A Decision Support System for the Home Management of Patients with Chronic Obstructive Pulmonary Disease (COPD) using Telehealth

Mas Sahidayana Mohktar

A thesis submitted for the degree of

Doctor of Philosophy

Graduate School of Biomedical Engineering

The University of New South Wales

October 2012
The increasing use of telehealth technologies to remotely monitor patients with chronic obstructive pulmonary disease (COPD) has enabled preemptive management of these patients by clinical teams. However, the altered monitoring workload imposed on the clinical care team associated with using a telehealth management strategy has been given less attention.

This thesis describes the design of a decision support system (DSS) to assist clinical care teams in managing COPD patients. The development of an overall DSS framework for a prospective application is firstly described. An analysis of home telehealth data contained in retrospective databases is used to develop the DSS’s knowledge base (rules engine) to facilitate COPD management. Moreover, a preliminary exploration of the effect of data quality on DSS operation is also presented.

The proposed DSS design was implemented using a business process management system with a rules engine as the core component. The objective of the rules engine is to assist the clinical team with the detection of possible COPD exacerbation events, thus facilitating referral decisions. The rules were constructed with two separate clinical measurement databases (termed Database I and Database II), collected from COPD patients enrolled in home telehealth intervention groups as part of two randomised controlled trials. The data were pre-processed and features were extracted, then a classification and regression tree (CART) technique was used to generate the rules.

Four types of CARTs were constructed using four different reference standards, two from each database. The accuracies of the COPD exacerbation algorithms were 79.00% and 76.72%, and the referral recommendation CARTs kappa values were 0.52 and 0.45, for Database I and Database II, respectively. The results showed that the CARTs constructed using home telehealth data were capable of detecting COPD exacerbations as well as generating referral recommendations. In addition, data quality analysis that was performed on the data used by one of the CART algorithms confirmed that data quality issues did affect the reliability of the particular algorithm.

In summary, this thesis presents a DSS that specifically could be used to facilitate the remote monitoring and management of COPD patients. More generally, it helps inform how similar DSSs linked to telehealth systems could improve the management of patients suffering chronic disease.
Originality Statement

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgment is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception or in style, presentation and linguistic expression is acknowledged.’

Mas Sahidayana Mohktar
October 2012
Copyright Statement

‘I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.’

Mas Sahidayana Mohktar

October 2012
Authenticity Statement

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.’

Mas Sahidayana Mohktar
October 2012
Acknowledgement

First and foremost, a special appreciation and heart-felt gratitude goes to my supervisors, Prof. Dr. Nigel H. Lovell, Dr. Stephen James Redmond and Dr. Jim Basilakis for being such a great supervisors, and for their invaluable advice and guidance, especially for their patience while attending to me throughout the whole duration of this research. I am honoured to have been offered the opportunity to learn from them and work with them. The knowledge and experience I’ve gained has been extremely valuable and I hope I will get the opportunity to work with them again in the future.

Equally important and appreciated are the support and assistance from the TeleMedCare Pty. Ltd., Sydney, Australia, especially to Mr. Khang Huynh, Ms. Vicki Clothier and Mr. Dean You. A big thank you to the Standard Best Practice Care (SBPC) nurses, especially Ms. Janette Gogler from Austin hospital in the state of Victoria, and the Respiratory Ambulatory Care Service (RACS) team, especially Ms. Mary Roberts in Blacktown hospital in the state of New South Wales (NSW). Thank you for providing the data in which the analyses were based on.

Thank you to all of the students and staff at the GSBME. To fellow colleagues from the telehealth group, in particular, thank you to Jumadi, Arni Ariani, Tal, Claire, Derrick, and Dave for providing assistance throughout my research.

Special mention goes to Juliana, Zainal, Haziq, Normah, Napi, Bowo and my little Haadi the closest family we have in Sydney and our best friends. On behalf of Sham and Mizan, I can’t thank you enough for all the help and support you have given all these years.

I also wish to record my appreciation to the University of Malaya and the Ministry of Higher Education (MOHE) for funding my doctoral study; lecturers and staff of the Faculty of Engineering, and in particular, the Department of Biomedical Engineering; and all involved directly or indirectly towards this study.

To my beloved parents, in-laws, and family, my heartfelt thank you for all your love and support. For my mum (Mrs. Sapiah Jantan), my dad (Mr. Mohktar Abdullah), my brothers (Shahid, Safuan and Shahriman), and my aunties and uncles thank you for waiting patiently for almost four years for us to come back from Australia, and for always being there whenever we needed you.
And lastly, to my dear husband Hishamudin Abu Bakar, thank you for your love, support and especially patience for putting up with me throughout my struggle in completing this research. To my precious little one, Muhammad Mizan, you are the greatest gift I could ever have. Your smile and laughter always makes my day, always. Mummy loves you very much.

Mas Sahidayana Mohktar  
*University of New South Wales*  
*October 2012*
Abstract

The increasing use of telehealth technologies to remotely monitor patients with chronic obstructive pulmonary disease (COPD) has enabled pre-emptive management of these patients by clinical teams. However, the altered monitoring workload imposed on the clinical care team associated with using a telehealth management strategy has been given less attention.

This thesis describes the design of a decision support system (DSS) to assist clinical care teams in managing COPD patients. The development of an overall DSS framework for a prospective application is firstly described. An analysis of home telehealth data contained in retrospective databases is used to develop the DSS’s knowledge base (rules engine) to facilitate COPD management. Moreover, a preliminary exploration of the effect of data quality on DSS operation is also presented.

The proposed DSS design was implemented using a business process management system with a rules engine as the core component. The objective of the rules engine is to assist the clinical team with the detection of possible COPD exacerbation events, thus facilitating referral decisions. The rules were constructed with two separate clinical measurement databases (termed Database I and Database II); collected from COPD patients enrolled in home telehealth intervention groups as part of two randomised controlled trials. The data were pre-processed and features were extracted, then a classification and regression tree (CART) technique was used to generate the rules.

Four types of CARTs were constructed using four different reference standards, two from each database. The accuracies of the COPD exacerbation algorithms were 79.00% and 76.72%, and the referral recommendation CARTs kappa values were 0.52 and 0.45, for Database I and Database II, respectively. The results showed that the CARTs constructed using home telehealth data were capable of detecting COPD exacerbations as well as generating referral recommendations. In addition, data quality analysis that was performed on the data used by one of the CART algorithms confirmed that data quality issues did affect the reliability of the particular algorithm.

In summary, this thesis presents a DSS that specifically could be used to facilitate the remote monitoring and management of COPD patients. More generally it helps inform how similar DSSs linked to telehealth systems could improve the management of patients suffering chronic disease.
Related Publications and Presentations

ACADEMIC JOURNAL


ACADEMIC CONFERENCES


BOOK CHAPTER

PRESENTATIONS

Oral presentations


Design of a decision support system using open source software for a home telehealth application. Presented at the 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering, Bandung, Indonesia, 8 November 2011 to 9 November 2011. Organizer: Bandung Technology Institute and Indonesia Sensor and Actuator System Society.
Table of Contents

Originality Statement ........................................................................................................... i
Copyright Statement .......................................................................................................... ii
Authenticity Statement ....................................................................................................... iii
Acknowledgement ............................................................................................................. iv
Abstract.............................................................................................................................. vi
Related Publications and Presentations ............................................................................. vii
Table of Contents ............................................................................................................... ix
List of Tables....................................................................................................................... xiv
List of Figures..................................................................................................................... xvii
Abbreviations ..................................................................................................................... xviii

Chapter 1: Introduction

1.1 Overview ...................................................................................................................... 1
1.2 Contributions .............................................................................................................. 3
1.3 Structure ...................................................................................................................... 4

Chapter 2: Background and Literature Review

2.1 Introduction ................................................................................................................. 6
2.2 Chronic Obstructive Pulmonary Disease................................................................. 7
   2.2.1 Introduction and Definition of COPD ......................................................... 7
   2.2.2 Prevalence and Burden of COPD .............................................................. 8
   2.2.3 Aetiology of COPD ................................................................................. 11
   2.2.4 Symptoms, Diagnosis and Investigations of COPD .............................. 11
   2.2.5 COPD Exacerbations .......................................................................... 17
      2.2.5.1 Stages of COPD Exacerbations ..................................................... 17
      2.2.5.2 Management of COPD Exacerbations .................................... 18
   2.2.6 Ambulatory Standard Care and Hospital at Home Schemes for COPD .......................... 21
2.3 Home Telehealth ................................................................................................. 23
   2.3.1 Definition ................................................................................................. 23
   2.3.2 Home Telehealth Application of COPD .......................................... 24
2.4 Decision Support System ..................................................................................... 31
Chapter 3: Problem Statements

3.1 Introduction ................................................................. 52
3.2 Problem Statements ...................................................... 52
3.3 Objectives ........................................................................ 57

Chapter 4: Decision Support System Design

4.1 Introduction ....................................................................... 58
4.2 Background ....................................................................... 58
4.3 Design concept .................................................................... 59
4.4 Design architecture ........................................................... 66
4.5 Technological solution ...................................................... 68
4.6 Laboratory testing ............................................................. 69
4.6.1 Data access layer ......................................................... 69
4.6.2 Logic layer .................................................................... 70
4.6.3 Front end layer ............................................................. 71
4.7 Discussion ........................................................................ 74

Chapter 5: Rules Engine Development Methodology

5.1 Introduction ....................................................................... 76
5.2 Materials ........................................................................... 76
5.2.1 Database ....................................................................... 76
5.2.2 Data ............................................................................. 77
5.2.2.1 Physiological Measurements Data Types ....................... 77
5.2.2.2 Questionnaires Data Types ......................................... 80
5.3 Methodology ...................................................................... 82
Chapter 6: Predicting the Risk of Exacerbation in Patients with COPD using Home Telehealth Data

6.1 Introduction ........................................................................................................ 98
6.2 Background ........................................................................................................ 98
6.3 Methodology ...................................................................................................... 100
  6.3.1 Exacerbation Risk Health Status Reference Standard ...................... 100
    6.3.1.1 Database I ................................................................................. 100
    6.3.1.2 Database II .............................................................................. 101
6.4 Results ............................................................................................................... 102
  6.4.1 Exacerbation Risk Health Status Reference Standard ...................... 102
    6.4.1.1 Database I ................................................................................. 102
    6.4.1.2 Database II .............................................................................. 103
  6.4.2 Data Analysis .............................................................................................. 103
    6.4.2.1 Exacerbation Classification Performance for Database I.............. 103
    6.4.2.2 Agreement between Exacerbation Risk Health Status Reference Standard I from Database I and Reference Standard II from Database II .......................... 105
Chapter 7: Generating Clinical Assessment Referral Recommendations from Home Telehealth Data

7.1 Introduction ........................................................................................................... 110
7.2 Background ............................................................................................................ 110
7.2 Methodology .......................................................................................................... 111
    7.2.1 Referral Status Reference Standard .............................................................. 111
        7.2.1.1 Database I ......................................................................................... 111
        7.2.1.2 Database II ...................................................................................... 112
7.3 Results ................................................................................................................... 113
    7.3.1 Referral Status Reference Standard .............................................................. 111
        7.3.1.1 Database I ......................................................................................... 113
        7.3.1.2 Database II ...................................................................................... 114
    7.3.2 Data Analysis .................................................................................................. 111
        7.3.2.1 Database I ......................................................................................... 115
        7.3.2.2 Database II ...................................................................................... 116
7.4 Discussion .............................................................................................................. 117

Chapter 8: Effect of Home Telehealth Data Quality on Decision Support System Performance

8.1 Introduction ............................................................................................................. 119
8.2 Background ............................................................................................................. 119
    8.2.1 Pulse Oximetry Signal Quality Analysis ....................................................... 119
    8.2.2 Blood Pressure Signal Quality Analysis ....................................................... 121
8.3 Methodology .......................................................................................................... 122
    8.3.1 Statistical Analysis ....................................................................................... 122
    8.3.2 HM-DSS Performance .................................................................................. 123
8.4 Results .................................................................................................................... 126
    8.4.1 Examples of Corrupted Signals .................................................................... 126
    8.4.2 Statistical Analysis ....................................................................................... 127
    8.4.3 Analysis Using Only Pulse Oximetry Features ............................................ 129
    8.4.4 Analysis Using Only Blood Pressure Features ............................................ 130
List of Tables

Table 2-1: The number of associated causes for COPD deaths in people aged 55 years and over as identified on death certificates, Australia, 1997–2003, adapted from (Australian Centre for Asthma Monitoring, 2006)............................9

Table 2-2: Deaths in which COPD was the underlying cause or was mentioned anywhere on the death certificate of people aged 55 years and over, Australia, 1997–2003, adapted from (Australian Centre for Asthma Monitoring, 2006).............................................................................................9

Table 2-3: Indices that can be derived from a forced spirometry measurement...........13

Table 2-4: Five sets of staging criteria for COPD.............................................................15

Table 2-5: A list of tests and investigations that could be performed by a specialist on COPD patients. ........................................................................................................16

Table 2-6: Staging of COPD exacerbations based on health care utilisation and symptoms..........................................................................................................................18

Table 2-7: Clinical indicators in deciding the need to treat an acute exacerbation at hospital. .........................................................................................................................................................19

Table 2-8: Medication types and usage for managing COPD exacerbations..................20

Table 2-9: A summary of studies that used home telehealth applications with COPD patients.........................................................................................................................................................25

Table 2-10: The checklist of necessary criteria for a decision support architecture. ......33

Table 2-11: The definition of accuracy, sensitivity, specificity, PPV and NPV ..........43

Table 4-1: Circumstances in the home management of COPD patients using telehealth and HM-DSS.................................................................................................................................61

Table 4-2: The description of the tables that have been mapped from the TeleMedcare Pty. Ltd. Sydney, Australia database..............................................................69

