveillance program, with each visit involving duplex ultrasound and the measurement of ABPI. Based on the duplex ultrasound result, we used a modification of Rutherford’s recommendations to the Society of Vascular Surgery to classify grafts into mild, moderate, or critical stenosis and grafts that were occluded (Table I). The frequency of surveillance visits (6 weeks, 3, 6, 9, 12, 18, 24 months), duplex ultrasound scanning protocol, and criteria for intervention were similar in both vascular surgical units. Data from these 2 units were used to develop a decision tree model to identify vein grafts at high risk of developing stenosis or occlusion.

Patient demographics, type of operation, conduit, and follow-up information were recorded in a computerized database (Microsoft Excel, Seattle, WA). Data analysis was performed retrospectively. Results were analyzed and reported in accordance with the reporting standards of the Society for Vascular Surgery and the North American Chapter of the International Society for Cardiovascular Surgery.\(^{22}\)

**Study End Points**

The primary end point for the study was development of severe vein graft stenosis or graft occlusion during postoperative follow-up, while risk of amputation was the secondary end point of the study.

Cox regression analysis was used in parallel to the decision tree analysis to identify potential factors associated with infrainguinal vein bypass graft failure, using Statistical Package for Social Sciences version 12 (SPSS, Chicago, IL) software. Patency and limb loss rates were determined using Kaplan–Meier analysis.

**Decision Tree Analysis**

A Classification and Regression Tree Model (CART) was used to identify patients who are likely to
develop significant vein graft stenosis or occlusion (within 2 years of duplex ultrasound follow-up), with the aim of identifying patients who would benefit from more intensive duplex ultrasound-based follow-up. The variables used were clinical and duplex ultrasound findings, which were available at the time of the first postoperative assessment. The Gini coefficient was used to assess heterogeneity in the distribution of patients from the parent node to children nodes. The Gini coefficient is zero when all observations at a node belong to one level of a dependent variable and is 0.5 when observations are equally distributed in various levels of a dependent variable. The best split on a variable will be the one that minimizes the Gini coefficient. This process continues until a termination criterion is met.

The level of significance of each variable will be adjusted by Bonferroni’s correction, where the decision tree is first fitted to the data, with the most complex topography and the greatest number of nodes and hence accuracy. Prediction at each node is made based on the weight of observations in each category of the dependent variable and misclassification cost. The CART then undergoes a process of pruning, during which the terminal nodes will be deleted only if its elimination causes a misclassification cost which is significantly lower than the reduction in complexity. This process continues until the root node. The optimal CART is selected as the final predictive model. The optimal CART selected was one that prioritized sensitivity over specificity, as its purpose was to relate the decision or an attribute to its consequences while maintaining the simplest topography.

The aim of the model was to repeatedly divide patients into subgroups, each of which consisted of patients with or without severe vein graft stenosis or occlusion thereby stratifying patients into high-risk groups who should undergo duplex ultrasound-based surveillance and low-risk groups who do not require such intensive follow-up. The CART model was then used to describe the possible pathways of different cohorts during the first 2 years of vein graft surveillance.

RESULTS

There were 796 infrainguinal vein graft bypasses performed in 760 patients, of whom 501 (66%) were men and 259 (34%) were women. The median age of the patients was 72 years (range 37–92 years). Three hundred sixty-seven patients (48.2%) were current smokers, 308 (42.9%) had diabetes, and 67 (8.8%) had chronic renal failure.

Critical limb ischemia was the primary indication for surgery (n = 619, 77.7%), with the remaining patients undergoing infrainguinal bypass procedures for intermittent claudication (n = 124, 15.7%) or popliteal artery aneurysms (n = 53, 6.7%). In 21 patients the indication was short distance claudication with occasional rest pain but no tissue loss. Six hundred ninety-six grafts (87.4%) had their origin at the femoral artery in the groin, 25 (3.1%) had their origin at the external iliac artery, and 54 (6.8%) had the proximal anastomosis from the superficial femoral artery. The distal anastomoses were to the above-knee popliteal in 275 bypasses (34.5%), below-knee popliteal in 313 (39.3%), and crural arteries in 208 (26.1%). Seven hundred thirty-three bypasses (92.1%) were reversed vein grafts and 63 (7.9%) were in situ vein graft bypasses.

We utilized uninterrupted ipsilateral or contralateral long saphenous vein as the venous conduit in 729 bypasses (91.5%). In 34 bypasses (4.3%).

<table>
<thead>
<tr>
<th>Table 1. Duplex ultrasound classification—modified from recommendations by Rutherford et al. 1997, Society of Vascular Surgery²²</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSV at the site of stenosis</td>
</tr>
<tr>
<td>Mild stenosis</td>
</tr>
<tr>
<td>Moderate stenosis</td>
</tr>
<tr>
<td>Critical stenosis</td>
</tr>
<tr>
<td>Occluded</td>
</tr>
</tbody>
</table>

Poststenotic PSV of blood flow in the graft downstream from the stenotic lesion.

N/A, not applicable; PSV, peak systolic velocity; ABPI, ankle brachial pressure index.

Fig. 1. Distribution of duplex scanned stenoses and flow-related abnormalities along the length of the graft.
A composite long saphenous vein graft was used, while in 22 (2.8%), short saphenous venous conduit was used; in 11 patients (1.4%) upper limb veins were used to perform the bypass procedure.

The first postoperative vein graft surveillance scan was performed at a median of 6 weeks (range 4–9 weeks). By the time of this scan, 64 grafts had occluded and were excluded from further study, leaving 732 vein grafts, which were entered into a surveillance program.

Figure 1 illustrates the distribution of critical, moderate, and mild stenosis within the bypass grafts. Two-thirds of critical stenosis occurred in the proximal third of the graft. The CART model stratified 299 grafts (40.8%) as low-risk and 433 (59.2%) as high-risk grafts (Fig. 2).

During the vein graft surveillance period, 279 (35.1%) grafts developed a significant flow disturbance (moderate or critical), of which 126 (45.2%) were classified as critical vein graft stenosis. The majority of all significant flow disturbances, 249 (89.2%), were present in the first follow-up scan. Ninety-nine (78.6%) grafts with a critical stenosis were intervened upon, using open surgical techniques in 23 grafts and endovascular intervention in 76. A further 27 grafts were not intervened upon or occluded before scheduled intervention.

Overall, 30-month primary patency, primary-assisted patency, and secondary patency rates were 76.2%, 83.6%, and 85.3%, respectively (excluding the 64 grafts that occluded before the first clinic visit). The results of Cox regression analysis of potential risk factors associated with infrainguinal
vein graft bypass graft stenosis or occlusion are shown in Table II.

### Table II. Cox regression analysis of potential factors associated with infrainguinal vein graft bypass graft revision

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio (95% CI)</th>
<th>Significance (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.03 (0.62–1.36)</td>
<td>0.81</td>
</tr>
<tr>
<td>Gender</td>
<td>0.98 (0.71–1.10)</td>
<td>0.72</td>
</tr>
<tr>
<td>Critical ischemia</td>
<td>1.41 (0.79–2.60)</td>
<td>0.18</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.86 (1.65–4.97)</td>
<td>0.002</td>
</tr>
<tr>
<td>Smoking (within 3 months of bypass)</td>
<td>2.613 (1.51–4.53)</td>
<td>0.006</td>
</tr>
<tr>
<td>Previous failed ipsilateral bypass</td>
<td>2.51 (1.41–4.32)</td>
<td>0.008</td>
</tr>
<tr>
<td>Nonreversed or in situ saphenous vein</td>
<td>0.87 (0.51–1.21)</td>
<td>0.52</td>
</tr>
<tr>
<td>Nonsingle segment saphenous vein</td>
<td>0.79 (0.36–1.80)</td>
<td>0.63</td>
</tr>
<tr>
<td>Infragenicular distal anastomosis</td>
<td>2.40 (1.31–4.47)</td>
<td>0.01</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0.80 (0.36–1.80)</td>
<td>0.6</td>
</tr>
<tr>
<td>Significant flow abnormalities on early postoperative duplex ultrasound scan</td>
<td>3.22 (1.63–4.69)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Decision Tree Analysis**

Seven hundred thirty-six patients were classified through 5 items into 10 different categories with 2-year stenosis or occlusion rates ranging from 5% to 95%.

Figure 2 illustrates the final topography of the decision tree and the proportion of patients assigned to each group. The receiver operator characteristic (ROC) curve for the decision tree is shown in Figure 3. The model was able to stratify high-risk grafts that were at risk from significant vein graft stenosis or occlusion with good accuracy. Area under the ROC curve for prediction of graft stenosis or occlusion was 0.88 (95% confidence Interval [CI] 0.81–0.94), and the sensitivity of the decision tree for prediction of graft stenosis or occlusion was 92% (95% CI 89–93). The specificity of the decision tree for prediction of graft stenosis or occlusion was 75% (95% CI 69–82).

The decision tree has a negative predictive value of 97%, reassuring clinicians that patients categorized as low risk have a low incidence of graft failure and limb loss following the early postoperative duplex ultrasound scan and review.