Table 4-3: Example of rules embedded in the Data evaluation node.................................71

Table 4-4: Important JSF elements in the email templates ................................................71

Table 5-1: List of the measurement parameters and the units ..............................................78

Table 5-2: Questions and answers asked during the trials..................................................80

Table 5-3: κ value interpretations (Cohen, 1960, Viera and Garrett, 2005) ....................87

Table 5-4: Database I subjects baseline data. ....................................................................88

Table 5-5: Total number of days each measurement device was used by all 12 patients .........................................................................................................................................................89
Table 5-6: Total number of available and removed parameter values from all 12 subjects ........................................................................................................... 90
Table 5-7: Database II subjects baseline data. ................................................................. 91
Table 5-8: Total number of days each measurement device was used by 18 patients. ..... 93
Table 5-9: Total number of available and removed parameter values from all 18 subjects. .......................................................................................................... 93
Table 5-10: The percentage of parameters values removed for each parameter type for Database I and Database II. .............................................................. 96
Table 6-1: Multiple feature classification using all features from each measurement type with κ value estimated using cross validation. ..................................... 104
Table 6-2: Confusion matrix for CART classifiers estimated using cross validation using data from all 12 patients from Database I................................................. 104
Table 6-3: Confusion matrix comparing ‘high risk’ and ‘low risk’ labelled using exacerbation risk health status reference standard I and II. .......................... 105
Table 6-4: Multiple feature classification using all features from each measurement type with κ value estimated using cross validation. .................................. 106
Table 6-5: Confusion matrix for CART developed using data from...................................... 106
Table 6-6: Comparisons between the approach taken by (Jensen et al., 2012) and the approach described in chapter six.......................................................... 108
Table 7-1: κ value for classification using data from each measurement technique, estimated using cross validation................................................................. 115
Table 7-2: Confusion matrix for HM-DSS developed using data from Database I.............. 115
Table 7-3: κ value for classification using features from each measurement technique, estimated using cross validation. ............................................................ 116
Table 7-4: Confusion matrix for HM-DSS developed using data from Database II......... 116
Table 8-1: The descriptions of the nine types of HM-DSSs. ........................................... 125
Table 8-2: p-values obtained using the Wilcoxon signed rank test performed on each feature from each subject’s data without* and with signal quality analysis. ........................................................................................................ 128
Table 8-3: The κ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS I, HM-DSS II and HM-DSS III ................................................................. 129
Table 8-4: The κ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS III and HM-DSS IV .................................................................................. 131
Table 8-5: The κ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS VII, HM-DSS VIII and HM-DSS IX ......................................................... 132
List of Figures

Figure 2-1: Mortality rates for pulmonary and cardiovascular diseases in the USA from 1965 to 1998, adapted from (Rodriguez-Roisin, 2009) ........................................ 7
Figure 2-2: The percentage of direct health system expenditure for COPD spent on hospitalisations, by country. ................................................................. 10
Figure 2-3: The process of confirming a COPD case .................................................. 11
Figure 2-4: Spirometry flow-volume curve ............................................................... 12
Figure 2-5: The prevalence of the influenza vaccination, adapted from (Lemon et al., 2003)................................................................. 40
Figure 2-6: The IMS architecture ............................................................................. 45
Figure 4-1: Information flow between human and systems ........................................ 63
Figure 4-2: Schematic of overall workflow for the TMC-Home ............................... 65
Figure 4-3: HM-DSS overall design ................................................................. 67
Figure 4-4: A process in jPDL ................................................................................. 70
Figure 4-5: The email template for generating the report ............................................ 72
Figure 4-6: The measurement data visualised in a time series graph ......................... 73
Figure 5-1: TMC Home automatic and manual measurements .................................... 77
Figure 5-2: General methodology for the development of the rules engine ............... 82
Figure 5-3: A box plot to illustrate the characteristic of a patient’s heart rate data .... 84
Figure 5-4: The timeline shows an example of how the mean, the standard deviation, the percentage changes and the z-score are calculated .................. 85
Figure 5-5: Triaging of the remote monitoring patient .............................................. 89
Figure 5-6: Gender distribution in Database I and Database II ................................. 94
Figure 5-7: Severity of patients in Database I and Database II ................................. 94
Figure 5-8: The percentage of parameters values removed for each parameter type for Database I and Database II ................................................................. 95
Figure 6-1: Distribution of 621 days containing both respiratory medication records and measurement data, for the 12 patients in Database I ........................ 102
Figure 6-2: Distribution of 348 consultation days containing both RACS diagnosis and measurement data, for the 16 patients in Database II ................................ 103
Figure 6-3: Illustration of a CART constructed using data from 11subjects (one of the 12 cross validation run) in Database I ......................................................... 109
Figure 7-1: Distribution of the 82 referred cases ..................................................... 113
Figure 7-2: Distribution of the home management and the referral cases ................. 114
Figure 8-1: A typical PPG signal in the form of light absorption intensity that corresponds to arterial blood volume changes ................................................. 120
Figure 8-2: Examples of cuff pressure, oscillometric and Korotkoff sound signals, 121
Figure 8-3: Methods to construct nine types of HM-DSSs, 124
Figure 8-4: Example of a PPG signal with sections of acceptable and poor quality waveform segments, 126
Figure 8-5: Example of a raw Korotkoff sound signal affected with noise and thus potential yielding unreliable diastolic blood pressure estimation, 127
Figure 8-6: Bar charts represent HM-DSS I, HM-DSS II and HM-DSS III performance, 129
Figure 8-7: The bar charts detail the performance of HM-DSS IV, HM-DSS V and HM-DSS VI, 130
Figure 8-8: The bar charts detail the performance of HM-DSS VII, HM-DSS VIII and HM-DSS IX, 131
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6MWD</td>
<td>Six-minute walking distance test</td>
</tr>
<tr>
<td>AECOPD</td>
<td>Acute exacerbations of COPD</td>
</tr>
<tr>
<td>ATS</td>
<td>American Thoracic Society</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>BPM</td>
<td>Business process management</td>
</tr>
<tr>
<td>BR</td>
<td>Breathing rate</td>
</tr>
<tr>
<td>CART</td>
<td>Classification and regression tree</td>
</tr>
<tr>
<td>CDR</td>
<td>Clinical decision rules</td>
</tr>
<tr>
<td>CHF</td>
<td>Congestive heart failure</td>
</tr>
<tr>
<td>COPD</td>
<td>Chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>COPD-X</td>
<td>Australian and New Zealand guidelines for the management of Chronic Obstructive Pulmonary Disease</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision support system</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiography</td>
</tr>
<tr>
<td>ED</td>
<td>Emergency department</td>
</tr>
<tr>
<td>ERS</td>
<td>European Respiratory Society</td>
</tr>
<tr>
<td>FEV$_1$</td>
<td>Forced expired volume in one second</td>
</tr>
<tr>
<td>FEV$_6$</td>
<td>Forced expired volume in six seconds</td>
</tr>
<tr>
<td>FN</td>
<td>False negative</td>
</tr>
<tr>
<td>FP</td>
<td>False positive</td>
</tr>
<tr>
<td>FVC</td>
<td>Force vital capacity</td>
</tr>
<tr>
<td>GELLO</td>
<td>Object-oriented guideline expression language</td>
</tr>
<tr>
<td>GOLD</td>
<td>Global Initiative for Chronic Obstructive Lung Disease</td>
</tr>
<tr>
<td>GPs</td>
<td>General practitioners</td>
</tr>
<tr>
<td>Hb</td>
<td>Haemoglobin</td>
</tr>
<tr>
<td>HbO$_2$</td>
<td>Oxygenated haemoglobin</td>
</tr>
<tr>
<td>HEARTS</td>
<td>Home-based Everyday Activities analysis and Response Telecare System</td>
</tr>
<tr>
<td>IC</td>
<td>Inspiratory capacity</td>
</tr>
<tr>
<td>IMS</td>
<td>In-Home Monitoring System</td>
</tr>
<tr>
<td>jBPM</td>
<td>JBoss Business Process Management</td>
</tr>
<tr>
<td>jPDL</td>
<td>jBPM process definition language</td>
</tr>
<tr>
<td>JSF</td>
<td>JavaServer Faces</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>LTHM</td>
<td>Lung Transplant Home Monitoring</td>
</tr>
<tr>
<td>mmHg</td>
<td>Millimetres of mercury</td>
</tr>
<tr>
<td>NIBPM</td>
<td>Non-invasive blood pressure measurement</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative predictive value</td>
</tr>
<tr>
<td>NSW</td>
<td>New South Wales</td>
</tr>
<tr>
<td>OSS</td>
<td>Open source software</td>
</tr>
<tr>
<td>PaO₂</td>
<td>Arterial partial pressure of oxygen</td>
</tr>
<tr>
<td>PEF</td>
<td>Peak expiratory flow</td>
</tr>
<tr>
<td>PPG</td>
<td>Photoplethysmography</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>RACS</td>
<td>Respiratory Ambulatory Care Service</td>
</tr>
<tr>
<td>SaO₂</td>
<td>Oxygen saturation</td>
</tr>
<tr>
<td>SBPC</td>
<td>Standard best practice care</td>
</tr>
<tr>
<td>SEBASTIAN</td>
<td>System for evidence-based advice through simultaneous transaction with an intelligent agent across a network</td>
</tr>
<tr>
<td>SpO₂</td>
<td>Saturation of peripheral oxygen</td>
</tr>
<tr>
<td>SVM</td>
<td>Support vector machine</td>
</tr>
<tr>
<td>SWAHS</td>
<td>Sydney West Area Health Service</td>
</tr>
<tr>
<td>Ti/Ttot</td>
<td>The time of inspiration to total breath duration</td>
</tr>
<tr>
<td>T-IDDM</td>
<td>Telematic Management of Insulin-Dependent Diabetes Mellitus</td>
</tr>
<tr>
<td>TMC-Home</td>
<td>TeleMedCare Home Health Monitor</td>
</tr>
<tr>
<td>TN</td>
<td>True negative</td>
</tr>
<tr>
<td>TP</td>
<td>True positive</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>VE</td>
<td>Minute ventilation</td>
</tr>
<tr>
<td>VT</td>
<td>Tidal volume</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Chapter 1 : Introduction

1.1 Overview

Chronic obstructive pulmonary disease (COPD) is a complex and a poorly understood disease caused mainly by smoking (Fishman, 2005). The World Health Organization (WHO) reported in year 2004 that approximately 64 million people worldwide were affected by this disease and the prevalence was most common in the elderly population (Hanania et al., 2010).

Elderly people favour living independently and prefer to stay in their own homes (Mynatt and Rogers, 2001). Therefore, it is important to provide a service that can maintain functional independence in elderly patients with COPD, especially those with a higher degree of adverse health effects caused by the disease (Hanania et al., 2010). Moreover, home management is considered a cost effective solution, given that the number of elderly people is expected to increase to 26% of the total worldwide population by 2051 (Murray and Lopez, 1997).

Patients with COPD should initially be advised to end smoking behaviour and to avoid exposure to tobacco smoke and polluted atmosphere. This advice is necessary to avoid worsening of their health condition. Management plans such as ‘hospital at home’ or ambulatory standard care should also be promoted to elderly patients who stay at home in order to deal with exacerbations that necessitate vital intervention. Accordingly, to assist with the management plan, telehealth technologies have been used (Glaser, 2009).

The purpose of telehealth is to provide healthcare services to remotely located patients. It enhances the power of multidisciplinary teams using communication technology to allow healthcare to be delivered to patients at distant locations. The usage of telehealth technologies has offered inclusive access to healthcare services and has reduced the cost of their delivery (Darkins and Cary, 2000).

Home telehealth systems have been beneficial to elderly patients who are in need of healthcare at home. These electronic monitoring systems enable patients to monitor their own physiological signs, unsupervised at home. Using telecommunication technologies, home telehealth personnel (often skilled nurses) can monitor the patient’s information and provide necessary clinical consultation, from a distance (Finkelstein et al., 2006). Moreover, the ability
to remotely access patient data also enables the discussion of health concerns and health conditions between the patient and disparately located clinical experts. This capability also allows a single nurse carer to proactively monitor more than one patient (Mekhjian et al., 1999).

When COPD patients are affected with an exacerbation of their disease, extra medical attention may be required. (Burge and Wedzicha, 2003). During the normal process of remote monitoring integrated standard care, the home telehealth nurse carer monitors physiological measurement data (self-measured by the patients), and/or contacts the patients, to detect an exacerbation event. When patients have been detected to be experiencing a COPD exacerbation, the nurse carer has to recommend that either the patients be managed at home or be treated in the hospital (Roberts et al., 2008b).

Most studies of home telehealth reveal benefits to using telehealth, in terms of patient outcomes. However, the acceptance of home telehealth among ambulatory care staff is a challenging issue (Koff and Westfall, 2008). The reason being that there is a much wider variation of data captured by the home monitoring device (Darkins and Cary, 2000) and this quantity and mixture of data cannot be efficiently processed without computerised or semi-computerised support (Clifford et al., 2009). Consequently, the home telehealth nurse workload increases, owing to the need to analyse and synthesise the remotely acquired data. Therefore, to assist the ambulatory care team in the task of organising visits and monitoring and assessing the patient’s telehealth measurement data, a customised decision support system (DSS) is proposed.