**DISCUSSION**

The surveillance of infrainguinal vein bypass grafts remains a controversial issue. A significant proportion of these grafts will develop graft-threatening stenosis during the course of the first postoperative year.6,17,23–26 If left untreated these stenoses are associated with a 3 to 6-fold increase in graft occlusion.4,17,27,28 The outcome for an occluded graft is bleak, with secondary patency rates significantly reduced compared with revision of a patent but failing graft.19,29 This, together with the good outcomes achieved following corrective intervention for graft stenosis, lends support for a process of active identification of failing vein grafts to improve the long-term patency and limb salvage rates for infrainguinal bypass surgery.17,23,30–32 This is in contrast to prosthetic grafts where duplex ultrasound-based surveillance is poor at detecting lesions before occlusion.32

Figure 4 illustrates the Kaplan–Meier survival curves for primary graft patency and amputation rates for grafts that have been classified as high and low risk based on the model. They show that grafts classified as high risk had significantly lower primary patency rate (log rank 186, P < 0.001) and were associated with higher risk of amputation (log rank 118, P < 0.001) during follow-up.

Figure 3. The receiver operator characteristic curve for the decision tree.

**Figure 4** illustrates the Kaplan–Meier survival curves for primary graft patency and amputation rates for grafts that have been classified as high and low risk based on the model. They show that grafts classified as high risk had significantly lower primary patency rate (log rank 186, P < 0.001) and were associated with higher risk of amputation (log rank 118, P < 0.001) during follow-up.
intensive duplex ultrasound-based scanning compared with clinical follow-up. In VGST, patients were randomized following the first postoperative review which was conducted 4–12 weeks postoperatively. This early postoperative review involved a duplex ultrasound scan for both study and control populations, to assess graft patency before enrollment into the VGST trial.

We have previously demonstrated that the 6-week scan is essential in identifying grafts at risk of failure, with 33% of grafts displaying a significant flow abnormality at this time. This finding has been reinforced by a number of studies, reporting incidence of vein graft stenosis up to 40%, within the early postoperative period.5,6,7 As well as indicating an early threat to patency, detection of abnormalities early in the life of the vein graft predict the natural history of that graft and the outlook for the limb in the medium and longer term. In this study, the vast majority (89.2%) of grafts that occluded or required intervention exhibited significant flow disturbances by the time of the early postoperative review, although in the majority these flow abnormalities were not severe enough to require intervention and in some patients mild and moderate flow abnormalities regressed during follow-up.

Graft surveillance programs are resource intensive, yet their value among certain groups of patients is indisputable. By defining high- and low-risk groups, resources can be used to target those who will benefit from more intensive surveillance and safely improve the cost effectiveness of vein graft surveillance programs. Mills et al. suggested that grafts with a normal early duplex ultrasound scan require a less concentrated surveillance program, as they exhibit a low incidence of stenosis development or occlusion. Although 59.7% of our patients were categorized as high risk, 40.3% of patients could have been discharged from hospital-based follow-up, at a significant saving to vascular departments’ budgets. It must be remembered that the primary indication for intervention in both this study and in the United Kingdom is critical ischemia, which defines a high-risk group of patients in itself.

The size and quality of the venous conduit available for bypass are important factors affecting the outcome of infrainguinal bypass. Although the use of smaller veins does not invariably result in graft failure, grafts with suboptimal vein are at increased risk of occlusion. Preoperative duplex ultrasound assessment of quality of venous conduit would be a potentially valuable information for inclusion in a predictive model. Regrettably, the quality of vein used was not routinely collected in one of the vascular units involved in the study; therefore it could not be included in the decision tree model.

**Fig. 4. (A)** Kaplan-Meier survival curve showing the primary patency rate in the grafts that have been selected to be high risk versus low risk by the decision tree. **(B)** Kaplan–Meier survival curve showing amputation rates in the grafts that have been selected to be high risk versus low risk by the decision tree.

<table>
<thead>
<tr>
<th>Follow up/ months</th>
<th>Low risk</th>
<th>High risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>299</td>
<td>433</td>
</tr>
<tr>
<td>10 Months</td>
<td>234</td>
<td>314</td>
</tr>
<tr>
<td>20 Months</td>
<td>178</td>
<td>241</td>
</tr>
<tr>
<td>30 Months</td>
<td>122</td>
<td>153</td>
</tr>
<tr>
<td>40 Months</td>
<td>73</td>
<td>101</td>
</tr>
</tbody>
</table>
Multiple reports have indicated that although some cardiovascular risk factors increase the risk of amputation and mortality, they do not directly influence vein graft patency.\textsuperscript{34–38} A decision tree tool has the benefit of creating a method of identifying high-risk grafts using pertinent risk factors, which are then subjected to an intensive duplex ultrasound-based graft surveillance program.

Once the high-risk group is identified, the next stage is to attempt to identify and correct lesions responsible for graft failure. This process is best achieved using intensive duplex ultrasound-based surveillance, as vein graft stenosis is often asymptomatic for most of its natural history. The Kaplan–Meier curves (Fig. 4) clearly illustrate that grafts selected as low risk by the decision tree model have good long-term primary patency and reinforce why intensive graft surveillance should be targeted to the high-risk group. However, within the high-risk group there remains a 50% limb loss rate.

There are well-established standards for the development of clinical decision rules.\textsuperscript{39} The present decision tree fulfills the standards required for the development of a clinical decision rule before implementation. It is recommended that each variable and the rule as a whole be assessed for interobserver agreement. However, each category in our decision tree analysis is highly reproducible as all are objective in nature. Therefore, it is anticipated that different clinicians will categorize patients as high or low risk using this rule with a high degree of accuracy. This is currently being established through external validation in a new population of patients. There are limitations to the decision tree analysis. In this and all recent studies on infrainguinal vein graft surveillance, stenoses thought to threaten graft patency have undergone prophylactic revision. Therefore, the natural history of vein graft stenosis cannot accurately be determined, as all subsequent studies have a selected subgroup of patients and these selected groups may not be fully representative of the population.

CONCLUSION

In conclusion, a clinical decision rule based on readily available clinical data and the findings of significant flow abnormalities on an early postoperative duplex ultrasound scan successfully identifies grafts at high risk of failure. Implementation of such a decision tree will allow the resources of a graft surveillance program to be channeled toward patients and grafts which are at highest risk of failure and avoid a large number of unnecessary duplex ultrasound scans.

Thanks go to the support of the vascular technicians from the Royal Infirmary of Edinburgh and Ninewells Hospital Dundee for the scanning of patients.

REFERENCES

Validation of a Decision Tree to Streamline Infrainguinal Vein Graft Surveillance

Reza Mofidi,1 Olivia M.B. McBride,2 Barnabas R. Green,1 Tracey Gatenby,1 Paul Walker,3 and Simon Milburn,3 Middlesbrough, and Edinburgh, United Kingdom

Background: Duplex ultrasound (DU)–based graft surveillance remains controversial. The aim of this study was to assess the ability of a recently proposed decision tree in identifying high-risk grafts which would benefit from DU-based surveillance.

Materials and Methods: Consecutive patients undergoing infrainguinal vein graft bypass from January 2008 to December 2015 were identified from the National Vascular registry and enrolled in a duplex surveillance program. An early postoperative DU was performed at a median of 6 weeks (range: 4–9 weeks). Grafts were classified into high risk or low risk based on the findings of the earliest postoperative scan and 4 established risk factors for graft failure (diabetes, smoking, infragenicular distal anastomosis, and revision bypass surgery) using a classification and regression tree (CRT). The accuracy of the CRT model was evaluated using area under receiver operator characteristic (AROC) curve.

Results: About 278 vein graft bypasses were performed; 29 grafts had occluded by the first surveillance visit; 249 vein grafts were entered into surveillance. Sixty-four (23%) developed critical stenosis. Overall 30-month primary patency, primary-assisted patency, and secondary patency rates were 71.2%, 77.2%, and 80.1%, respectively. AROC for prediction of graft stenosis or occlusion was 83% (95% confidence interval [CI]: 78–87%). The sensitivity and specificity of the CRT model for prediction of graft stenosis or occlusion were 95% (95% CI: 88–98%) and 52.2% (95% CI: 45–60%).

Conclusions: A prediction model based on commonly recorded clinical variables and early postoperative DU scan is accurate at identifying grafts which are at high risk of failure. These high-risk grafts may benefit from DU-based surveillance.

INTRODUCTION

Despite advances in endovascular techniques, surgical infrainguinal lower limb revascularization remains the gold standard treatment for lower limb ischemia.1 Maintaining patency of infrainguinal vein bypass grafts has been a challenging task for vascular surgeons as these grafts are prone to the development of stenosis which may precipitate failure of the bypass.2,3

Features associated with the development of vein graft stenosis are recognizable by Duplex Ultrasound (DU) scanning which has been widely used for graft surveillance. The rationale for graft surveillance is that correction of stenotic lesions improves graft patency and reduces the risk of amputation.2,3 Nonetheless, DU-based vein graft surveillance remains controversial as repetitive ultrasound examinations over long periods of time is expensive and time consuming, and many of the scanned grafts

1Department of Vascular Surgery, The James Cook University Hospital, Middlesbrough, UK.
2Department of Vascular Surgery, Royal Infirmary of Edinburgh, Edinburgh, UK.
3Department of Interventional Radiology, The James Cook University Hospital, Middlesbrough, UK.