The benefits of using DSSs as an addendum to home telehealth services is to assist with workload management by completing vast numbers of automated tasks in a short time frame and improving the storage and the management of information. Data acquired from home telehealth systems are typically digitised and stored on a centralised database. These data are easily made available to the healthcare professionals via an internet service. The remotely acquired data are usually stored until the general practitioners (GPs) or nurses are prepared to perform a data evaluation. An effective DSS would provide notifications to these clinical support staff to alert them of patients who require urgent review or intervention (Lovell et al., 2010). Thus, enabling GPs and/or nurses to prioritise their patient workload based on monitoring results highlighted for abnormal deviations for each particular patient, rather than having to routinely monitor all data from every patient.
This thesis describes the initial path in designing a DSS to assist the management of patients with COPD and their nurse carers, using a home telehealth system. The DSS architectural design has been developed using an open source software (OSS) technology and a business process management (BPM) concept. The DSS’s rule engine has been developed using two retrospective home telehealth databases. The data were obtained from randomised controlled clinical trials comparing standard care with remote monitoring integrated standard care. The design has included an algorithm to predict exacerbation and decision rules to assist nurse carers in making referral decisions. To ensure that the guidance provided by the DSS is acceptable, this thesis has also considered the issue of home telehealth data quality, which is a significant issue when physiological measurements are performed in an unsupervised environment.

1.2 Contributions

The main contributions of this thesis are:

- Design of a DSS that could provide immediate feedback and be integrated within a remote monitoring integrated standard care model;
- Analysis of home telehealth physiological measurement data for the early detection of exacerbation events among COPD patients;
- Development of referral recommendation decision rules that could help stratify high priority cases among COPD patients that required urgent clinical assessment;
- An initial investigation of the effect of home telehealth data quality on DSS performance.

There are also a number of other contributions presented in this thesis, which are:

- The introduction of a DSS architectural design for the application in home telehealth using OSS and BPM concepts;
- Classification and regression tree (CART) algorithms developed from the home telehealth data that could be used as a real-time exacerbation and referral event detector;
- A DSS that could help nurse carers to better manage their consultation workload by limiting the time expended in consulting low priority cases;
- Data quality assessment tools that can improve the performance of the DSS.
1.3 Structure

The second chapter of this thesis presents the literature review and background information on COPD, home telehealth and DSSs. Chapter three synthesises the content of chapter two and presents the problem statements and the objectives of this thesis.

Chapter four illustrates the DSS design architecture. The design used a three-tier system architecture that consisted of a data layer, a logic layer, and a front-end layer. The DSS employed a BPM system and used a rule engine for its logic and knowledge base. This chapter also discusses design considerations and illustrates how a system could be developed entirely using an OSS approach.

Chapter five describes the methodology employed in the development of the rule engine. The analysis used a technique called a CART. The CART features were extracted from the home telehealth measurement data. The analyses in chapter six and seven used two different types of health status standards from two retrospective databases. The data collections were based around two hospitals in Australia - the Austin hospital in the state of Victoria and the Blacktown hospital in the state of New South Wales (NSW).

Specifically, in chapter six, we explore the ability of using DSS algorithms to parse home telehealth measurement data and predict exacerbation in patients with COPD. The algorithm classifies patients into two classes: ‘low risk’ and ‘high risk’ of COPD exacerbation. The exacerbation risks reference standard for the Austin site (hereafter referred to as Database I) was generated from patient symptom and medication questionnaires, while the exacerbation risks standard for the Blacktown site (hereafter referred to as Database II) was generated from clinical diagnoses. This chapter demonstrates that when specific rules were applied to home telehealth data, it was possible to predict COPD exacerbation.

The objective in chapter seven was to provide a recommendation to assist the home telehealth nurse in making a decision on whether to refer the patient for clinical assessment or to adjust a patient’s home management plan. The analysis in this chapter also used the CART technique, as per chapter four. This chapter also discusses the potential of using a DSS approach to help nurse carers to better manage their consultation workload by limiting the time dedicated to low priority cases.
Chapter eight discusses the effect of poor home telehealth data quality on the DSS performance. The model performance was evaluated (as developed in chapter seven) with a data set that had been ‘cleaned’ using available signal quality tools.

The conclusions for chapters four, five, six, seven and eight are presented in chapter nine. This chapter also includes a commentary on the outcomes from the thesis and proposed future work.
Chapter 2: Background and Literature Review

2.1 Introduction

COPD is a complex disease that has been acknowledged as one of the top five leading causes of fatality in the world (Rodriguez-Roisin, 2009). The major causes of morbidity in patients with COPD are COPD exacerbations and its association with other co-morbidities. Both factors have caused increases in health care costs in managing COPD. Moreover, the burden of COPD has increased in tandem with an expanding elderly population (Kinsella, 1998, Koch, 2006) and rising tobacco and cigarette sales (Lennon, 2005). To ease the burden, there is an emerging trend for suitable COPD patients to be managed at home by targeted health care services, typically called ambulatory standard care and ‘hospital at home’ (Ansari et al., 2009). In addition, some plans have been proposed to use telehealth monitoring systems to provide continuous monitoring and to access patients regardless of location.

However, the addition of a home telehealth system to an ambulatory care model has resulted in changes in ambulatory care workload management. Thus, a customised DSS is introduced in this thesis to assist the remote care team in the task of monitoring and assessing a patient’s telehealth measurement data. The DSS could facilitate the work of clinical staff at the point of care by stratifying patients into different risk groups in order to simplify the decision making process.

This chapter discusses the topics related to COPD, home telehealth and DSS. Section 2.2 defines and describes COPD disease, including the burden of COPD and ambulatory care schemes for COPD management. A home telehealth definition and its application in COPD are discussed in section 2.3. Section 2.4 characterises the DSS and explains clinical DSSs in detail. Lastly, the conclusion of this chapter is presented in section 2.5.
2.2 Chronic Obstructive Pulmonary Disease

2.2.1 Introduction and Definition of COPD

COPD is a disease that is responsible for a significant amount of individual suffering and has been recognised as a foremost source of morbidity in developed countries. In the United States of America (USA), COPD associated mortality rates increased by 163% over the 30 years up to 1998 (Figure 2-1). Many people who suffer from this long-term disease die suddenly from its complications (Rodriguez-Roisin, 2009).

![Figure 2-1: Mortality rates for pulmonary and cardiovascular diseases in the USA from 1965 to 1998, adapted from (Rodriguez-Roisin, 2009).](image)

COPD is characterised by a fixed or minimally reversible airway obstruction that is, in general, progressive. The airflow obstruction is linked with an unusual inflammatory reaction of the lungs to toxic particles or gases, mainly caused by cigarette smoking. While COPD affects the lungs, it also produces significant systemic consequences (Celli et al., 2004).

The chronic airflow limitation attribute of COPD is generally an association of chronic bronchitis and parenchymal destruction (emphysema). Chronic bronchitis happens when there is an inflammation in the bronchial tube, which leads to an increase in mucus production, resulting in narrowed or blocked airways. Meanwhile, emphysema is caused by a loss of alveolar attachment, occurring when the lung tissue loses its recoil ability and the air sacs become enlarged (Pauwels and Rabe, 2004).
In addition, many patients with COPD also have a variety of other diseases such as lung cancer, myocardial infarction, malnutrition, osteoporosis, anaemia, increased gastro-oesophageal reflux and clinical depression and anxiety. Likewise, patients with COPD are often elderly and regularly present with co-morbidities that also necessitate medical attention (Barnes and Celli, 2009). These co-morbid diseases multiply the mortality risk.

### 2.2.2 Prevalence and Burden of COPD

COPD affects 64 million people worldwide, striking men at the rate of nine per 1,000 individuals and women at seven per 1,000 individuals. These rates are expected to increase relative to the rising trend in population ageing and smoking behaviour. This is proving to be the case in Australia, where 62% of the more than one million people who are affected by COPD are aged 60 or over, and where older people are projected to comprise 44% of the Australian population by 2051 (Crockett et al., 2003, Wouters, 2003, Murray and Lopez, 1997, Australian Bureau of Statistics (ABS), 2008, Catalogue no. 3222.0, ABS, Canberra). Sales of tobacco and cigarettes are also rising in Australia, from approximately $7.06 billion in the 1997–1998 financial year to $9.31 billion in 2004 (Lennon, 2005). That is an increase in sales of 32%. Consequently, COPD has been forecasted to be the third leading cause of mortality in Australia by 2020 (MacNee and Rennard, 2009).

The morbidity and hospitalisation of COPD patients with co-existing diseases is increasing (Barnes and Celli, 2009). Table 2-1 and Table 2-2 show the distribution of causes of mortality in COPD patients according to percentage and gender. Only 7% of deaths were caused by COPD alone, while 93% of deaths had at least one associated disease (Australian Centre for Asthma Monitoring, 2006).
Table 2-1: The number of associated causes for COPD deaths in people aged 55 years and over as identified on death certificates, Australia, 1997–2003, adapted from (Australian Centre for Asthma Monitoring, 2006).

<table>
<thead>
<tr>
<th>Number of associated causes</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (underlying cause reported alone)</td>
<td>7</td>
</tr>
<tr>
<td>One</td>
<td>26</td>
</tr>
<tr>
<td>Two</td>
<td>30</td>
</tr>
<tr>
<td>Three</td>
<td>20</td>
</tr>
<tr>
<td>Four or more</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 2-2: Deaths in which COPD was the underlying cause or was mentioned anywhere on the death certificate of people aged 55 years and over, Australia, 1997–2003, adapted from (Australian Centre for Asthma Monitoring, 2006).

<table>
<thead>
<tr>
<th>COPD</th>
<th>Males</th>
<th>Females</th>
<th>Persons (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying cause of death</td>
<td>23,398</td>
<td>15,672</td>
<td>39,070 (30.9)</td>
</tr>
<tr>
<td>Multiple causes of death</td>
<td>54,997</td>
<td>32,406</td>
<td>87,403 (69.1)</td>
</tr>
</tbody>
</table>

COPD is a very costly disease. Australia reports a total expenditure of $82,925 per patient per year, making COPD the third most expensive disease after cancer and multiple sclerosis (Crockett et al., 2003).

Consequently, in most countries, including Australia, Canada, Italy, Spain, the United Kingdom (UK) and the USA, the largest contributor to the cost of managing patients with COPD has been hospital-based care (Sullivan et al., 2000). All of the above-mentioned countries spent more than 50% of their direct health system expenditure for COPD on inpatient hospitalisations (see Figure 2-2).
The main cause of hospitalisation among patients with COPD is acute exacerbations of COPD (AECOPD), which frequently occurs in patients with moderate to severe COPD (MacNee, 2003, Burge and Wedzicha, 2003). Every day in Australia, an average of 1,000 patients are hospitalised (approximately 350,000 hospitalisations per year (Australian Institute of Health and Welfare, ICD-10-AM, Australia 2006-2007)); with an average length of stay of 7.2 days and a basic standard cost of $3,700 per admission. The cost of two or more hospital admissions exceeds the total health expenditure allocated for each person in Australia, which is around $4,001 per person yearly (Crockett et al., 2003, Australian Institute of Health and Welfare, 2007).

COPD places a similar burden on the health care systems of other countries. The USA funds 16 million GP visits and 500,000 hospitalisations, spending $18 billion annually in direct health care costs for AECOPD (Snow et al., 2001). Yet, since most AECOPD exacerbations are not reported (Trappenburg et al., 2010), the numbers of GP consultations does not reflect the true burden.

Consequently, this significant economic burden has triggered the development of plans for managing exacerbations.
2.2.3 Aetiology of COPD

Smoking has been established as the most significant cause of COPD. Approximately 50% of smokers will develop some airflow limitation and 15 to 20% will develop a significant clinical disability (Walters, 2010). Moreover, people who smoke are at risk of developing lung cancer and other diseases such as cardiovascular disease. While smoking cessation has a minimal impact on the improvement of lung function, it is essential for the conservation of remaining lung function and to impede the onset of disability (McKenzie et al., 2009, Fletcher and Peto, 1977).

Besides cigarette smoking, there are several other acknowledged risk factors for COPD. COPD may arise from gene–environment interaction. It has been linked to the total burden of inhaled particles (Midgley, 2008). Twenty to 30% of COPD cases are caused by occupational and biological dust exposure in the workplace (Matheson et al., 2005). The risk of developing severe COPD is also increased when a person is exposed to air pollution (Andersen et al., 2011).

In addition, people with asthma are also at risk of developing COPD later in life (Silva et al., 2004). Other causes of COPD include oxidative stress, gender, age, socioeconomic status, nutrition and co-morbidities (McKenzie et al., 2009).

2.2.4 Symptoms, Diagnosis and Investigations of COPD

Figure 2-3 shows the typical process used to confirm a COPD case in the Australian healthcare context. First, the patient presents to a GP clinic, depending on when he or she feels the need for treatment. The GP performs an initial diagnosis and refers the patient to a respiratory specialist. The specialist conducts some tests to confirm the diagnosis of COPD.

![Figure 2-3: The process of confirming a COPD case.](image-url)
The level of urgency given to the seeking of medical assistance is generally determined by the degree of a patient’s symptoms. Breathlessness is the symptom that causes the most discomfort and is the most common reason for the patient to undergo a health check. Conversely, although chronic cough and sputum production may come first, they are often not associated with the development of airflow limitation. These symptoms vary across cases, depending on the stage of COPD (Barnett, 2006).

In general, the GP diagnoses COPD cases based on symptoms and airflow limitation measured using spirometry. The measurement is performed on any patient who fits the following criteria:

a. Over 35 years of age;
b. Smoker or ex-smoker;
c. Exhibits any of the distinctive symptoms of COPD, such as constant and progressive breathlessness, coughing, dyspnoea or sputum production;
d. Has a history that includes COPD risk factors such as exposure to cigarettes, exposure to environmental or occupational pollutants, or a family history of chronic respiratory illness.

Spirometry is a physiological test that assesses the volume and the flow of air in the lung. These values are typically measured during a forced spirometry manoeuvre, which can be described as the subject using maximum force to blow all the air out of their lungs as hard and as fast as possible (Bellamy and Booker, 2006). In general, a spirometry device produces a graph of flow versus volume (flow–volume curve), where the flow rate is plotted to the maximum time used to reach maximum exhalation. Figure 2-4 shows an example of a flow-volume curve obtained from a COPD patient.
The volumes exhaled at various time points are expressed both in litres and as a percentage of the predicted value for the patient; this predicted value is based on the patient’s gender, age and height. The derived indices are described in Table 2-3.