Correspondence to: Reza Mofidi, Department of Vascular Surgery, The James Cook University Hospital, Marton Road Middlesbrough, TS4 3BW, UK; E-mail: rezamofidi@nhs.net
Ann Vasc Surg 2017; 40: 216–222
http://dx.doi.org/10.1016/j.avsg.2016.07.082
© 2016 Elsevier Inc. All rights reserved. Published online: 24 November 2016
do not require any intervention over their surveil-

cidence cycle, thus questioning the validity of routine

canning.\textsuperscript{4} The difficulty, however, is that graft ste-
nosis which may lead to graft failure and possible

limb loss is often clinically asymptomatic but could

detected on DU.\textsuperscript{3,5} Therefore, there remains

much room for rationalization of vein graft sur-

veillance.

Decision tree analysis is a decision support and

machine-learning paradigm that is based on decision

theory. It strives to correctly classify members of a

population based on several dichotomous vari-

ables.\textsuperscript{9,10} Decision (classification) trees are often

used as a clinical decision support tool where they

map clinical and therapeutic decisions and the

consequences of these decisions in a graph or

tree like manner, from which a therapeutic algo-

rhythm can be created. Such a model is ideally suited

to the problem of identifying grafts that are at the

highest risk of developing graft stenosis and failure.

McBride et al.\textsuperscript{11} have designed such a decision tree

that was accurate at stratifying high-risk infrain-

guinal vein grafts that would benefit from DU sur-

veillance from low-risk grafts which could be

followed up clinically. The aim of this study is to

assess the external validity of this decision tree

model in a contemporary population of patients

from a different tertiary vascular surgical department.

\section*{METHODS}

Consecutive patients who underwent lower limb

bypass procedures between January 2008 and

December 2015 in a regional tertiary referral

vascular department were identified from the Na-

tional Vascular Registry (NVR).\textsuperscript{12} The NVR was

established to measure the quality and outcomes of

care for patients who undergo major vascular sur-

gery in the United Kingdom.\textsuperscript{12} Its primary purpose is

to provide comparative figures on the performance of

vascular services in National Health Service

hospitals and local benchmarking and quality

improvement. To that end, it collects detailed infor-

mation about the patients who undergo vascular

procedures, including co-morbidities and athero-
sclerotic risk factors, indications for surgery proced-

ures performed, and outcomes of treatment.\textsuperscript{12}

Anonymized data can be obtained from the NVR

for analysis under strict information governance

guidelines. Data collected on the NVR were cross-

referenced and supplemented by information ob-

tained from the hospital administrative, radiologi-

cal, and laboratory databases. The study was

registered with the audit department of South Tees

Hospitals National Health Service Foundation Trust.

Approval to perform the data analysis was obtained

before the commencement of the study.

Infralingual vein bypass grafts were performed

with a variety of techniques including reversed

and in situ bypass grafts, with the ipsilateral long

saphenous vein being the conduit of choice. If un-

suitable, the contralateral long saphenous vein

was used, followed by the short saphenous and

arm veins, if required. Patients were then enrolled

into a vein graft surveillance program, with each

visit involving DU and the measurement of ankle

brachial pressure index. Based on the DU result,

we used a modification of Rutherford’s recommen-

dations to the Society of Vascular Surgery to classify

stenotic lesions into mild, moderate, or critical ste-

nosis and grafts that were occluded.

Patient demographics, type of operation, conduit,

and follow-up information were recorded in a

computerized database (Excel; Microsoft, Redmond,

WA). Statistical analysis was performed using the

Statistical Package for Social Sciences, version 12

(SPSS, Inc., Chicago, IL). Results were analyzed

and reported in accordance with the reporting stan-

dards of the Society for Vascular Surgery and the

North American Chapter of the International Soci-

ey for Cardiovascular Surgery.\textsuperscript{13} Primary outcome

measures were compliance with vein graft surveil-

lance, stenosis, need for reintervention, graft

patency, and amputation rate. The differences in

patency and limb salvage rates between high-

and low-risk grafts were determined using Kaplan-

Meier analysis.

\section*{DECISION TREE ANALYSIS}

A Classification and Regression Tree (CRT) model

was used to identify patients who were likely to

develop significant vein graft stenosis or occlusion

(within 2 years of ultrasound follow-up), with the

aim of identifying patients who would benefit

from more intensive DU follow-up. The design,

development, and validation of this decision tree

have been previously reported.\textsuperscript{11} Precision Tree\textsuperscript{TM}

(version 7) decision tree analysis software (Palisades

Corporation\textsuperscript{TM}, Middlesex, United Kingdom) was

used to map the decision tree. The topography of

the decision tree used is illustrated on Figure 1.

The ability of CRT model in correctly classifying

infralingual vein grafts into high-and low-risk

grafts was expressed using sensitivity, specificity,

positive, and negative predictive values as well as

area under receiver-operator characteristic curves

(AROC).
RESULTS

Between January 1, 2008, and December 31, 2015, 278 infrainguinal vein grafts were performed. The median age of the patients was 68 years (range 29–89 years). One hundred seventeen patients (42%) were current smokers, 121 (43.5%) had diabetes, and 34 (12.2%) had chronic kidney disease stage 4 or 5. Critical limb ischemia was the primary indication for surgery (n = 226, 81.3%), with the remainder of patients undergoing infrainguinal bypass procedures for intermittent claudication (n = 31, 11%) or popliteal artery aneurysms (n = 21, 7.6%). Two hundred eleven patients were already on statin or started statin treatment during the index episode of care, and 232 patients (83.5%) were on antiplatelet agents, and 21 (7.6%) were fully anticoagulated with warfarin or low-molecular weight heparin on discharge. The median duration of hospital stay was 12 days (range 3–73 days).

Two hundred forty-one grafts (86.7%) had their origin at the common femoral artery, 5 (1.8%) had their origin at the external iliac artery, 18 (6.5%) had the proximal anastomosis from the superficial femoral artery, and 14 (5%) grafts had their origins from popliteal artery. The distal anastomoses were to the above-knee popliteal in 98 bypasses (35.2%), the below-knee popliteal in 88 (31.7%), and crural or pedal arteries in 92 (33%). Two hundred and twenty bypasses (79%) were reversed vein grafts, and 58 (21%) were in situ vein graft bypasses.

The first postoperative vein graft surveillance scan was performed at a median of 6 weeks (range 4–12 weeks). By the time of this scan, 29 grafts (10.4%) had occluded and were excluded from further analysis leaving a further 249 which were entered into graft surveillance.

During vein graft surveillance, 81 grafts (29.1%) developed significant flow abnormalities of which 64 (23%) were classified as or progressed to severe vein graft stenosis. Forty-five grafts (70.3%) were intervened upon using a combination of endovascular (n = 38) and surgical techniques (n = 6) to maintain graft patency; a further 19 grafts were either not intervened upon or occluded before scheduled endovascular or surgical intervention. Figure 2 illustrates DU and angiographic images of vein graft stenosis both prior and after endovascular intervention. Figure 1 illustrates DU and angiographic images of vein graft stenosis both prior and after endovascular intervention. A total of 19 grafts (42.2%) developed recurrent stenosis of which 14 underwent repeat attempt at endovascular correction with technical success (standard balloon angioplasty in 6 grafts, drug eluting balloon angioplasty in 5 grafts, and treatment with a covered stent in 3 grafts) and 2 underwent surgical correction; 2 occluded before surgical intervention, and 1 was not intervened upon.
due to other comorbidities. The median number of interventions on grafts which developed critical graft stenosis was 1 (range: 1–3). Overall 30-month primary patency, primary-assisted patency, and secondary patency rates were 71.2%, 77.2%, and 80.1%, respectively.

**DECISION TREE ANALYSIS**

Two hundred and forty-nine grafts were classified through 5 items into 10 different categories with 2-year stenosis or occlusion rates ranging from 5 to 90%. Figure 1 illustrates the topography of the decision tree and the number of patients assigned to each category.

Area under the receiver-operator characteristic curve (AROC) for prediction of graft stenosis or occlusion was 83% (95% confidence interval [CI]: 78–87%); the sensitivity of the decision tree for prediction of graft stenosis or occlusion was 95% (95% CI: 88–98%). The specificity of the decision tree for prediction of graft stenosis or occlusion was 52.2% (95% CI: 45–60%). The decision tree has a negative predictive value of 96.3% (95% CI: 90–99%), reassuring clinicians, that patients categorized as low-risk have a low incidence of graft failure and limb loss after the early postoperative DU scan and clinical assessment (Table I).

Figures 3 and 4 illustrate the Kaplan-Meier survival curves for primary graft patency and amputation rates for grafts that have been classified as high- and low-risk based on the model. They show that grafts classified as high-risk had significantly lower primary patency rate (log rank = 70, P < 0.001) and were associated with higher risk of amputation (log rank = 42, P < 0.001) during follow-up.
Table I. Indicators of accuracy of classification of the CRT (decision tree) analysis for prediction of development of vein graft stenosis or occlusion within the first 2 years of follow-up

<table>
<thead>
<tr>
<th>Value</th>
<th>Estimated value</th>
<th>95% confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of high-risk grafts</td>
<td>0.29</td>
<td>0.24–0.35</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.95</td>
<td>0.88–0.99</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.52</td>
<td>0.45–0.60</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.45</td>
<td>0.38–0.54</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.96</td>
<td>0.90–0.99</td>
</tr>
<tr>
<td>Positive likelihood ratio</td>
<td>2.01</td>
<td>1.71–2.36</td>
</tr>
<tr>
<td>Negative likelihood ratio</td>
<td>0.079</td>
<td>0.26–0.24</td>
</tr>
</tbody>
</table>

DISCUSSION

Avedis Donabedian, who is considered as the father of health care improvement, defined delivery of health care in terms of the triad of structures, processes, and outcomes. Structure includes the material and human resources as well as the organizational structure within which care is delivered. The process denotes the steps involved in providing care and outcomes are effects of care on the health status of patients and populations. Care pathways involved in the delivery of care to patients with critical lower limb ischemia or disabling claudication fall within the definition of the processes which are followed to deliver safe and effective treatment to this cohort of patients. Care pathways provide a portfolio approach for delivery of care which requires exactly the same steps to be followed in management of each cohort of patients and have been credited with improving patient safety in delivery of health care.

Individualizing patient risk to inform the clinicians and patients as to the relative importance and necessity of each step in the care pathway is a compelling argument especially with regards to identifying conditions which are likely to remain asymptomatic until they cause a seminal event such as vein graft stenosis, or planning prophylactic procedures such as angioplasty, stenting, or surgical revision where the indication for intervention is prevention of a future event, for example graft occlusion and limb loss. Individualizing patient care to suit their condition runs counter to following a rigid vein graft surveillance care pathway which involves the same step in follow-up for each patient who has undergone infrainguinal bypass surgery using autologous vein as conduit. With regards to vein graft surveillance, the situation is further complicated by the fact that most grafts do not develop stenosis, and there is no evidence that beyond the early postoperative scan (performed at around 4–6 weeks), further DU surveillance of vein grafts has any impact on graft patency or amputation rates.

On the other hand, there is compelling evidence that vein graft stenosis can be reliably diagnosed with duplex ultrasound examination and a secondary intervention does significantly improve graft patency and lowers amputation rates in patients who have undergone infrainguinal lower limb revascularization using vein grafts. Rehfuss et al. reported a strong correlation between ultrasound characteristics of vein grafts, in particular, flow and velocity characteristics at the site of stenosis, and the risk of graft failure.

Duplex ultrasound-based graft surveillance is resource intensive and requires repetitive scans of grafts, many of which will not develop graft stenosis during their surveillance cycle. By defining high- and low-risk groups, the decision tree model allows for resources to be focused on those who will benefit from more intensive surveillance and improve the cost efficacy of vein graft surveillance. Mofidi et al. reported that an early postoperative DU scan could predict the natural history of a vein graft throughout the screening cycle. Similarly, Mills et al. suggested that grafts with a normal early DU scan require a less concentrated surveillance, as they exhibit a low incidence of stenosis development or occlusion. However, development of vein graft stenosis or occlusion is multifactorial in nature and the implications of vein graft occlusion vary depending on what the indication for the procedure originally was. This would mean that a screening test to identify high-risk grafts would also need to be multifactorial in nature.

There are well-established standards for the development of clinical decision rules. This would include that the model is validated on an external cohort of patients before implementation. The previous publication by McBride et al. developed the classification and regression tree model for the identification of high-risk grafts, whereas this publication completes the cycle by externally validating their proposed model in an unselected cohort of patients who were treated in a different hospital in a different country. It is also recommended that each variable is assessed for interobserver agreement. However, each category in the decision tree analysis is highly reproducible as all are subjective in nature. Therefore, it is anticipated that different clinicians will categorize patients as high- or low-risk using this rule with a high degree of accuracy.

In conclusion, this study validates the previous findings of McBride et al. that a decision tree based...
Fig. 3. Kaplan-Meier survival curve comparing primary assisted patency (y axis) of grafts which were predicted to be high risk (red) using the decision tree model compared with low-risk grafts (black), (log rank = 70, \( P < 0.001 \)).

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>10 Months</th>
<th>20 Months</th>
<th>30 Months</th>
<th>40 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>95</td>
<td>83</td>
<td>76</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>High risk</td>
<td>154</td>
<td>79</td>
<td>63</td>
<td>41</td>
<td>28</td>
</tr>
</tbody>
</table>

Fig. 4. Kaplan-Meier survival curve comparing amputation rates in patients who had undergone infrainguinal bypass grafts which were predicted to be high risk (red) using the decision tree model compared with amputation rates in low-risk grafts (black), (log rank = 42, \( P < 0.001 \)).
on readily available clinical data which are available from the NVR and the findings of the early postoperative DU successfully identifies grafts at high risk of failure. Implementation of such a decision tree will allow the resources of a graft surveillance program to be channeled toward patients and grafts which are at the highest risk of failure.

REFERENCES


Prediction of the Exact Degree of Internal Carotid Artery Stenosis Using an Artificial Neural Network Based on Duplex Velocity Measurements

Reza Mofidi, MB, MCh, FRCSI,1 Thomas I. Powell, MB, BCh, FFRRCSI,2 Anthony Brabazon, BCom, MSc, MBA,3 Denis Mehigan, MCh, FRCSI, FRCS(Ed),1 Stephen J. Sheehan, MD, FRCSI,2 Donal P. MacErlaine, FFRRCSI, FRCR,2 and Thomas V. Keaveny, MCh, FRCSI,1 Dublin, Ireland

Duplex ultrasound criteria use a combination of velocity measurements to evaluate internal carotid artery (ICA) stenosis. These evaluations divide ICA stenosis into broad categories. The aim of this study was to design an artificial neural network (ANN) capable of predicting the exact degree of ICA stenosis based on duplex velocity measurements. Consecutive patients with significant carotid atherosclerosis underwent carotid duplex ultrasound and angiography. Peak systolic and end-diastolic velocities in the ICA and common carotid artery were measured. Multilayered perceptron ANNs were constructed and trained to predict the degree of ICA stenosis and band the degree of ICA stenosis into 10% intervals based on these measurements. The accuracy of the ANN models in predicting the degree of ICA stenosis and classifying the ICA stenosis was compared with the angiographic degree of ICA stenosis and duplex velocity criteria. A total of 208 carotid bifurcations were studied. ANNs were able to accurately predict the degree of angiographic ICA stenosis (R² = 0.9374, p < 0.0001) and band the ICA stenosis into the predefined 10% intervals [sensitivity 97.3% (95% CI 90.7-99.3), specificity 97.7% (95% CI 93.6-99.2), accuracy 97.5%]. The ANN model was more accurate [discriminant power (DP) = 4.11] in banding the degree of ICA stenosis than duplex velocity criteria (DP = 1.67) (p < 0.05). The accuracy of the ANN in correctly identifying >70% ICA stenosis was 98.4% [sensitivity 96.4% (95% CI 93.8-99.3), specificity 98.7% (95% CI 93.4-99.8), DP = 4.21]. ANNs can accurately predict the degree of ICA stenosis. With further refinement, ANNs could replace velocity criteria in the assessment of ICA stenosis using duplex ultrasound.

INTRODUCTION

Randomised trials have confirmed the value of carotid endarterectomy compared with the best medical treatment for patients with significant extracranial internal carotid artery (ICA) stenosis.1,2 These trials have required confirmation of the degree of stenosis with contrast angiography, which has been used as the gold standard for the measurement of ICA stenosis and to confirm the presence of ICA occlusion. However, angiography has a significant morbidity and a low but significant mortality rate.3 As the technology has improved and experience has broadened, duplex ultrasound
scanning has been found to provide sufficient information for clinical decision making in patients with ICA stenosis. This is performed by obtaining peak systolic (PSV) and end-diastolic velocities (EDV) at the stenosing lesion, ICA/common carotid artery (CCA) velocity ratios, and subjective assessment of the Doppler waveform at the site of maximal stenosis and the B-mode appearance of the atherosclerotic plaque. Cut-off values for these measurements are used to define the degree of ICA stenosis. Whilst these cut-off values are calculated using receiver operator characteristic curves, the combination of these measurements is based on empirical selection, with not enough statistical data to support which values best represent the degree of ICA stenosis, especially in the presence of confounding variables such as contralateral ICA occlusion. With the increase in use of duplex ultrasound, there has been a proliferation in the number of duplex velocity criteria for measurement of the degree of ICA stenosis. These classifications vary between laboratories both in their method of calculation and in the range of degrees of stenosis used to describe each stratification or band.

Banding of the degree of ICA stenosis using duplex criteria usually involves classification of the degree of ICA stenosis into 20% bands. Earlier broad categories did not allow an adequate estimation of the degree of ICA stenosis to provide all the information needed for clinical decision making in accordance with the Asymptomatic Carotid Atherosclerosis Study (ACAS) and the North American Symptomatic Carotid Endarterectomy Trial (NASCET). A recent paper by Filis et al. suggested that it is possible to accurately band the degree of ICA stenosis into 10% intervals for stratifying 50-100% carotid stenosis. Although this duplex velocity criterion does not estimate the exact degree of ICA stenosis, it provides the closest clinical method with which other models attempting to achieve a similar degree of accuracy can be compared.