### Table 2-3: Indices that can be derived from a forced spirometry measurement.

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
<th>Expression (unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forced expired volume in one second (FEV₁)</td>
<td>The value of the forced expired volume in the first second</td>
<td>Litre (L) Percentage (%)*</td>
</tr>
<tr>
<td>Force vital capacity (FVC)</td>
<td>The total volume of air that can be exhaled from maximal inhalation to maximal exhalation</td>
<td>Litre (L) Percentage (%)*</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>The ratio of FEV₁ to FVC</td>
<td>L/L</td>
</tr>
</tbody>
</table>

*Percentage calculated from the value in litres and the predicted value for that individual.
Equation 1 and 2 show the calculation to predict the individual value of FEV₁ and FVC (Quanjer et al., 1993).

\[
\text{FEV}_1 \text{ predicted} \quad (1) \\
\begin{align*}
\text{Men: } & 4.30(\text{Height (m)}) - 0.029(\text{Age (yr)}) - 2.49 \\
\text{Women: } & 3.95(\text{Height (m)}) - 0.025(\text{Age (yr)}) - 2.60
\end{align*}
\]

\[
\text{FVC predicted} \quad (2) \\
\begin{align*}
\text{Men: } & 5.76(\text{Height (m)}) - 0.026(\text{Age (yr)}) - 4.34 \\
\text{Women: } & 4.43(\text{Height (m)}) - 0.0256(\text{Age (yr)}) - 2.89
\end{align*}
\]

A post-bronchodilator¹ FEV₁/FVC of less than 0.7 confirms the existence of airflow obstruction or restriction that is not fully reversible (Celli et al., 2004, McKenzie et al., 2009, Pearson, 1997). Using the spirometric value and the symptoms presented by the patients, five stages of COPD severity can be distinguished.

Table 2-4 shows a summary of five sets of staging criteria from four available guidelines for COPD: the American Thoracic Society (ATS) guidelines, the British Thoracic Society guidelines, the Australian and New Zealand guidelines for the management of Chronic Obstructive Pulmonary Disease (COPD-X), and the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines. The categories of stages are: 0 (at risk), I (mild), II (moderate), III (severe) and IV (very severe) (Sutherland and Cherniack, 2004).

¹ A bronchodilator is a drug that causes the widening of the airways that pass air to and from the lungs. The usage of this drug is explained in section 2.2.5.
Table 2-4: Five sets of staging criteria for COPD.

<table>
<thead>
<tr>
<th>Guidelines</th>
<th>Stage</th>
<th>0 At risk</th>
<th>I Mild</th>
<th>II Moderate</th>
<th>III Severe</th>
<th>IV Very severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Thoracic Society (Celli et al., 2004)</td>
<td>FEV(_1) (%)</td>
<td>≥ 50</td>
<td>35–49</td>
<td>&lt;35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>British Thoracic Society (Pearson, 1997)</td>
<td>FEV(_1) (%)</td>
<td>≥80</td>
<td>50–79</td>
<td>30–49</td>
<td>&lt;30**</td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>±</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COPD-X (McKenzie et al., 2009)</td>
<td>FEV(_1) (%)</td>
<td>60–80</td>
<td>40–59</td>
<td>&lt;40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>±</td>
<td>+</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLD (Rodriguez-Roisin, 2009)</td>
<td>FEV(_1) (%)</td>
<td>≥80</td>
<td>50–79</td>
<td>30–49</td>
<td>&lt;30</td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>-</td>
<td>±</td>
<td>++</td>
<td>+++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The predicted percentage of FEV\(_1\) is extracted from a spirometry manoeuvre performed after administration of bronchodilator.

- no symptoms, ± variable symptoms, + mild to moderate symptoms,
++ symptoms that limit physical exertion, +++ symptoms that limit functional activities

** or FEV\(_1\) <50% in patients with respiratory failure (insufficient gas exchange by the respiratory system)

The general spirometric staging criteria for COPD require post-bronchodilator measurement of lung parameters. However, retrospective studies that did not use a bronchodilator, used modified GOLD criteria for COPD. The mutually exclusive categories are: mild (FEV\(_1\) >80%), moderate COPD (FEV\(_1\) >50 to <80% predicted) and severe COPD (FEV\(_1\) <50% predicted) (Mannino et al., 2003, Probst-Hensch et al., 2010). The use of the modified criteria are only suitable for confirmed COPD cases that do not have completely reversible airway obstruction (FEV\(_1\)/FVC <0.70) (Celli et al., 2003).

Specialist knowledge and further investigation is required to confirm the COPD case and to isolate the patient’s condition from chronic asthma or other airway diseases. Table 2-5 shows the list of tests that could be performed by the specialist on patients diagnosed as having COPD by their GP (McKenzie et al., 2009).
Table 2-5: A list of tests and investigations that could be performed by a specialist on COPD patients.

<table>
<thead>
<tr>
<th>Test/investigations</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex lung function tests</td>
<td>To assess if patients have more complex respiratory disorders.</td>
</tr>
<tr>
<td>Exercise testing</td>
<td>To differentiate between breathlessness resulting from cardiac and respiratory disease, and to identify other causes of exercise limitation.</td>
</tr>
<tr>
<td>Sleep studies</td>
<td>To check if patients have a co-existent sleep disorder.</td>
</tr>
<tr>
<td>Chest x-rays</td>
<td>To detect changes in COPD patient’s condition.</td>
</tr>
<tr>
<td></td>
<td>To exclude other disorders.</td>
</tr>
<tr>
<td></td>
<td>To detect lung cancer.</td>
</tr>
<tr>
<td>High resolution computed tomography</td>
<td>To consider bullectomy or lung reduction surgery.</td>
</tr>
<tr>
<td>Ventilation and perfusion scans</td>
<td>To assess whether patients are suitable for lung resection or lung volume reduction surgery.</td>
</tr>
<tr>
<td>Arterial blood gas measurement</td>
<td>To consider domiciliary oxygen therapy</td>
</tr>
<tr>
<td>Sputum examination</td>
<td>To start or change antibiotic therapy or to identify resistant organisms.</td>
</tr>
<tr>
<td>Haematology and biochemistry</td>
<td>To identify the overproduction of red blood cells secondary to COPD.</td>
</tr>
<tr>
<td>Electrocardiography (ECG) and echocardiography</td>
<td>To confirm clinically suspected arrhythmias.</td>
</tr>
</tbody>
</table>
2.2.5 COPD Exacerbations

A patient with COPD may experience an exacerbation, which is one of the most common conditions requiring hospital admission. While there is no widespread agreement on the definition of exacerbation, the description generally used is, an acute deterioration of respiratory function that typically requires a patient to seek medical assistance or to change their treatment (Rodriguez-Roisin, 2000). One-third of exacerbation events are caused by infections (infective exacerbations). Non-infective exacerbation is caused by the following triggers: environmental pollution and inhaled irritants (Wedzicha and Seemungal, 2007, Madison and Irwin, 1998).

Exacerbations become more frequent as the severity of COPD increases. In patients with moderate to severe COPD, there are an average of two episodes per year. A study by Miravitlles et al. showed that exacerbation frequency is a significant determinant of the wellbeing status of patients with COPD. The findings of the study were that patients with more recurrent exacerbations had a notably poorer quality of life than those with less frequent exacerbations (Bellamy and Booker, 2006, Miravitlles et al., 2004).

2.2.5.1. Stages of COPD Exacerbations

The stages of COPD exacerbations in patients with confirmed COPD can be characterised based on health care utilisation and symptoms (see Table 2-6) (Rodriguez-Roisin, 2000, Burge and Wedzicha, 2003).
Table 2-6: Staging of COPD exacerbations based on health care utilisation and symptoms.

<table>
<thead>
<tr>
<th>Severity/Type</th>
<th>Level of health care utilisation</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>Patients have an increased need for medication, which they can manage in their own normal environment</td>
<td>A single symptom from: An increase in sputum volume An increase in sputum purulence An increase in dyspnoea(^) and at least one minor symptom*</td>
</tr>
<tr>
<td>Moderate</td>
<td>Patients have an increased need for medication and feel the need to seek additional medical assistance</td>
<td>Any two from: An increase in sputum volume An increase in sputum purulence An increase in dyspnoea</td>
</tr>
<tr>
<td>Severe</td>
<td>Patients or caregivers recognise obvious and/or rapid deterioration in condition, requiring hospitalisation</td>
<td>An increase in sputum volume and An increase in sputum purulence and An increase in dyspnoea</td>
</tr>
</tbody>
</table>

\(^\) Dyspnoea means breathlessness  
* Minor symptoms include: sore throat or nasal discharge within the past 5 days, fever without other cause, increased wheezing, increased coughing, increased respiratory rate >20% above baseline, increased heart rate >20% above baseline

Most of the criteria described in Table 2-6 are qualitative measures of patients’ conditions, measured using questionnaires (Leonie et al., 2010). Therefore, there is a need for quantitative measures that can be used as a surrogate to assist in the diagnosis of exacerbations. Regrettably, little is known of the biological markers in both stable and exacerbation phases of COPD (Franciosi et al., 2006, Miravitlles, 2009).

2.2.5.2 Management of COPD Exacerbations

Patients in the mild stage of COPD can usually be managed at home, however, patients in the moderate or severe stages should be referred to health care personnel or should be hospitalised. Table 2-7 lists the current clinical indicators used in determining the need for hospitalisation in patients experiencing COPD exacerbation (MacIntyre and Huang, 2008, Bellamy and Booker, 2006, Pearson, 1997).
Table 2-7: Clinical indicators in deciding the need to treat an acute exacerbation at hospital.

<table>
<thead>
<tr>
<th>Clinical indicators</th>
<th>Home treatment</th>
<th>Hospital treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Able to cope at home</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Breathlessness</td>
<td>Mild</td>
<td>Severe</td>
</tr>
<tr>
<td>General condition</td>
<td>Good</td>
<td>Poor, deteriorating</td>
</tr>
<tr>
<td>Level of activity</td>
<td>Good</td>
<td>Poor, confined to bed</td>
</tr>
<tr>
<td>Cyanosis(^a)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Worsening peripheral oedema(^b)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Level of consciousness</td>
<td>Normal</td>
<td>Impaired</td>
</tr>
<tr>
<td>Already receiving long-term oxygen therapy</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Social circumstances</td>
<td>Good</td>
<td>Living alone/not coping</td>
</tr>
<tr>
<td>Acute confusion</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Rapid rate of onset</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Measurements available in hospitals and clinics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes on the chest x-ray</td>
<td>No</td>
<td>Present</td>
</tr>
<tr>
<td>Arterial pH level</td>
<td>&gt;7.35</td>
<td>&lt;7.35</td>
</tr>
<tr>
<td>Arterial partial pressure of oxygen (PaO(_2))</td>
<td>&gt;7 kPa/52.5mmHg</td>
<td>&lt;7 kPa/52.5mmHg</td>
</tr>
</tbody>
</table>

\(^a\) Cyanosis is the manifestation of a blue or purple blush of the skin due to the tissues being low in oxygen.  
\(^b\) Peripheral oedema is the enlargement of tissues, frequently in the lower limbs, due to the accretion of fluids.

By examining the clinical signs in the provisional guidelines in Table 2-7, health care personnel can decide whether to treat the patient at the hospital or to supervise home management. Patients are returned to and managed at home if they match less of the above-mentioned criteria and have an adequate social support network. Conversely, patients who qualify for admission to hospital may be deemed fit for an early discharge and be returned home for management in their own normal environment under certain conditions (Barnett, 2006).

Smoking cessation is the first and most important step for patients with COPD. They should also avoid exposure to tobacco smoke or other air pollutants to prevent exacerbations (Glaser, 2009). However, when exacerbations occur, drug therapy can and has been used to reduce symptoms and complications, even though there is no existing medication that can modify the long-term rate of decline in lung function.

Table 2-8 describes the three main types of medication used and their application in managing COPD exacerbations.
## Table 2-8: Medication types and usage for managing COPD exacerbations.

<table>
<thead>
<tr>
<th>Medication type</th>
<th>Description</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronchodilators</td>
<td>This is the main type of drug for symptomatic management. It is given as required or on a regular basis to reduce symptoms. An inhaled course is generally preferred.</td>
<td>Used in the home management of COPD exacerbation. A high-dose of nebulised therapy can be given in severe cases.</td>
</tr>
<tr>
<td>(drugs that widen the bronchi and bronchioles, thereby decreasing resistance in the respiratory airway and increasing airflow to the lungs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>The inhaled type can be used orally for short-term treatment.</td>
<td>Used in the management of AECOPD. It helps to shorten the recovery time and to restore lung function more quickly.</td>
</tr>
<tr>
<td>(also known as steroids, they act as anti-inflammatory agents)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotics</td>
<td>This drug is used to treat infectious exacerbations.</td>
<td>It is effective only when patients with worsening dyspnoea and cough also have increased sputum volume and purulence.</td>
</tr>
<tr>
<td>(drugs that kill bacteria)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Typically, patients with regular infective exacerbations are provided with an emergency standby supply of antibiotics, which they can administer immediately if they experience an increase in symptoms. Likewise, for patients who have previously benefited from courses of oral steroids, it is sensible to provide them with a supply to take if an urgent situation arises such as the worsening of breathlessness or wheezing over a 24-hour period (Bellamy and Booker, 2006).