The aim of this study was to examine the hypothesis that an artificial neural network (ANN) based on the most commonly used duplex velocity measurements is capable of measuring the degree of ICA stenosis and to evaluate its performance in correctly classifying the degree of ICA stenosis with a comparable, clinically used method.

ANNs are a family of computational algorithms which are modeled on the capabilities of the human nervous system. They permit recognition of patterns in data that cannot be detected with linear statistical analysis.

The neuron (node) is the basic computational unit of an ANN. It receives a variety of inputs from other neurons through connections that resemble synaptic structure and has a binary (all or nothing) output. This output is determined by the sum of the inputs as well as the weight (synaptic strength) attached to these input variables. The weight attached to each node is altered through a period of familiarization “training” until an optimal weight which best describes the influence of the data set, represented by the node, is reached. Each of the input nodes corresponds to a single input variable. Initial values are set for the “weights” associated with each link in the network. Input data for which an output is known are presented to the network. If the predicted output from the model does not equal the known output, the weights within the network are changed so as to narrow this difference. This process continues until the prediction errors are minimised. Once the network is trained and validated, it may be used on unseen data for prediction or classification purposes.

A wide variety of neural network designs with varying degrees of complexity have been described. The simplest and most commonly used form is the multilayered perceptron (MLP). The structure of this neural network model is illustrated in Figure 1.

METHODS

Between April 1998 and April 2000, consecutive patients who were admitted for assessment of ICA stenosis were assessed. All patients underwent carotid duplex ultrasound scanning as well as intra-arterial digital subtraction angiography (IADSA) for assessment of the degree of extracranial carotid artery stenosis during the same hospital stay (within 3 days). At the time, angiography was routine practice in our department for patients with severe ICA stenosis (prior to carotid endarterectomy) or who had recent ocular or hemispheric symptoms together with a suspected ICA occlusion on previous duplex scanning.

Carotid Duplex Scanning

Duplex ultrasound imaging was performed using an Acuson 128 XP duplex ultrasound system with a 7.5 MHz linear array probe (Acuson, Mountain View, CA) and a 60-degree angle of insonation. PSV and EDV were measured through spectral analysis at the stenotic portion of the ICA. Representative recordings of PSV and EDV were also
performed in a nonstenotic portion of the distal CCA. ICA stenosis severity was graded using haemodynamic evaluation of index stenotic vessels based on standard grading criteria. These evaluations were performed retrospectively by two experienced observers, who at the time of evaluation were unaware of the identity of the patient, clinical findings, and the angiographic measurement of the degree of ICA stenosis.

**Angiography**

Standard IADSA was performed by retrograde catheterization via the femoral arteries. Two projections of carotid bifurcation were obtained routinely. Digitally magnified views were used to obtain the angiographic measurements.

Blinded angiographic evaluation was performed by two observers (R.M. and T.P.). The diameter of the ICA at the site of maximal stenosis and the diameter of the poststenotic ICA were measured using precision Vernier calipers to the nearest 0.1 mm. The degree of stenosis was obtained by comparing the residual angiographic lumen with the diameter of the normal distal ICA.

Interobserver agreement in the measurement of ICA stenosis was assessed using linear regression and Bland-Altman analysis.

**Neural Network Design**

MLP ANN models with a back-propagation circuit were constructed using Neuro-Solutions™ version 4 software (NeuroDimension, Gainesville, FL). This neural network design was selected because of its good analytical power and relative simplicity of design. The MLP models which were developed had six input nodes, consisting of entries corresponding to PSV and EDV in the ICA and prestenotic CCA, presence of contralateral ICA occlusion, and a bias node. The bias node performed a function similar to that of a regression constant in a standard linear regression model. The neurons in the hidden layer of these ANNs used a sigmoid activation function, which ensures that node output will lie in the range 0-1. The input data did not undergo any preprocessing steps prior to input into the neural network apart from randomisation of data sets. Two sets of ANN models were created (one of which had as its output the actual degree of ICA stenosis, whilst the other had an output which corresponded to the degree of ICA stenosis stratified to bands of predefined 10% intervals).

The available data were randomly divided into two mutually exclusive data sets. One of these was used to train the model; the other was used to validate the model by testing the out-of-sample performance of the constructed MLP (assessment of the accuracy of predictions of the trained ANN). The data sets were organised into eight different randomly selected re-cuts (reruns) of input data, each used to train eight identical MLP models (four identical MLPs for prediction of actual ICA stenosis and four to stratify the stenosis according to 10 bands).

**Training.** The input variables used were PSV and EDV in ICA as well as CCA and presence of contralateral ICA occlusion. All of the input variables apart from the presence of contralateral ICA occlusion were continuous. Contralateral occlusion was a categorical variable. Sixty percent of the data set was used to train the neural network. The
training rule was back-propagation of error. This adjusts the weights associated with each node during the training process until the mean squared error was reduced to a minimum.

Validation. Forty percent of the available data set was held out and not used for training of the neural network model. These data were used to evaluate the ability of the fully trained neural network to predict the degree of angiographic stenosis from which the neural network had been blinded. The mean squared error was used to express the variation of the degree of predicted ICA stenosis compared to the angiographic measurement of ICA stenosis. A classification matrix was used to assess the predictive ability of the neural network. The accuracy of the neural network in correctly classifying the degree of ICA stenosis was compared with duplex velocity criteria alone.

Statistical Analysis

The degree of ICA stenosis predicted by the ANN was compared with the angiographic measurements using linear regression and Bland-Altman analysis.

The accuracy, sensitivity (Sen), specificity (Spec), and likelihood ratios for positive and negative tests were calculated for duplex velocity criteria and the ANN. To compare the duplex velocity criteria with the ANN in correct banding of the degree of ICA stenosis, a measure of the discriminant power (DP) of a test was used:

\[ \text{DP} = \left( \sqrt{3} \right) / \pi \left\{ \ln [\text{Sen} / (1 - \text{Spec})] + \ln [\text{Spec} / (1 - \text{Sen})] \right\} \]

McNemar’s test was used to compare the relative accuracy of the ANN model and duplex velocity criteria in correctly stratifying the degree of ICA stenosis to 10% intervals.

RESULTS

Two hundred and eight carotid bifurcations in 104 patients were studied. Eighty-three patients had ipsilateral symptoms prior to presentation (hemispheric, 51; ocular, 27; both symptoms, 5), and 21 patients were asymptomatic. The mean age of subjects was 70.1 years (standard deviation = 10.11); 38 were female and 66 were male. No patient was excluded due to inadequate angiography or duplex ultrasound.

Seventeen carotid arteries were occluded. Carotid duplex examination was able to identify occluded ICA in all patients. In the remaining ICA bifurcations, the median degree of ICA stenosis was 64% (range 0-99%). One hundred and forty-one carotid bifurcations (68%) studied possessed \( \geq 50\% \) degree of ICA stenosis and 114 (55%) had \( \geq 70\% \) stenosis, as measured by angiography.

Interobserver Agreement

There was good agreement between the two observers in the angiographic measurement of degree of ICA stenosis \( (R^2 = 0.95) \). Bland-Altman analysis confirmed the high degree of interobserver reproducibility of the angiographic measurement of the degrees of ICA stenosis (Fig. 2).

Interobserver variability in determining the correct banding for the degree of ICA stenosis was acceptable \( (\kappa = 0.79) \).
Duplex Velocity Criteria and the Degree of ICA Stenosis

The overall agreement between the results of duplex ultrasound velocity measurement and angiography in determining the correct banding for the degree of ICA stenosis was 91.8% (191/208), with a sensitivity of 89.7% [95% confidence interval (CI) 82.8-94] and specificity of 93.5% (95% CI 86.5-97) (DP = 1.67).

The accuracy of duplex ultrasound criteria at correctly identifying >70% ICA stenosis was 95.2%, with sensitivity of 96.3% (95% CI 90.9-98.6), specificity of 94% (95% CI 87.5-97.3), and DP of 3.3. Carotid duplex examination was also able to identify over 50% ICA stenosis with reasonable accuracy [sensitivity 96.38% (95% CI 89.9-97.5), specificity 91.6% (95% CI 82.8-96.1), and accuracy 94.7% (DP = 3.09)].

Neural Network Analysis

Two hundred and eight data vectors were randomly divided as follows: 124 (59.6%) were used to train the neural network model, and 84 (40.4%) were used for model validation and out-of-sample testing.

Following completion of training, the ANN model was able to predict the degree of ICA stenosis with a high degree of accuracy (mean squared error 0.012); cross-validation confirmed the validity of the ANN model at predicting the degree of ICA stenosis (Fig. 3). Good correlation was observed between the degree of ICA stenosis predicted by the ANN and that measured angiographically ($R^2 = 0.9374, p < 0.0001$) (Fig. 4). Bland-Altman analysis revealed a high degree of agreement between the results of ANN and angiographic measurements (Fig. 5). Furthermore, the ANN model was able to accurately band the ICA stenosis according to the predefined 10% intervals with a sensitivity of 97.3% (95% CI 90.7-99.3), a specificity of 97.7% (95% CI 93.6-99.2), and an overall accuracy of 97.5%. The ANN model was significantly more accurate at correctly banding the degree of ICA stenosis. The DP of ANN was 4.11, while that of duplex velocity criteria alone was 1.67 (McNemar’s test, $p < 0.05$).