Further, to help ease the high demand for hospital beds and to assist in the management of COPD exacerbation in respiratory patients, there is a growing trend for suitable patients experiencing exacerbation to be managed at home by groups of health care personnel. Schemes to perform this management have been deployed to support COPD patients either after hospitalisation (ambulatory care teams or community nurses) or instead of hospitalisation (‘hospital at home’ schemes). In general these services provide home visits to monitor the patient’s condition (Ansari et al., 2009).
2.2.6 Ambulatory Standard Care and Hospital at Home Schemes for COPD

Primarily people with COPD have been managed within tertiary hospital settings. However, recently, ambulatory standard care have positively influenced the service model (Hermiz et al., 2002). The aim of the service is to reduce recurrent hospitalisations for people with COPD and to manage bed demand (Kliethermes and Brown, 2011).

Moreover, another scheme called the hospital at home is also a reliable alternative to hospitalisation for COPD. The scheme has been shown to be a safe and useful approach for suitable patients with exacerbations of COPD. This scheme involves support from hospital doctors, GPs, nurses, physiotherapists and other health care professionals. Most of the plan types offer specialised respiratory nurses providing care at home (Stevenson, 2007, Cecile et al., 2010, Taylor et al., 2007).

Whenever patients with COPD choose to accept either ambulatory standard care or hospital at home support, they receive visits from one of the support team members (typically an ambulatory nurse). During the home visits, the nurse records the patient’s symptoms such as dyspnoea, cough, sputum colour and amount and performs physiological measurements on the patient. The collected information is used to assess the patient’s condition. The nurse gives copies of the clinical notes to the patient to assist if an emergency call is made (Stevenson, 2007).

Several standard care randomised control studies involving COPD, congestive heart failure (CHF) and diabetes patients have been conducted to assess the advantages of the scheme. The findings have shown benefits in terms of reducing the risk of hospital re-admission and lowering health care costs, while also providing a better quality of life for enrolled patients (Ansari et al., 2009, Frick et al., 2009, Aimonino Ricauda et al., 2008).

Despite consistent reported benefits, in their community-based cohort study of a hospital at home and early or assisted discharge scheme, Matheson et al. found a limitation. They revealed that when these types of short-term monitoring methods were used, continuous monitoring was not provided to the patients after the monitoring period ended. This had caused many patients with COPD to go undertreated (Matheson et al., 2006). Consequently, this made it difficult to detect health deterioration in patients with COPD (Stevenson, 2007, Lamothe et al., 2006).

Despite this, with a motivation to deliver an early and effective intervention for COPD exacerbation, a Respiratory Ambulatory Care Service (RACS) in Blacktown in the Western
Sydney area of Australia has provided a 24-hour help-line service for patients with COPD that were staying at home. The RACS team operated a 24 hour, 7 day a week telephone line for patients with at least one prior admission caused by COPD exacerbation. As a result, the patients who used the service had reduced hospital presentations due to COPD exacerbation (Roberts et al., 2008a).

However, the standard care services need to be appropriate not only for patients with COPD living in urban areas, but also for those living in rural areas. This is important in a large and sparsely populated country like Australia, which has patients living in remote areas. A study by Goodridge et al. reported on the differences between the respiratory illness home care services provided to residents in municipal and remote areas. Not unexpectedly, the study showed that patients in urban areas had more visits from GPs and had easier access to standard care services than patients living in rural areas (Goodridge et al., 2010).

Based on positive evidence from past studies, it is thought that the sustained development of standard care models will become a necessary element of future health care schemes for managing patients with COPD (Nicholson et al., 2001). However, full time monitoring leads to a time management issue, caused by the limited number of qualified nurses that can perform ambulatory care (Cowen and Moorhead, 2006).

To solve this time management and nursing resource issue, and to provide equal health care utilisation opportunities to all patients regardless of location, a telehealth technology is used (Meade and Dunbar, 2004, Glaser, 2009).
2.3 Home Telehealth

2.3.1 Definition

Telehealth is defined as the support of health care by using electronic information and communication technologies in situations in which the user and provider are in separate locations. Advances in technology allow various patient data to be monitored and transmitted using either a phone line or broadband system, with information transmitted from distant locations to a hospital or clinic (Koch, 2006, Bowles and Baugh, 2007).

Telehealth is one of the approaches being taken to solve the international care crisis arising from the aging population, lack of human resources and lack of access to health care facilities. The rate of health care service usage is four times greater for an elderly person than a younger person. In addition, it is likely that there will be an inadequate number of nurses to cater for the growing elderly population. Likewise, it is expected that there will be a shortage of health facilities to accommodate them (Botsis et al., 2008).

A home telehealth system provides remote health care to patients in their home environment. The principle of the home telehealth system is to monitor the patient’s condition based on tailored care protocols. This system requires patients to use a storage and communications device to enter information on their condition regularly. The device, which is installed in the patient’s home, transfers this information to a central database, allowing the remote clinical team to observe the patient’s condition and provide care accordingly (Basilakis et al., 2010, Sicotte et al., 2011).

There are two categories of home telehealth: telephone support (including videoconference technology) and home telemonitoring. In the telephone support plan, the health care supplier provides support to patients in the course of telephone contact. This type of home telehealth does not involve electronic transmission of patient’s clinical measurement data (Polisena et al., 2010). Conversely, home telemonitoring involves patients interacting with measurement devices, generating clinical information that is eventually stored on a remote database, which is then available to clinical users for review. A home telemonitoring device provides users with the facility to measure and monitor their vital signs and symptoms on a regular basis, non-invasively (Basilakis et al., 2010, Collinge and Liu, 2009).
2.3.2 Home Telehealth Application of COPD

This section reviews available studies that have used home telehealth applications with patients with COPD.

Table 2-9 displays the summary of the studies. Five studies divided patients into control and intervention groups; one study evaluated patients before and after they used home telehealth; and the remaining six studies assessed the outcomes of patients using the system.
Table 2-9: A summary of studies that used home telehealth applications with COPD patients.

<table>
<thead>
<tr>
<th>Study</th>
<th>Method/description</th>
<th>Monitoring/assessment</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Paré et al., 2006)</td>
<td>A comparison between two groups: COPD patients under the hospital at home care scheme, receiving access to a home telehealth system. COPD patients receiving regular hospital at home care.</td>
<td>A web phone was used by patients to complete data entry forms that documented their spirometry measurement, symptoms and medication records.</td>
<td>Fewer visits were made by nurses to the Group 1 patients’ homes, and the cost of providing care decreased by 15%.</td>
</tr>
<tr>
<td>(Lewis et al., 2010)</td>
<td>Group 1 was asked to record their symptoms and physical observations (body temperature and pulse oximetry measurements) two times per day. These data were automatically transmitted to a database once a day. Nurses accessed the data via a website and received alert email messages if two or more of the subsequent conditions occurred: any symptoms scoring ‘much worse than usual’, pulse rate greater than 120 beats per minute, oxygen saturations less than 88%, or body temperature greater than 38.5°C.</td>
<td>Home telemonitoring was reported to be safe. There was no significant difference in quality of life as measured using questionnaires between stable COPD patients monitored by standard care alone, and those receiving home telemonitoring augmented standard care.</td>
<td></td>
</tr>
<tr>
<td>(Vontetsios et al., 2005)</td>
<td>An investigation of ambulatory care nurses or skilled nurses equipped with home telemonitoring devices.</td>
<td>In each visit, the nurse used the equipment to assess the patient’s vital signs, to perform medical tests and to search for signs of exacerbation. Collected data were transmitted to the hospital via the internet. The visiting nurse was also able to perform videoconferences with distant health professionals.</td>
<td>There was a 23% reduction in the number of emergency and scheduled visits. There was a 67% decrease in the number of patient hospitalisations. Details of the statistical analysis were not reported in this paper.</td>
</tr>
<tr>
<td>(Finkelstein et al., 2006)</td>
<td>A comparison of three methods: The control group only received the usual home healthcare by a skilled nurse. Group 1 received home healthcare by a skilled nurse and virtual visits</td>
<td>All patients were enrolled in home healthcare where skilled nurses monitored their underlying disease (congestive heart failure, or COPD, or chronic wound care). Patients were required to perform their physiological measurement and record their health data.</td>
<td>Virtual visits between a home health care nurse and patients at home using home telemonitoring devices improve patient outcomes at a lower cost; with a cost of $22.11 and $48.27 for telemonitoring and</td>
</tr>
</tbody>
</table>
using video-conferencing technology.

Group 2 received the same intervention as Group 1 in addition to receiving physiologic monitoring.

<table>
<thead>
<tr>
<th>(Koizumi et al., 2005)</th>
<th>A real-time system connected three remote locations: the residence of chronic respiratory failure patients, the physician’s hospital and the pulmonary specialist’s hospital.</th>
<th>A multi-station telemedicine support system was investigated. Rehabilitation programs were provided to patients who had difficulty visiting hospital regularly. The interactions between patients, physicians and pulmonary specialists were scheduled once a week. The patient’s medical history and biologic variables were shared by both the physicians and the pulmonary specialist in real-time.</th>
<th>An appropriate way for rehabilitation programs to care for remotely located patients by involving respiratory specialists and physicians was established.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Maiolo et al., 2003)</td>
<td>A study conducted in two phases: 1st Phase: For the first 12 months, patients were observed and treated using face-to-face medical visits. 2nd Phase: During the consequent 12 months, patients were supplied with a telemonitoring device that could measure their arterial oxygen saturation and heart rate.</td>
<td>Telemonitoring services were provided for patients with severe respiratory illness. Measurements were performed twice a week and were automatically transmitted to the hospital’s processing centre.</td>
<td>There was a 50% decrease in the numbers of hospital admissions and a 55% drop in acute home exacerbations. Both results were statistically significant ($p&lt;0.05$). 96% (22 out of 23) of the patients were pleased with the quality of the telemonitoring. Telmedicine can facilitate high-quality home care for patients with severe respiratory illness.</td>
</tr>
<tr>
<td>(Dale et al., 2003)</td>
<td>This was a three-month study of a home monitoring service for patients with COPD.</td>
<td>The monitoring centre automatically received physiological data on a daily basis from 55 recruited patients. The system incorporated a pulse oximeter and a weight-monitoring device. Every day, nurses called the patients and asked health related questions. In the event of an escalation,</td>
<td>There was a decrease (approximately 50%) in the rate of hospital admissions.</td>
</tr>
</tbody>
</table>
an increase in symptoms or deterioration in physical conditions, the nurse contacted the patient’s GP, who made a decision about treatment.

| (Lamothe et al., 2006, Gagnon et al., 2006) | This home telecare program targeted patients who required home care visits after hospitalisation. | Enrolled patients were supplied with a home monitor (HomMed Sentry, Honeywell, Brookfield, WI) connected to peripheral apparatus such as a weight scale, thermometer, sphygmomanometer and pulse oximeter. Patients were required to perform and send the required measurement on a daily basis. Data from the home monitor could be transmitted wirelessly or wired to the central monitor, located in the primary care organisation. The central monitor was operated by nurses who responded to alerts or abnormal measures and under certain circumstances, called or visited the patients. | The presentation of weekly or monthly data readings in graphs helped doctors in identifying patients’ conditions. Qualitative feedback revealed that the program increased patients’ access to health services and led to less patient emergency visits. |

| (Trappenburg et al., 2008) | This multicentre prospective controlled non-randomised study took place across 4 hospitals. The control group included patients from two hospitals received care as usual without telemonitoring for 6 months. The intervention group included patients from two hospitals who received care as usual and used the telemonitoring device for 6 months. | In this and the next two studies, intervention group was supplied with Health Hero’s network technology (Health Buddy), a pulse oximeter, a FEV\textsubscript{1} monitor and a pedometer. Each of the additional instruments came from a different vendor. Patients were asked to answer questions about their disease symptoms and medication compliance every weekday morning. Patients were required to enter their FEV\textsubscript{1}, results from a six-minute walking distance test (6MWD) measured using the pedometer and their resting arterial oxygen saturation measured by pulse oximetry data. | There was a significant reduction in hospital admission (\(p = 0.02\)). The study revealed a potential method for the early detection and treatment of exacerbations. |

| (Koff et al., 2009) | A group of 40 COPD patients were randomly assigned to receive proactive integrated care or usual care over a 3-month period. | There were two main finding for this study: (1) Changes in quality of life, measured using the St George’s Respiratory Questionnaire, was higher in the proactive |
| (Dang et al., 2006) | This study involved participants aged 60 or over.
19 patients had congestive heart failure, 23 had diabetes mellitus and 17 had COPD. | integrated care group compared to the usual care group; (2) A decline in healthcare costs by US$1,401 in the proactive integrated care group compared with an increase of cost by US$1,709 in the usual care group. | It was found that telehealth models had no effect on hospital admission or emergency department visits in elderly patients with COPD. There were also no significant different in the number of beds used in the hospital for the COPD cases ($p = 0.24$). |
The above-mentioned studies used a variety of home telehealth technology and overall demonstrated mixed outcomes. Some studies only used symptom-based monitoring devices (Trappenburg et al., 2008, Dang et al., 2006), while others only focused on videoconferencing (Vontetsianos et al., 2005, Finkelstein et al., 2004, Finkelstein et al., 2006, Finkelstein et al., 2001). Although such studies are beneficial, diverse instruments should be combined to give a thorough representation of a patient’s health status (Leonie et al., 2010). In all of the studies that collected the patient’s physiological measurement data automatically (Vitacca et al., 2009, Paré et al., 2006, Vontetsianos et al., 2005, Finkelstein et al., 2006, Dale et al., 2003, Lewis et al., 2010, Koizumi et al., 2005, Maiolo et al., 2003, Gagnon et al., 2006, Lovell et al., 2010), nurses had to be made available either for supervising vital sign measurements, or for performing medical charting or analysis of data. In two of the studies, the comprised components—the physiological measurement devices (i.e., pulse oximetry, weight scale, spirometry and pedometer) and questionnaire modules—did not coordinate automatically (Dale et al., 2003, Koff et al., 2009). This is because each component was developed by a different manufacturer (Koff et al., 2009, Sakka et al., 2004). This limitation has been solved in the TeleMedCare Home Health Monitor–TMC-Home system, in which all of the components (i.e., vital sign monitoring and health questionnaires) have the same communication protocols and data format and are automatically synchronised (Basilakis et al., 2010, Lovell et al., 2010).

The TMC-Home system was developed by our research group over a decade ago. The system has been commercialised by TeleMedCare Pty. Ltd., Sydney, Australia. The system’s function is to integrate a range of e-health services to manage chronic disease in the patient’s home. The central component of the TMC-Home system is a workstation that connects to a mobile computing platform, which has internet connectivity. Patient users can use the system to measure and monitor their vital signs, including blood pressure, blood oximetry, weight, lung function (forced and relaxed spirometry), temperature and blood glucose levels, on a regular basis. The patient users also have access to a range of lifestyle and health questionnaires, which can be used to measure general wellness and fitness (Basilakis et al., 2010, Lovell et al., 2010).

In addition, the TMC-Home system has been used in two randomised controlled studies comparing standard care with remote monitoring augmented standard care. One study was based around the Austin hospital in the state of Victoria and the other was based around the Blacktown hospital in the state of NSW. The ‘standard care’ is an inclusive respiratory outreach program for patients at high risk of admission, using frequent home visits, action plans and telephone support, in addition to pulmonary rehabilitation. These plans were conducted by: (i) the ambulatory nurses in the case of the Austin data collection and (ii) the RACS team for the Blacktown data collection. In the first study, patients were asked to answer survey
questionnaires after the completion of the study. Ninety-four per cent of the patients found the home monitoring system easy to use, and 82% felt that the system had improved their management of their condition (Rochford et al., 2009).

Patients enrolled in the second study were asked to complete perception questionnaires after the second week and after the fourth month using the home monitoring unit. Ninety-three per cent of the patients agreed that the TMC-Home system had played an important role in managing their health and that the system had given them more control over their lung disease conditions (Shany et al., 2010). However, the supervision of vital signs and data analysis tasks increased the RACS staff workload (Rochford et al., 2009, Shany et al., 2010, Roberts and Robinson, 2011). The likely reason for this is the much wider disparity of data captured by the home monitoring device requiring additional time by the ambulatory care staff to interpret the data (Darkins and Cary, 2000). This was in agreement with another home telehealth study among chronic respiratory failure patients conducted by Vitacca et al. in which a home telehealth nurse was required to dedicate an extra 73 minutes per month for each newly enrolled patient (Vitacca et al., 2010).

Therefore, with a motivation to provide an efficient way for the ambulatory care team to perform the task of monitoring and assessing the patient’s telehealth measurement data, a computerised support system is proposed.
2.4 Decision Support System

2.4.1 Definition

A DSS, is a computer based system that helps decision makers extract the most from data and enables model formation to resolve unstructured problems (Sprague Jr, 1980). DSSs have been used in a wide range of domains in which decisions are made, including the clinical domain.

DSSs can range from simple alert systems, to complex, long-term guideline-based clinical management systems. Most of these services are enabled through the combination of interacting logical components, with a pattern classification system at its heart. For instance, a certain characteristic of the patient is recorded and interpreted to decide upon a course of action, in order to make a decision about the clinical management of the patient. Most pattern classification systems fit a standard model, comprising data acquisition, pre-processing and feature extraction, classification and decision making stages (Duda et al., 2000).

Acquisition of data makes up a significant component of the information generated in most current telehealth systems to date. Telehealth systems that support the automatic collection of data can be inherently supported by DSSs. The choice of inputs to a DSS is tightly intertwined with the classification task to be performed requiring a priori knowledge of the clinical domain.

Blood pressure, weight, ECG data, body temperature, spirometry, medication records and patient responses to questionnaires are among the most common types of data acquired from patients using a home telehealth system. These types of data typically use a variety of collection formats ranging from signal waveforms to coded medical record items. The challenge for any DSS in home telehealth is to be able to seamlessly process and integrate these widely varying data sets.

Pre-processing describes any type of processing performed on the raw data to prepare it for another processing procedure, while feature extraction is the process of compressing the data into a number of features. The types of features extracted are completely dependent on the classification task. In home telehealth, clinical information such as clinical history and medication records are important features that must be considered in interpreting clinical data. A sudden deviation in a patient’s features may be the result of a change in health status or medication. The more information a classifier incorporates about the clinical context surrounding the generated measurement data, the better the model for reasoning correctly about the clinical domain during classification.
Classification involves assigning a patient to a particular class, based on their most recently acquired data. The simplest classification in home telecare might be into the classes of ‘unwell’ versus ‘healthy’. Typically, most of the difficult work will have been performed during the feature extraction phase, and it remains for the classifier to ‘divide up’ the feature space into regions, with each region defining a class. In order to achieve this in any meaningful or robust manner, the classifier must have some outcome results against which to be measured. These could be objective or subjective clinical outcomes such as admission to hospital, visits to the GP or simply the patient’s report of feeling unwell.

The definition of a DSS is that it supports the final decision making process. This may involve taking inputs from a number of integrated classifier systems using a coordinated approach, together with accurate reasoning about the clinical context surrounding these inputs. The aggregation of input features from the classification modules and subsequent decision-making itself encompasses the definition of a pattern classification system at a higher level, having the same techniques available as those employed by lower level classification systems. However, the support provided by a DSS has to be interpreted by the medical expert within a broader clinical context, taking into account consultations with the patient and their family members and other contextual cues that may not have been encoded in the DSS algorithms (Luger, 2005).

2.4.1.1 Necessary criteria for a decision support architecture

Wright et al. have identified the necessary criteria for decision support architecture (Wright and Sittig, 2008). The essential criteria are:

(i) shareable: to enable the sharing of the contents of a DSS,
(ii) content viewable: to allow DSS users to review the contents of a DSS;
(iii) separation of rules from the operation codes: to allow separate editing of each component in a DSS,
(iv) automatic updates: to ensure the DSS is up to date,
(v) integrated into workflow: to reduce the gap between the DSS software design and actual implementation,
(vi) multiple user: to support the evaluation of many patients at a time,
(vii) support separation of responsibilities: to define clearly the role of the knowledge editor, the domain expert and the end user,
(viii) support composition of rules: to allow new rules to be added in the DSS,
(ix) free choice of knowledge representation syntax: to allow multiple knowledge editor using different syntax to be involve with the DSS design.

Table 2-10 shows the comparisons between some existing clinical DSS systems. All systems listed in Table 2-10 have a similar objective, which is to offer an assessment tool to support decision makers (Jun Hua Li et al., 2009). Except SEBASTIAN (system for evidence-based advice through simultaneous transaction with an intelligent agent across a network), other systems (Arden syntax and GELLO (object-oriented guideline expression language)) have the same limitations.

Table 2-10: The checklist of necessary criteria for a decision support architecture.

<table>
<thead>
<tr>
<th>Criteria (Wright and Sittig, 2008)</th>
<th>Arden Syntax¹</th>
<th>GELLO²</th>
<th>SEBASTIAN³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shareable</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Content viewable</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Separation of rules from the operation codes</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Automatic updates</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Integrated into workflow</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Decision support for multiple patient user</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supports separation of responsibilities</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Supports composition of rules</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Free choice of knowledge representation syntax</td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

¹ Arden Syntax is a type of language that represents algorithms in clinical information systems (Jenders et al., 2003).
² GELLO is a type of independent expression language platform that can be used by any clinical application (Sordo et al., 2003).
³ SEBASTIAN is a Web service-based framework (Kawamoto and Lobach, 2005).
2.4.3 Application of Clinical Decision Support Systems

A clinical DSS is defined as an automatic process that compares patient-specific characteristics against a programmed knowledge base to present recommendations or reminders to the provider at the time of clinical decision making.

The key components of a clinical DSS, which are exclusive in almost all implementations, are the automated delivery process of the alerts or reminders at the point of care, and the patient-specific content resultant from the evaluation of the patient’s information against a set of knowledge ‘rules’ or guidelines (Bryan and Boren, 2008).

While the central attribute of the clinical DSS is its anticipated interaction with and provision of recommendations to health care providers, GPs and nurses (Greenes, 2007), the scope can be applied to the entire health care service (Lovell et al., 2010). Important tools in designing a clinical DSS are: access to data, information and knowledge; models or algorithms to provide the intelligence; and a user interface to enable queries, reports and graphing functions (Shim et al., 2002).

Clinical DSSs have been built to address clinical problems such as diagnosis, prognosis, therapy and side effects (Heathfield and Wyatt, 1993), and the application of clinical DSSs affects the clinician’s performance and clinical outcomes. In the area of drug dose determination, studies of computerised aids have shown improvement in achieving therapeutic levels. There are mixed outcomes in terms of diagnosis, with some studies failing to show the effect of clinical DSSs on diagnostic accuracy, while others have claimed success in identifying high-risk patients (Chase et al., 1983, Er and Temurtas, 2008). Contrary to the findings for clinical DSS for diagnosis, most studies of clinical DSSs that were designed to improve the quality of preventive care have shown a significant effect on clinician performance (McDowell et al., 1989, Rosser et al., 1992). The outcomes of clinical DSSs in the area of active medical care are also generally positive (McDonald, 1976). A number of studies have shown that recommendations generated by clinical DSSs have enhanced clinician fulfilment of practice guidelines for preventive and active care (Johnston et al., 1994, Hunt et al., 1998).

In designing a clinical DSS, it is crucial to develop the models or algorithms to provide the intelligence. In many clinical DSS applications, sophisticated and complex strategies have become essential in handling situations involving uncertainties and probabilities. The type of design techniques employed include artificial intelligence, neural networks, logistic regression
and CART (Gardner, 2004, Ebell, 2010). The following are examples of applications using such techniques.

In the Netherlands, a clinical DSS for asthma and COPD called AsthmaCritic was developed. The system was used in a clinical practice environment to assist physicians in using clinical guidelines in their daily practice. AsthmaCritic relied on electronic patient record data. It was used whenever the physician consulted with an asthma or COPD patient. However, the system only started analysing the recorded data after the consultation had finished. The generated critique reports were presented in the same way as the medical-record system. Thus, from the physician’s viewpoint, AsthmaCritic was a part of the medical-record system. There were two types of feedback provided by AsthmaCritic: critique and transformed clinical measurement. Critique feedback was a patient-specific remark based on his or her present clinical situation, while transformed clinical measurement feedback provided calculations on patient measurements to the physicians. AsthmaCritic’s knowledge base was primarily derived from the asthma and COPD guidelines of the Dutch College of General Practitioners. The system generated approximately three feedback comments per patient. This study found that the most common comment for patients aged over 12 years old was non-compliance to prescription (Kuilboer et al., 2002, Kuilboer et al., 2003).

A clinical DSS system to automatically categorise COPD cases and non-COPD cases was developed by Er et al.. This clinical DSS used a multilayer neural network method. The COPD dataset was obtained from patient’s epicrisis reports stored in a hospital database. Using a ten-fold cross-validation approach, the clinical DSS system obtained an accuracy of 95.43% (Er and Temurtas, 2008).

In a different study, researchers constructed a clinical DSS to supervise physical training in patients affected by COPD. The knowledge base module was developed from evidence in literature reviews and from expert consultations. The developed software offered a new method of rehabilitation in patients with COPD. However, the system is still under development and has not yet been used with patients with COPD (Song et al., 2010). Another study by Zhu et al. investigated the ability of the Support Vector Machine (SVM) and K-Nearest Neighbours methods in a clinical DSS to guide rehabilitation planning for home care clients. This study analysed data collected from eight home care programs in Ontario, Canada. The developed algorithms were expected to classify a client according to their need to perform rehabilitation or to be discharged. The paper concluded that both methods can be used in
developing algorithms in guiding rehabilitation planning for home care clients (Zhu et al., 2007).

In a study by Fonarow et al., a CART was used to develop a risk-stratification clinical DSS for patients hospitalised with acute decompensated heart failure. The CART technique selected the best variable and categorised cut-off values for low, intermediate and high-risk patients. The result was a clinical decision tree that was applicable and practical for mortality risk stratification and that had been tested on a group of 32,229 hospitalised patients. The decision tree identified up to 19.8% of patients who were at risk of in-hospital death (Fonarow et al., 2005, Ebell, 2007).

In a clinical setting, point of care is defined as the place where health personnel and patients come together to make clinical decisions about the patient’s health status (Ebell, 2010). Meanwhile, in a home telehealth scenario, the place where point of care occurs is in a virtual space (i.e., telephone conversation or videoconferencing). Thus, the decision rules of a DSS applied at the point of care should be simple and comprehensible by those human users (Ebell, 2010, Ebell, 2007). The clinical DSS is best designed using clinical decision rules (CDR) as compared to a neural network or SVM techniques. This is because these latter methods are known as ‘black box’ methods, as they do not provide any real-life interpretation that a human user could understand (Ayer et al., 2010, Núñez et al., 2002).
2.4.3.1 Methods for the Development of Clinical Decision Rules

A. Expert opinion and/or evidence

The applications described above used either knowledge from experts or information extracted from original research (meaning that they were evidence-based) to construct the CDR. Experts ranked a series of variables from patients’ histories, physiologic values or tests based on their assessment of the clinical importance of the values (Ebell, 2010). Examples of clinical DSSs that used this method include: a clinical DSS that separately controlled COPD patients’ training sessions (Song et al., 2010); an expert system for diagnosis and therapy in lung transplant patients (Prasad et al., 1996); and SAPHIRE, a system for remote cardiac rehabilitation (Laleci et al., 2008).