The accuracy of the ANN at correctly identifying >70% ICA stenosis was 98.4%, with sensitivity of 96.4% (95% CI 93.8-99.3), specificity of 98.7% (95% CI 93.4-99.8), and DP of 4.21. This was comparable to duplex velocity criteria ($p = $ nonsignificant). The accuracy of the ANN at correctly identifying >50% ICA stenosis was 97.4%, with a sensitivity of 98.7% (95% CI 93.8-99.3 and) specificity of 98.2% (95% CI 90.4-99.7), which was not significantly better than the duplex velocity criteria ($p = $ nonsignificant). Table I illustrates the comparison between duplex velocity criteria and ANN at identification of >50% and >70% stenosis as well as correct banding of ICA stenosis for >50% ICA stenosis.

The use of ANN was associated with altered clinical decision in 2.9% of cases (5/208) compared with angiographic findings. This was not significantly different from the duplex velocity criteria, use of which was associated with altered clinical decision in eight carotid bifurcations (3.8%) compared with angiography ($p = $ NS).

DISCUSSION

Duplex scanning is the principal method of imaging carotid bifurcation disease, owing to its noninvasive nature and widespread availability. It has been shown that carotid endarterectomy can be...
safely performed without preoperative angiography.4,19

Proliferation of duplex velocity criteria has resulted in lack of uniformity in the measurement of ICA stenosis using duplex imaging, in contrast to the standardized nature of angiographic assessment of ICA stenosis during randomised controlled trials, upon which the efficacy of carotid endarterectomy is based.1,2 The randomised controlled trials which form the evidence base for the management of symptomatic and asymptomatic carotid atherosclerosis have identified different threshold values for ICA stenosis as representing the indication to perform carotid endarterectomy.20-22 Several previous authors have changed their velocity criteria in order to comply with these studies.23,24

With respect to symptomatic patients, both NASCET and the European Carotid Surgery Trial (ECST) have demonstrated a gradient effect of benefit with increasing degree of stenosis.20,22 Therefore, in patients with moderate degrees of stenosis (50-70%), knowing the exact degree of stenosis in conjunction with preoperative clinical risk factors may influence decision making with regard to surgery.12 Such information is of value in follow-up of patients who are treated conservatively and do not undergo carotid endarterectomy.

![Fig. 4. A plot of the angiographic degree of ICA stenosis (x axis) and ANN prediction of ICA stenosis in the validation sample (y axis).](image1)

![Fig. 5. Bland-Altman analysis applied to the comparison of predicted degree of ICA stenosis by ANN and the angiographic degrees of ICA stenosis.](image2)
Table I. Comparison between accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of duplex velocity criteria and ANN at identification of >50% and >70% stenosis as well as correct banding of ICA stenosis

<table>
<thead>
<tr>
<th>Modality</th>
<th>Identification of &gt;70% stenosis (p = NS)</th>
<th>Identification of &gt;50% stenosis (p = NS)</th>
<th>Correct 10% banding for stratifying 50-100% stenosis (p &lt; 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>Duplex velocity criteria</td>
<td>95.2</td>
<td>96.3 (90.9-98.6)</td>
<td>94 (87.5-97.3)</td>
</tr>
<tr>
<td>ANN</td>
<td>98.4</td>
<td>96.4 (93.8-99.3)</td>
<td>98.7 (93.4-99.8)</td>
</tr>
</tbody>
</table>

NS, nonsignificant; DOR, diagnostic odds ratio (95% CI in parentheses).

In order to better stratify the degree of ICA stenosis measured with duplex ultrasound, Filis et al.12 developed velocity criteria with 10% interval categories for >50% ICA stenosis. They reported that such criteria have similar accuracy to the old 20% intervals. Using the duplex velocity criteria described by Filis et al., we found good degree of accuracy and interobserver agreement for the stratification of 50-100% ICA stenosis (into 10% bands); therefore, it was adapted retrospectively by us for comparison against the ANN model.

The use of ANNs represents a novel approach in the assessment of degree of ICA stenosis. Neural network algorithms have been applied to a variety of clinical conditions where complex relationships within the data set preclude the use of conventional linear statistical analysis.25 Accurate measurement of the degree of ICA stenosis using velocity of blood flow contains an inherent degree of variability. This variability increases with progressively lower degrees of stenosis in the ICA.26,27 It would be impossible to account for this variability through conventional statistical models. ANNs can be used to separate background noise (variability) from information embedded in the data set. ANN algorithms can be built into new duplex devices. In this way, after selection of the appropriate sample volume by the operator, the machine could automatically calculate the duplex parameters for predicting the probability of threshold ICA stenosis or measure the exact degree of ICA stenosis.

Our results suggest that ANNs have an acceptable degree of accuracy in predicting the exact degree of ICA stenosis and represent a modest but significant improvement in classification of ICA stenosis into 10% bands. Carotid duplex criteria have been proven to be excellent tools in identification of threshold ICA stenosis at the clinically significant 70% level used for identification of surgical candidates amongst patients with symptomatic disease5-8 and 60% stenosis screening for asymptomatic carotid artery stenosis.23,24 In this study, we did not find a significant difference in the accuracy of ANN in the identification of patients with clinically significant ICA stenosis compared with carotid duplex criteria.

Other approaches have been adapted as alternatives to duplex velocity criteria. Zbornikova and Johansson28 found that multivariate regression based on PSV, EDV, and pulse pressure was more accurate than univariate analysis based on each variable at identifying >50% ICA stenosis and suggested that this could replace duplex velocity criteria in the identification of candidates for carotid endarterectomy.28 More recently, Hwang et al.9 used a multiple regression model to predict the exact degree of ICA stenosis using duplex velocity measurements as independent input variables. They found a reasonable correlation between the predicted and actual degrees of ICA stenosis. Apart from the current study, the publication by Hwang et al.9 is the only study in the literature which attempts to measure the exact degree of ICA stenosis using duplex vari-
ables. Their work suggests that a linear regression model based on the equation ICA stenosis = 20.2PSV -7.4EDV + 0.4SCR + 8.5SCR [where SCR is (peak) systolic carotid ratio] can be used to predict the exact degree of ICA stenosis, albeit with a moderate degree of correlation with angiographic controls (correlation coefficient = 0.75). Such a model, whilst novel, does not offer an improvement over duplex velocity criteria\textsuperscript{12} or an ANN.

Direct measurement of the residual lumen using a combination of B-mode ultrasound and colour flow imaging components of duplex investigation has been used to measure ICA stenosis.\textsuperscript{29,30} These measurements have been found to correlate well with angiographic measurements\textsuperscript{29} and examination of endarterectomy specimens.\textsuperscript{30} However, it is important to point out that B-mode ultrasound and angiography measure the degree of vessel stenosis in different ways: duplex ultrasound assesses stenosis by percentage area reduction, whereas angiography measures the percentage reduction in luminal diameter. Therefore, despite close correlation, the two values are not interchangeable.

In our study, data analysis was performed using a neural network with a simple design and five input variables (PSV and EDV measurements in the ICA, CCA, and the presence of contralateral ICA occlusion). The reason for limiting the number of input variables and using simple neural network topography was to avoid complexity in the design of the neural network and the resultant data analysis. Nonetheless, it was able to predict the degree of ICA stenosis with a high level of accuracy.

We did not analyse the influence of factors such as heart rate, presence of cardiac arrhythmias, and systemic blood pressure, which are known to be responsible for alteration of haemodynamic findings in the carotid bulb region. These findings are hard to characterise in the linear assessment of the degree of ICA stenosis but would be ideally suited to ANN analysis. Their addition would significantly increase the complexity of the neural network model, which would require a larger pool of subjects to train and validate. However, it is likely that the future introduction of these data into the ANN model would improve its predictive ability in the assessment of ICA stenosis.

Previous authors have suggested that duplex velocity profiles may overestimate the degree of ICA stenosis in the presence of contralateral occlusion.\textsuperscript{31,32} This is difficult to adjust for when duplex velocity criteria are used alone. However, such an adjustment is possible in an ANN model in the form of another input variable (node).

This study suggests that using ANN analysis it is possible to predict the exact degree of ICA stenosis using Doppler velocity parameters. We believe that with further refinement and the addition of new input variables ANNs can be used to perform this task in a standardized manner.

REFERENCES


Prediction of survival from carcinoma of oesophagus and oesophago-gastric junction following surgical resection using an artificial neural network

R. Mofidi*, C. Deans, M.D. Duff, A.C. de Beaux, S. Paterson Brown

Department of Surgery, Royal Infirmary of Edinburgh, Old Dalkeith Road, Edinburgh EH16 4SA, UK

Accepted 27 February 2006
Available online 18 April 2006

Abstract

Aim: The aim of this study was to assess the ability of artificial neural network (ANN) in predicting survival in patients undergoing surgical resection for carcinoma of oesophagus and oesophago-gastric junction.