B. Multivariate model

A multivariate model based on logistic regression represents the relationship between a vector of predictor variables (\( \mathbf{x} \)) and a dependent variable (\( Y \)) that is dichotomous (binary) (Equation 3).

\[
Y = \ln \left( \frac{P}{1-P} \right), \text{ } P \text{ is the probability that the dependent variable (Y) is true,}
\]

\[\beta_0, \beta_1, \beta_i \text{ are the regression coefficients of } X_1, X_2, X_3, ..., X_i \text{ respectively and,} \]

\[i = \text{maximum number of the regression coefficients.} \]

Equation 4 is an example of a multivariate model, developed by Lee et al. to predict influenza cases in patients with febrile respiratory illness (Lee et al., 2011).

\[
Y = -3.1 - 0.5x_{\text{sore throat}} + 0.6x_{\text{running nose}} + 0.2x_{\text{ocular symptoms}} - 0.3x_{\text{nausea}} + 0.4x_{\text{chills}} - 0.7x_{\text{photophobia}} + 0.5x_{\text{fever} \geq 37.8^\circ C} + 0.8x_{\text{fever} \geq 38^\circ C} - 0.4x_{\text{injected pharynx}}
\]
Some other studies have assigned points based on the $\beta$ coefficient values allied with the predictor variables in a multivariate model. The usage of a point score method is illustrated by the CDR to identify children at low risk for appendicitis as constructed by Kharbanda et al.. By using the output of a multivariate logistic model, the study created six types of score from six types of independent variable: nausea (two points), a history of focal right lower quadrant pain (two points), migration of pain (one point), difficulty walking (one point), rebound tenderness and pain with percussion (two points) and an absolute neutrophil count (six points). A patient with a total score of less than five was categorised as low risk (Kharbanda et al., 2005).

C. Classification and regression tree (CART)

A CART is a nonparametric classifier model that uses a tree structure, formed using a set of if-then-else logical conditions to assign an unknown vector of predictor values (or features) to a predefined class, or category. The training of the CART involves examining the predictor variables, finding those variables that are most predictive and applying specific cut thresholds to each predictor variable so as to split the data into two or more classes accurately.

Let us say that a training set contains $N$ instances, or observations, of $Y$, and $x$ is associated with each of the $N$ instances. Each instance is denoted by the 2-tuple $(Y, x)$ and $x = (x_1, x_2, ..., x_l)$. The training of the CART begins with a single root node containing all instances. The node is then split using a single predictor to create two leaf nodes. All predictors are evaluated at various threshold values to find which predictor (and corresponding threshold) should be used, so as to reduce a chosen global measure of impurity for the tree by the greatest margin (Lemon et al., 2003). A completely pure node contains only instances from one class. The Gini index is used as the global measure of impurity if $Y$ is dichotomous, while if $Y$ is a continuous value, the impurity criterion is the sum of squares (see Equation (5) (Breiman et al., 1984).
\[ I(t) = \begin{cases} \text{for binary } Y & 1 - \sum_{j=1}^{J} p^2(j|t) \\ \text{for continuous } Y & \sum_{j=1}^{J} (y_j - \bar{y})^2 \end{cases} \]

Where:

\( I(t) = \) Impurity value at branch \( t \)

\( J \subseteq N \)

\( p = \) proportion of both class at \( j^{th} \) instance

\( y_j = \) predicted \( y \) value at \( j^{th} \) instance

\( \bar{y} = \) mean of the value of \( y \) from \( j^{th} \) instance to \( J \)

The splitting process is repeated for all existing leaf nodes, until all leaf nodes reside no greater than some predefined depth from the root node (Breiman et al., 1984, Deconinck et al., 2006). The tree depth is defined as the maximum number of branches (a branch joins two nodes) on the path from any leaf node to the root node.

Another example of a CART was built by Lemon et al. to classify the prevalence of the influenza vaccination into five different groups. The CART is shown in Figure 2-5. This study used a CART technique as a promising research tool to identify populations at risk in public health research (Lemon et al., 2003).
When comparing CART and multivariate logistic regression methods, the former has a number of advantages over the later method;

I. CART is naturally non-parametric. It means no assumptions have to be made concerning the type of distribution of the predictor variables values. While, in multivariate logistic regression, predictor variables that are highly skewed or multi-modal, or in categorical form need to be transformed or normalised. Thus, this advantage allows the CART analysis to eliminate time used in the analysis which would then be spent defining whether variables are normally distributed, and making alterations if they are not (Lemon et al., 2003).

II. Secondly, logistic regression models are vulnerable to the multicollinearity problem, CART trees are not affected with this problem as they do not treat interactions between predictor variables as concisely as logistic models (Altman and Bland, 1994).

III. Finally, CART trees are comparatively simple for non-statisticians to understand. CDRs based on trees are more likely to be viable and practical, since the assembly of the rule and its inherent logic are obvious to the clinician (Lewis, 2000).
Prior to discussing the applications of clinical DSS’s in home telehealth systems, the next section will briefly discuss each sub-method towards the development of a CDR.

2.4.3.1.1 Pre-processing

In general, pre-processing describes any type of processing performed on the raw data to prepare it for another processing procedure. In the case of signal data, this usually involves filtering of the signals to reduce noise components, or to enhance prominent features. However, in the circumstance of raw data, this typically includes removing outlier data.

The earliest and the most typical algorithm used for outlier detection is the statistical method (Barnett and Lewis, 1984). Laurikkala et al. described a technique that used box plots to identify outliers in both multivariate and univariate data types. The technique creates a graphical illustration and permits a human assessor to visually identify far-off points. Box plots show the lower extreme, lower quartile, median, upper quartile and upper extreme points. The outliers are the points outside the lower and upper extreme values of the box plot; Laurikkala et al. recommend an empirical value of 1.5 times the inter-quartile range outside the upper and lower extremes for outliers, however, this could differ through different data sets (Laurikkala et al., 2000, Hodge and Austin, 2004).

2.4.3.1.2 Feature extraction

Feature extraction is the process of compressing the data into a number of features, which contain information directly relevant to the challenge of discriminating between a number of possible decision outcomes.

The types of features extracted are completely dependent on the classification task. Hence, domain knowledge is dominant in the design of this stage of the system. For example, cardiac patients showing decreased heart rate variability are known to be at a higher risk of sudden cardiac death (Kleiger et al., 1987). Therefore, if the task is to identify these patients, an interesting feature might be the standard deviation of the heart rate interval. Thus the feature extraction routine, for this feature, is to calculate the heart rate intervals and then find the standard deviation of this list of intervals.

The feature extraction becomes vastly more interesting when longitudinal records are involved; that is, when the measurement is performed at regular time intervals. In this case, we can utilise
any trend and distribution information of the features to identify a pattern that is indicative of a pathological development. These longitudinal records are sometimes required, since ‘snapshots’ of a feature treated independently from data recorded at other times, may not contain enough information to make a reliable decision. Longitudinally based features may prove useful as they can encode the normal physiological behaviour for each individual.

2.4.3.1.3 Cross validation

Cross-validation is a process for assessing the classification or predictive ability of an algorithm on unseen data. Presume that \( a \) data points are presented. The data set is divided into two parts; the first part contains a number of data points \( (b) \) used for constructing the algorithm, and the second part contains a number of data points \( (c) \) for assessing the analytical ability of the constructed algorithm. Nevertheless, the rule of thumb is; \( a = b + c \) (Shao, 1993). The most unbiased cross-validation technique is the leave-one-out cross-validation, as in Equation (6) (Cawley and Talbot, 2003, Cawley and Talbot, 2004).

\[
E[p_{\text{error}}^{\ell-1}] = E\left( \frac{L(x_1, y_1, x_2, y_2, \ldots, x_{\ell}, y_{\ell})}{\ell} \right)
\]

Where;
- \( p_{\text{error}}^{\ell-1} \) = the probability of test error of a classifier trained on a sample size \( \ell - 1 \)
- \( L(x_1, y_1, x_2, y_2, \ldots, x_{\ell}, y_{\ell}) \) = the method to measure the number of leave-one-out errors for a classifier trained on a set of input-target pairs \( \{(x_i, y_i)\}_{i=1}^{\ell} \), of size \( \ell \)

2.4.3.1.3 Performance measurement

The main reason to develop a DSS is to use it as a tool to assist the clinician or nurse, thus we need to know the probability that the recommendation given by the DSS is reliable. The accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) should be calculated. The definitions of these values are shown in Table 2-11 (Altman and Bland, 1994).
Table 2-11: The definition of accuracy, sensitivity, specificity, PPV and NPV

<table>
<thead>
<tr>
<th>Classifier output</th>
<th>Reference (Gold Standard)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reference Positive</td>
<td>Positive</td>
<td>Reference Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>True Positive (TP)</td>
<td></td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>Negative</td>
<td>False Negative (FN)</td>
<td></td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>

PPV = \frac{TP}{TP + FP}

Sensitivity = \frac{TP}{TP + FN}

Specificity = \frac{TN}{TN + FP}

Accuracy = \frac{TP + TN}{TP + FP + FN + TN}

In addition to the abovementioned values, there are several performance measurement techniques that can be used such as Receiver Operating Characteristics (ROC) curves and Cohen’s Kappa. An ROC curve is a graph of sensitivity against (1 – specificity) for all possible cut-off points. Using this technique, the classifier’s performance can be quantified by calculating the area under the curve (AUROC). An AUROC equal to one shows that the model gives a perfect test result, whereas a random guess would have an AUROC of 0.5 (Shao, 1993).

Cohen’s Kappa is a meter for classifier accuracy. Cohen’s Kappa is used to estimate the degree of agreement between the classifier and the reference standard (Cawley and Talbot, 2003). The Cohen’s kappa (κ) value can range from 0 (random classification) to 1 (perfect agreement). Equation 7 shows the calculation of κ (Kohavi, 1995);

\[ \kappa = \frac{p_0 - p_c}{1 - p_c} \]  \quad (7)

Where;

\( p_0 \) = the total agreement probability, or accuracy

\( p_c \) = the ‘agreement’ probability which is due to chance

The CART analysis searches for thresholds in the data space, it does not make available continuous response curves to demonstrate predictor-response relationships (Vayssières et al., 2009), thus ROC technique is not suitable to measure the performance of a CART model.
Even though Cohen’s Kappa is a single-scalar meter, it is useful in a multi-class classifier, which is the case of CART, as compared to ROC curves, which are only applicable to two dimensional classifiers (Cawley and Talbot, 2003).

The following section describes some applications of clinical DSS in home telehealth systems, in which the clinical DSS knowledge base is integrated with electronic health records containing individual patient data, facilitating automatic calculations and the streamlining of decision-making (Ebell, 2010).

### 2.4.4 Clinical Decision Support Systems and Home Telehealth

Most of home telehealth integrated with DSS have similar components, which are; 1) a monitoring station in the patient’s home; 2) a central server; 3) a web-based services; and 4) a knowledge based module (Finkelstein et al., 2000, Finkelstein et al., 1998, Baldinger et al., 2004, Demongeot et al., 2002, Rialle et al., 2003, Prasad et al., 1996, Karl et al., 2006, Finkelstein et al., 2005). However, some of the systems have used a slightly different concept. For example, a telehealth system developed in a project called the HEARTS project (Home-based Everyday Activities analysis and Response Telecare System) categorised the mechanism into only two components, local and remote. Data were collected locally, with processing and response mechanisms obtaining data, performing preliminary checks and conveying feedback to the patient when appropriate. Remotely, there were server and response mechanisms, which performed precautionary medical inference and medical decision making, based on the collected information (Raptis et al., 2009).

Furthermore, a system called In-Home Monitoring System (IMS) designed to monitor CHF patients in the home setting, commercialised by Honeywell HomMed consists of four separate remote monitoring units in the patient’s home. Each of the remote monitoring units includes a facility to provide the patient with audio prompts, which reminds the patient to perform measurements and leads the patient through a directed measurement routine. The measurement units communicates wirelessly with a local hub, which uses a conventional dial-up modem for relaying information to a central database over standard public switched telephone network (see Figure 2-6). The data are accessed by the healthcare facility server through either an internet protocol or a conventional modem (Peddicord and Tabor, 2002, Honeywell HomMed LLC, 2007).
Figure 2-6: The IMS architecture.

Part 1 is the data acquisition component. It includes a non-invasive blood pressure monitor unit, a pulse oximeter unit, a weight scale and an electronic thermometer; all except the thermometer have the ability to provide voice prompts.

Another example is the Telematic Management of Insulin-Dependent Diabetes Mellitus (T-IDDM) project that was designed for insulin dependent patients. The system was an integration of two modules: the Patient Unit and the Medical Unit. These modules communicated with each other via the Internet. Patients were able to upload their measurement data from a blood glucose monitoring device to a hospital database (Bellazzi et al., 2002, Bellazzi et al., 2001, Bellazzi et al., 1998).

A system called Mediville designed for elderly and cardiac disease patients who need continuous observation in their daily life comprised portable wireless monitoring equipment located in the patient’s home and a remote server located at the surveillance centre. The portable sensor contained ECG electrodes and an accelerometer. An alert was generated and transmitted to the surveillance centre if a cardiac event or a fall was detected, however, the study did not mention whether data could be accessed using a web interface (Baldinger et al., 2004).

In addition, a more comprehensive design is presented by an intelligent health care monitoring architecture called SAPHIRE. The design was developed to provide a clinical DSS for remotely training cardiac patients for rehabilitation purposes. The architecture was built such that all components understood each other, enabling communication with various clinical systems. The inputs for the clinical DSS not only came from patients’ vital signs received from wireless medical sensors, but also from electronic health care records from the database (Laleci et al., 2008, Nee et al., 2008, Hein et al., 2006).
Reported systems in general used either rule-based approaches, or trend and pattern recognition techniques for the DSS knowledge base. For example, in a trial of home monitoring of asthma severity collected spirometry and questionnaire data were analysed and trends or abnormalities, were extracted using pre-programmed algorithms. If an alert was detected, an alarm was triggered and both the patient and the physician were informed. The system also provided automatic tools to assess patients’ compliance and to provide significant reciprocal data exchange between patient and physician. Web-based services allowed the physician to review the test results, to analyse data trends and to compare past and recent results. Of the patients that used the system, 94.1% expressed their interest in continuing to use home asthma telemonitoring (Finkelstein et al., 2000, Finkelstein et al., 1998).