Methods: From January 1995 to August 2004 patients who underwent surgery for oesophageal and gastric carcinoma were identified. Biographical data, body mass index and pathological minimal cancer dataset were used to design an ANN. Post-operative survival was assessed at 1 and 3 years. Sixty percent of data was used to train and validate the ANN and 40% was used to evaluate the accuracy of trained ANN in predicting survival. This was compared with Union Internacional Contra la Cancrum UICC TNM classification system.

Results: Two hundred and sixteen patients underwent resectional surgery for oesophageal and OGJ carcinoma. The accuracy of the ANN in predicting survival at 1 and 3 years was 88% (sensitivity: 92.3%, specificity: 84.5%, DP = 2.3) and 91.5% (sensitivity of 94.61%, specificity: 88%, DP = 2.72), respectively. These figures were significantly better than 1- and 3-year survival predictions using the UICC TNM classification system 71.6% (sensitivity of 66.4%, specificity: 75.5%, DP < 1) and 74.7% (sensitivity of 70.5%, specificity: 74.9%, DP < 1), respectively (P < 0.01).

Conclusion: ANNs are superior to the UICC TNM classification system in correlating with survival following resection of carcinoma of oesophagus and OGJ junction and can become valuable tools in the management of patients with oesophageal carcinoma.

Keywords: Oesophageal carcinoma; Survival; Artificial neural network

Introduction

Union International Contra Cancrum (UICC) TNM classification system remains the most widely used method of staging for patients with oesophageal and OGJ carcinoma. The TNM classification system is based on anatomical extent of the disease, therefore, its ability as a predictive model is limited. Apart from the TNM stage of the lesion, other factors, which are amenable to pre-operative assessment, are known to influence survival following resection with curative intent. These include tumour length measured on contrast swallow, tumour differentiation and the presence of submucosal lymphocytic infiltration as well as involvement of the resection margins. Addition of these variables to a predictive model for oesophageal carcinoma based on conventional statistical models would significantly increase its complexity and limit its clinical utility.

ANNs have been increasingly used in medicine as decision support aids, for analysing radiological and histological images and to predict outcome from using complex multivariable prognostic indicators such as patients with cancer. Artificial neural networks ANN are a family of data analysis algorithms. Their structure and function is designed to resemble biological nervous systems. This novel design provides the ability to identify patterns within data that are associated with a particular outcome from a large population and apply the findings to an individual subject. ANNs are composed of data processing units, which like their biological counterparts are named neurons. The units are heavily interconnected to each other by weighted connections similar to synapses. During the process of training the ANN is introduced to a series of defining characteristics, which are associated with a particular outcome. By

---

* Corresponding author. Tel.: +44 131 2257346; fax: +44 131 2282661. E-mail address: rmofidi@doctors.net.uk (R. Mofidi).
performing repetitive cycles of training the weights associated with the synaptic connection are adjusted in order achieve the best predictive ability, after this the ANN should be able to accurately perform predictions on input data to which it was blinded during training. ANNs are useful in performing for interpreting data where the data is abundant, contradictory, biased or complex.¹⁹

The aim of this study was to assess the performance of an ANN based in a population of patients with carcinoma of the oesophagus and OG junction who have undergone surgical resection with curative intent and assess its ability in predicting 1- and 3-year disease free survival.

Methods

All patients who had undergone resection of carcinoma of the oesophagus and gastro–oesophageal junction in a specialist oesophago-gastric surgery unit between January 1995 and August 2004 were identified.²⁰ The surgical approach included a level 2 abdominal lymphadenectomy and extended en bloc mediastinal lymphadenectomy in all patients.

Neural network design

Three-layered multilayer perceptron (MLP) ANN models, with back propagation circuit were constructed using Neuro-Solutions²⁹ version-4 neural network software (NeuroDimension, Inc., Gainesville, FL, USA). This is a simple neural network design in which the neurons are arranged in parallel layers and each layer is fully connected to the previous layer through synaptic connections, leading to a single predictive outcome. The ANNs were trained through back propagation of error, which is a process by which the error of prediction is minimized by adjusting the weights associated with the synaptic connections in the hidden layers of the ANN (Fig. 1).

Two sets of artificial neural network models were created one set of ANNs was used to predict disease survival at 1 year, the other with disease free survival at 3 years. The input data did not undergo any pre-processing steps prior to input into the neural network. The available data were randomly assigned into two mutually exclusive datasets. One of these was used to ‘train’ the model, a second used to assess the accuracy of predictions of the trained ANN. Sixty percent of the available data were used to train each neural network. Forty percent of the data was withheld from the ANN during the training process and was used for cross validation of the ANN. In order to minimize bias based on selection of cases for training and evaluation datasets, panels of 10 identical ANNs were used for each prediction (disease free survival at 1 and 3 years). These ANNs were trained and evaluated independently by different, randomly selected, partitions of the dataset. The values pertaining to the accuracy of prediction were reported as average output for the ANN panel together with 95% confidence intervals. The number of training iterations (runs) was kept to eight the training process was considered complete at the end of

---

Figure 1. A simplified diagram depicting the structure of a multilayered perceptron (MLP). Neurons are arranged in layers (hidden layers) and receive synaptic input from the previous layers of neurons or input variables. Each neuron (represented by a circle) has a transfer function, which is equivalent to regression constant in conventional statistical models. The neuron converts the sum of weighted inputs into a binary (all or nothing) output. During the training process the weights associated with each synaptic connection are adjusted through the back propagation of error.
eight iterations, if the mean squared error of predictions was less than 0.025.

**Variable selection**

The optimum variables used to construct the artificial neural network were selected from 42 potential input variables available to the ANN using sensitivity analysis. Sensitivity analysis is a method for identifying the predictive value of input variables in formulating a correct outcome prediction. During the process of sensitivity analysis, the network learning is disabled such that the weights associated with the synaptic connections are not affected. The inputs to the network are shifted slightly and the corresponding change in the output is reported as a percentage. The values that cause the highest percentage of change are selected to create the model.

Using sensitivity analysis, input variables are categorized according to their ability in influencing outcome. The optimal number of variables for developing the ANN was assessed during backward variable selection, a process whereby the variables with the least predictive value are consecutively removed and the accuracy of prediction of the ANN model is evaluated by receiver operator characteristic analysis, until the ANN with the optimum accuracy of prediction is reached.

The predictive power of the UICC TNM classification system in assessing disease free survival at 1 and 3 years was analysed using stepwise linear discriminant analysis. The variables used to create this model are the T, N and M stage as well as clearness of the resection margins (R0: clear margins, R1: microscopically involved margins, R2: macroscopic residual disease).

The accuracy, sensitivity, specificity, positive and negative predictive values as well as likelihood ratios were calculated for the UICC TNM staging system and ANN as predictors of disease free survival. To compare the TNM classification with ANN in predicting disease free survival a measure of the discriminant power (DP) of a test was used, where:

\[
DP = (\sqrt{3}/(\ln(\text{SEN}/1 - \text{SPEC}) + \ln(\text{SPEC}/1 - \text{SEN})))
\]

where LN, natural logarithm; SEN, sensitivity; SPEC, specificity of each of the tests.

McNemar’s test was used to compare the relative accuracy of the ANN model and UICC TNM staging as predictors of disease free survival following potential curative resection.

**Results**

**Patient population**

Two hundred and sixteen patients underwent resectional surgery for oesophageal and OG junction carcinoma of whom 155 were male and 61 were female. Median age of the patients was 64 years (range: 30–82). Median ASA grade and BMI were 2 (range: 1–3) and 23 (range: 14–36), respectively. One hundred and sixty-one patients underwent primary surgical resection of the lesion and 55 received neoadjuvant, cisplatin and 5-flourouracil based chemotherapy prior to the definitive surgical procedure. One patient underwent Argon laser ablation of an early adenocarcinoma of OG junction followed by subsequent definitive surgical resection. One-year follow-up data was available in 198 patients and 126 patients had 3-year follow-up data available.

**Histological findings of the resected specimen**

One hundred and thirty-six patients presented with adenocarcinoma and 73 patients with squamous cell carcinoma, seven patients had other histological types (three small cell carcinoma, four undifferentiated).

Table 1 describes the pathological feature of the lesions based on the examination of the histological specimen. A median of 21 (range: 4–44) lymph nodes were retrieved from each specimen for analysis. The median number of histologically positive nodes was 2 (range: 0–21). Post-operative in hospital mortality was 15 patients. Median follow-up duration was 27 months (range: 2–64). The Kaplan–Meier survival curves based on the UICC classification are shown in Fig. 2.

**Artificial neural network analysis**

Table 2 illustrates the sample size and composition of training and evaluation datasets for the panel of neural networks used to predict disease free survival at 1 and 3 years. Using sensitivity analysis 14 input variables were
selected to construct the ANN. Table 3 illustrates the input values which were selected through sensitivity analysis to construct the ANN as well as their sensitivity for prediction of disease free survival at 1 and 3 years. Fig. 3 shows the change in the accuracy of area under ROC curve in response to the reduction in the number of input variables.

The process of ANN training was smooth with satisfactory minimization of MSE. The mean squared error (MSE) reached at the completion of training of the ANN for prediction of survival at 1 year was 0.0082, whilst the final training MSE for prediction of 3-year survival was 0.006. Fig. 4 shows the learning curve of the artificial neural network used for prediction of survival at 3 years with minimization of mean squared error during training process.

ANN was able to accurately predict disease free survival at 1 year (accuracy: 88%, DP = 2.3). This was significantly more accurate than the 1-year survival prediction using the UICC TNM classification system (accuracy: 71.6%, DP < 1) \( (P < 0.01) \) (Table 4).

The ANN panel was significantly more accurate at predicting survival at 3 years after surgery (accuracy: 91.5%, DP = 2.72) than the UICC TNM classification system (accuracy: 74.7%, DP < 1) \( (P < 0.05) \) (Table 4).

### Discussion

The use of artificial neural network is a relatively novel approach for the prediction of outcome in patients with carcinomas of oesophagus and OG junction. In this study,

**Table 2**
The sample size and composition of training and out of sample testing sets for the panel of neural networks used to predict survival at 1 and 3 years.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th></th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-Year</td>
<td>3-Year</td>
<td>1-Year</td>
</tr>
<tr>
<td>( N )</td>
<td>119</td>
<td>73</td>
<td>79</td>
</tr>
<tr>
<td>Age (range)</td>
<td>64 (41–80)</td>
<td>64.5 (30–82)</td>
<td>62 (36–78)</td>
</tr>
<tr>
<td>Stage I</td>
<td>15</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Stage II ( A )</td>
<td>32</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Stage II ( B )</td>
<td>14</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Stage III</td>
<td>55</td>
<td>33</td>
<td>35</td>
</tr>
<tr>
<td>Stage IV</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>( R_0 )</td>
<td>93</td>
<td>55</td>
<td>61</td>
</tr>
<tr>
<td>( R_1 )</td>
<td>19</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>( R_2 )</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Survival (%)</td>
<td>85</td>
<td>38</td>
<td>56</td>
</tr>
</tbody>
</table>

Sensitivity of each variable for 1- and 3-year survival. These variables were selected through sensitivity analysis. These variables were selected through sensitivity analysis.

---

Figure 2. Kaplan–Meier survival curves based on the UICC classification (log-rank test, \( P < 0.01 \)).

Figure 3. Changes in the accuracy of the ANN survival at 1-year (red line) and 3-year (black line) (shown as the area under ROC curve), plotted against the number of input variables used to create the ANN. With the reduction in the number of variables there is a progressive increase in the accuracy of the ANN (Y-axis) until the optimum number of variables is reached after which further reduction of the input variables results in diminished accuracy of ANN. The error bars indicate 5% confidence intervals.
the use of an artificial neural network with a simple design using clinical and pathological variables, which are readily available clinically was associated with significant improvement in early and 3-year outcome predictions compared with UICC TNM system. Artificial neural networks have been applied to other neoplastic processes and have been found to compare favorably with conventional statistical models or clinical judgment.\textsuperscript{14}

Superior accuracy of ANN models is partly related to superiority of analytical model and partly to the increased number of input variables used to develop the ANN. Most clinical analytical models such as TNM staging are heavily weighted towards the histopathological interpretation of the anatomical extent of disease. The ANN model in this study selected the appropriate input variables using sensitivity analysis. This by itself may be responsible for improved accuracy. A large-scale study by examining cancer specific 5-year survival in breast and colorectal carcinoma found that the ANN models constructed using TNM staging system alone significantly improved the predictive accuracy of TNM staging system.\textsuperscript{27}

Despite the fact that ANNs are well suited to outcome prediction, to date their routine clinical use has been limited to the management of prostatic carcinoma.\textsuperscript{28} Some of the resistance to the use of ANNs in the clinical setting relates to the fact that their introduction as predictive models involves addition of new nomenclature and methodology, with which the clinicians are not familiar. In addition, clinical use of ANNs is tempered by the innate suspicion of a computer based system that cannot explain its reasoning. Interestingly, there is a lot less reluctance in accepting the result of ANN analysis in dealing with problems such as gene filtering to study gene expression in at risk populations that due to their complexity defy conventional statistics\textsuperscript{13,29} or image analysis where data analysis is a simulation of a cognitive process performed by a trained observer.\textsuperscript{10,11} In these situations, the approach towards ANN is pragmatic and result orientated and based on adequate validation of the model without attempts at understanding the reasoning on case-by-case basis.

This is not to say that the development of an ANN model for outcome prediction in patients with invasive cancer is not without its potential pitfalls. A complex ANN with a large number of input variables can be very accurate at identifying patterns within the training dataset. This can result in over fitting of the model, whereby the ANN identifies patterns peculiar to the training dataset that does not represent the larger population of patients. This is overcome by limiting the iterations of training and using adequately sized samples of patients for training and validation of the ANN. Model instability is another potential source of inaccuracy, where identical independently trained ANNs can have varying predictions for individual patients. To minimize this, we used a panel of 10 ANNs trained on randomly selected subsections of data. The training and evaluation samples were randomly selected and were representative of the patient population (Table 3). The correct selection of input variables is also important, as a large number of potential input variables are often available in presence of limited number of training and evaluation cases. This is a potential source of model instability. In this study, variable selection was performed through the well established process of sensitivity analysis, which limited the number of input variables and selected the optimum combination of variables, which was associated with the most accurate predictions (Fig. 2). In this study, external validation of the ANN model was performed using a random selection of cases from the study population, which was withheld from the ANN during the training process. Whilst this is adequate to validate the concept, prior to clinical use any ANN will need to be validated on prospectively collected data in the patient population for which it is being considered.\textsuperscript{15,16}

Sato et al. used ANNs to predict outcome in patients who have undergone resection of oesophageal carcinoma.\textsuperscript{30} The

Table 4

<table>
<thead>
<tr>
<th></th>
<th>Prediction of 1-year survival</th>
<th>Prediction of 3-year survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ANN</td>
<td>TNM</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92.3</td>
<td>66.4</td>
</tr>
<tr>
<td>Specificity</td>
<td>84.5</td>
<td>75.5</td>
</tr>
<tr>
<td>PPV</td>
<td>93.1</td>
<td>85.1</td>
</tr>
<tr>
<td>NPV</td>
<td>83</td>
<td>47.7</td>
</tr>
<tr>
<td>Positive likelihood ratio</td>
<td>5.9</td>
<td>2.56</td>
</tr>
<tr>
<td>Negative likelihood ratio</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>Accuracy</td>
<td>88</td>
<td>71.6</td>
</tr>
<tr>
<td>Discriminant power</td>
<td>2.3</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.
ANN used in that study was designed using 60 and 65 clinical, pathological, biological and genetic variables and predicted survival at 1 and 5 years. The ANN was trained and evaluated using 418 patients who had undergone resection with curative intent from 1983 to 2001. The ANN was found to be an accurate predictor of survival and superior to conventional statistical techniques. The major shortcoming of the predictive model developed by Sato and colleagues is the large number of variables used. This is in direct contrast to the simplicity of conventional staging systems and would limit clinical utility.

The current study used post-operative pathological staging variables to create the ANN. The motivation for selecting pathological staging data was to train the ANN using input variables, which are reliable predictors of survival. One specific problem associated with such an approach arises in relation with patients who receive pre-operative neoadjuvant chemotherapy as the histological findings in this group of patients may have been down staged as a result of chemotherapy. This is an important group of patients who tend to have relatively advanced lesions and their exclusion would have resulted in the ANN model being unrepresentative and adversely effect its use in general population of patients. Pre-operative neoadjuvant chemotherapy was one of the variables used to train the ANN. Although neoadjuvant chemotherapy selected as an input variable through sensitivity analysis, rather than being selected by authors. Regardless of its method of selection as an input variable, its addition should improve the predictive power of the ANN by introducing the interaction between staging variables and pre-operative chemotherapy to the process of predicting outcome at 1 and 3 years.

In clinical practice, the most important decisions regarding patient management are made pre-operatively. Therefore, the next area which needs to be examined, is the ability of ANN in predicting survival based on pre-operative patient and disease specific variables including endoscopic ultrasound as well as multislice computerized tomography. Pre-operative staging data can be used to train an ANN in order to predict survival and aid decision-making regarding therapy, or with appropriate training predict who will respond to a particular therapy such as neoadjuvant chemotherapy.

Additional searches for prognostic variables for oesophageal carcinoma have identified cellular and histological markers of prognostic significance. These variables could easily be incorporated into the design of an ANN in order to improve its accuracy.

This study has identified the potential value of ANNs in predicting post-operative disease free survival from carcinoma of oesophagus and OG junction following surgery with curative intent. Further work is required if it is to be used as decision support aid pre-operatively. Such an artificial neural network needs to be trained prospectively and as part of a multicenter observational study and validated prospectively on pre-operatively collected clinical and imaging investigations. Information provided by such a model can be used to streamline patient management and ultimately improve survival rates from carcinomas of oesophagus and OG junction.

References