The Lung Transplant Home Monitoring (LTHM) system used a computerised rule-based clinical DSS for monitoring lung transplant recipients from their home. The system used portable electronic spirometers to record patients’ daily pulmonary function parameters and symptoms, such as high temperature, shortness of breath, wheezing, coughing, sputum amount and sputum colour. The LTHM was an internet-based system that issued reminders and provided access to educational material and guidelines from the health care provider to post-transplant patients. The associated clinical DSS system categorised patients into two conditions: stable or improving; or likely to develop a bronchopulmonary event. The implemented rule-based classification models were devised by clinical experts, according to clinical guidelines using the spirometry measurement values and the symptoms. The system’s sensitivity and specificity were both more than 90% when compared with nurse evaluation (Karl et al., 2006, Prasad et al., 1996, Finkelstein et al., 2005).

The HEARTS study used signal processing, data analysis algorithms and time varying pattern identification and prediction of important health variables in the system design. However, the authors did not elaborate on the algorithm development process (Raptis et al., 2009). Similarly, the IMS stores collected data and analyses the data for trends over selected time intervals. The data are sent to the clinician to be diagnosed and/or classified. If the clinician does not acknowledge these data within a pre-determined time, the server generates notification messages to alert the clinician (Peddicord and Tabor, 2002, Honeywell HomMed LLC, 2007).

The T-IDDDM system equipped the physician with a set of decision support tools that enabled them to visualise and analyse the data. Physicians were also empowered to send their therapeutic advice back to the patient. Patients could also consult with physicians about their
current diet and insulin plan. The T-IDDM decision support component employed a Rule-Based Reasoning methodology. First, the DSS calculated the blood glucose level modal day, which is an aggregation of all data for a week or month presented as if all measurements were performed on a single day. The modal day was an indicator of the patient’s average response to a therapy that they followed. If a problem was identified from the modal day calculation, an alarm was raised to indicate the patient tends to be either hyperglycaemic or hypoglycaemic during particular periods of the day. If a problem was detected, the system would provide a suggestion as to how the patient might alter their therapy. The rules revised and adjusted the current insulin therapy recommendation. In the situation where the system failed to control the patient’s condition, the rules were integrated with another Case Base Reasoning engine, which would retrieve and analyse the patient’s history in order to improve the outcome of the final decision (Bellazzi et al., 2002, Bellazzi et al., 2001, Bellazzi et al., 1998).

The Mediville system was later extended to the TelePat system, in 2006. Several IR sensors were placed in the patient’s bedroom to augment the existing ECG and accelerometer sensors; these IR sensors were intended to assist a falls detection algorithm. A second advancement was the addition of a Hidden Markov Model automatic ECG segmentation and classification routine, which could reportedly analyse the ECG for cardiac ischaemia or arrhythmias. Again, the data were relayed to a remote database server for archiving (J. Boudy, 2006).

The abovementioned studies have used a wide variety of approaches. Most DSSs that categorised patients into different stages of health conditions used rule-based approaches, while DSSs that detected an absolute event, such as a fall event, used pattern recognition techniques in the knowledge base module.
2.4.4.1 Clinical Decision Support System for COPD

The Health Buddy system used predefined criteria based on symptoms and measurement data, as entered by users, to stratify patients’ conditions. Patient names were automatically colour-coded by the Health Buddy system’s algorithms. The colour represented the degree of risk, with stable coded as green, caution as yellow and potential for change in health status coded as red. These indicators were based on symptom-based questionnaire responses, medication and monitored parameters. Respiratory nurses reviewed patient answers the morning after and called patients marked with red flags or discrete patients who had frequently been coded with red or yellow flags (Koff et al., 2009).

The integration of the DSS module in the TMC-Home system uses an enterprise application-server framework. The module has the ability to combine a rules engine with statistical analysis tools. The TMC-Home system framework was designed to enable researchers to perform analysis on the acquired data. An available function is the identification of individual measurements that exceed predetermined thresholds. Currently, ongoing investigations by others in our laboratory are searching for trends and shifts in specific measured parameter values that are associated with the user’s health status (Yang et al., 2010). The main objective of the TMC-Home system DSS is to provide a means of health risk stratification to manage the health of the patient. The TMC-Home system also records the raw signal data (the source of vital sign data values) as part of the measurement process. This allows for the exploration of signal quality in measurements performed in an unsupervised environment (Basilakis et al., 2010, Lovell et al., 2010, Lovell and Redmond, 2010).

A trial based around the Health Buddy system explored the possibility of early detection and treatment of exacerbations in the application of home telehealth for COPD. However, when self-reported symptom questionnaires were performed by elderly participants, concerns were raised that the tendency to misreport information about their health was high (Bush et al., 1989, Tisnado et al., 2006). Moreover, the Health Buddy system had to use other vendors’ vital sign measurement devices to collect data, which were then entered by patients into the system. Therefore, errors might have occurred from invalid data entry, again more likely when the device was used by elderly patients (Motulsky, 1995).

The TMC-Home system is able to provide the user’s carers or ambulatory health personnel with summarised clinical reports containing aggregated views of all measurement parameter data using a decision support approach. These reports include data highlighting, which direct the
attention of the viewer to important parts of the report and out of range data points. Moreover, the TMC-Home system allows for the exploration of signal quality in measurements performed in an unsupervised environment as the acquired raw data are also sent back to the central server and not just extracted features.

2.4.4.2 Clinical Decision Support System and Data Quality

Clinical DSSs linked to electronic medical records are dependent to the quality of data in the databases (Berner et al., 2005). There are two main characteristics of data quality; the accuracy and the completeness. Accuracy is the belief that data are correct, and completeness is the notion that data are recorded in the database (Hasan and Padman, 2006).

Berner et al. had assessed the impact of incomplete and incorrect medical data types on a clinical DSS that provides risk assessment for gastrointestinal bleeding cases. The results showed that 66% of the cases were incomplete and consequently resulted in 77% recommendation errors when used in the clinical DSS. On the other hand, data accuracy was not a problem with 94% of the data correctly recorded (Berner et al., 2005).

Another study by Aronsky et al. used a routinely recorded computerised patient record in a clinical DSS to assess the risk of mortality in patients with community-acquired pneumonia. The average completeness of all patients’ data was 98%. Data errors were predominantly caused during manual entry by the nurse (77%), with the remainder due to system errors (errors that originated from computerised evaluation as defined by Aronsky et al.) (23%). The data quality issue affected 27.9% of the DSS output, with the system underestimating the patient’s risk of mortality (Aronsky and Haug, 2000).

The studies above indicated that a DSS using poor data quality may generate and send false recommendations to clinical users. Consequently, the quality of data is a significant and a fundamental issue of the design of a DSS (Berner et al., 2005, Hogan and Wagner, 1997, Hasan and Padman, 2006, Aronsky and Haug, 2000, Wagner and Hogan, 1996).
2.5 Discussion

Regardless of the significant burden associated with COPD, the disease attracts less interest than other chronic conditions such as heart disease or cancer. This is because COPD is an incurable disease, considered self-inflicted and fairly resistant to treatment. If the burden of this disease is to be reduced, it is important to improve the management of the disease (Voelkel, 2000).

To improve the management of COPD, there is a need to understand the health conditions of COPD patients more completely. While the primary cause of COPD progression is COPD exacerbation events, there are no standardised concepts to recognise their occurrence. Consequently, these events often remain unreported (Caramori et al., 2009, Langsetmo et al., 2008). Therefore, it is vital to recognise markers, for use in patient risk stratification and to manage COPD exacerbation (Farkas et al., 2010, Miravitlles, 2009). Currently, general guidelines use qualitative symptoms or clinical indicators, with three types of measurement used in deciding the need to treat COPD exacerbation at the clinic or hospital. These measurements are chest x-ray, arterial pH level and PaO₂. However, these can only be performed at the hospital or clinic (Bellamy and Booker, 2006, Pearson, 1997, MacIntyre and Huang, 2008), making it is necessary to explore physiological measurement options that can be easily measured and made available in the patient’s home environment. This would assist in predicting exacerbation events, which can occur during the patient’s daily life.

At the same time, there is increasing demand for standard care schemes to assist in COPD management. These schemes provide home visits by ambulatory care teams or community nurses to monitor the enrolled patient’s condition (Ansari et al., 2009). An ideal standard care service should be continuously available and accessible to all patients, regardless of their residential location (Matheson et al., 2006, Stevenson, 2007, Lamothe et al., 2006, Goodridge et al., 2010). To achieve this purpose, home telehealth, a system that provides remote health care to patients in their home environment, has been used. The application of the home telemonitoring type of home telehealth with COPD patients has demonstrated significant positive outcomes (Polisena et al., 2010).

The integration of home telehealth system in a standard care environment has resulted in changes in ambulatory care workload management. Therefore, to assist the ambulatory team in organising visits and monitoring and assessing subjects’ telehealth measurement data, a customised DSS is proposed, to be integrated within the home telehealth system. The DSS
framework should be able to automatically gather electronic health data into a standardised format, synchronise automatically, validate the data quality and share the knowledge among human users in a way that suits the workflow route of patient care (Gardner, 2004).

To ensure the guidance provided by the DSS is reliable, the quality of the data used in the system should be assessed prior to utilisation. The data collected by the home telehealth system should be free from erroneous values and incompleteness of data should be addressed (Hasan and Padman, 2006, Aronsky and Haug, 2000).
Chapter 3: Problem Statements

3.1 Introduction

This chapter presents a synthesis and interpretation of the literature reviewed in chapter two and summarises the problem statements that underpin the thesis work. The four main issues and objectives that are explored are: 1) DSS design; 2) COPD exacerbation prediction; 3) workload management; and 4) home telehealth data quality.

3.2 Problem Statements

The first problem statement relates to:

1. The design of a DSS framework that can be integrated with available home telehealth systems

There are multiple human users involved in standard care management integrated with remote monitoring: GPs; nurses; patient carers, including relatives; and patient users. To reduce the complexity associated with the involvement of multiple stakeholders, two principal roles are proposed; patients and carers. The patient group includes the elderly or people diagnosed with COPD. The carer group includes relatives, nursing staff or GPs caring for the patient. A carer may be responsible for monitoring more than one patient and a patient may have more than one carer.

However, there are also several components in the management system, such as the measurement devices and remote databases. All of these elements, regardless of whether they are humans or technological agents, need to be integrated in the same workflow in order to maximise the home management outcomes (Cumberlidge, 2007).

To solve the integration and the issue posed by having multiple components, it has been proposed in this thesis that the DSS should be designed using a BPM approach that could manage both humans and systems (i.e., home monitoring device, web interface and database).
Moreover, to our knowledge, there are no studies that have attempted to:

2. Predict the onset of COPD exacerbation using home telehealth physiological measurement data

Section 2.2 in this thesis discussed the economic burden of COPD exacerbation (Sullivan et al., 2000). Higher rates of hospitalisation and duration of hospital stay occur among patients with COPD when there is any delay in treatment of an exacerbation episode. Unfortunately, recognition of the start of exacerbation is not a straightforward task, but early detection of this event would result in a reduction in the exacerbation recovery time and risk of hospitalisation for these individuals (Miravitlles, 2009, Wilkinson et al., 2004, Chandra et al., 2009, Seemungal and Wedzicha, 2009).

A study conducted by Trappenburg et al. used a clinical COPD questionnaire as a tool to detect exacerbation events. The patients in the study were asked to answer the self-administered questionnaire weekly, for six consecutive weeks. The study concluded that the questionnaire assessment method can detect exacerbations. However, the method was not designed for real-time daily exacerbation detection (Trappenburg et al., 2010), and data recording using diary cards results in inconsistencies, as each individual patient interprets his/her symptom severity differently (Calverley, 2005).

Moreover, a study by Marin et al. has shown the ability of the BODE index to forecast exacerbation in patients with COPD. The BODE index is a multidimensional index derived from body mass index (BMI), FEV$_1$, a dyspnea questionnaire and exercise capacity (6MWD) (Marin et al., 2009). To avoid any occurrence of adverse events during the 6MWD, the test needs to be conducted at a facility where a fast, proper response to any unfolding emergency is feasible and needs to be supervised by certified healthcare personnel (Brooks et al., 2003), thus it is not appropriate for performance at the patient’s home.

As a means of detecting exacerbation earlier, and at the resolution of a single day, it has been proposed that patients with COPD might use a home telehealth service daily to evaluate their health status (Seemungal and Wedzicha, 2009). The resulting longitudinal physiological records may prove useful in determining the stability of a COPD patient, and in predicting an exacerbation event (Barnett, 2006, McKenzie et al., 2009, Rodriguez-Roisin, 2009).
3. Assist the ambulatory care team or nurse carer in managing monitoring workload for patients with COPD by using home telehealth systems

Home telehealth does not intend to replace traditional contact between carer and patient, however, it can enable a clinical team to work more efficiently (Darkins and Cary, 2000). For example, one study has shown that standard care incorporating a home telehealth system allows nurse carers to effectively manage three times as many patients per day (15 to 22 patients per day) when compared to a conventional care paradigm which includes only home visits (5 patients per day). This improved productivity resulted in a 33% to 50% cost saving for home health care (Doolittle, 1997, Doolittle, 2001). Hence, home telehealth has demonstrated its potential as a means to deliver effective healthcare over large distances using less personnel than existing community care models (Darkins and Cary, 2000).

Conventional methods for assessing the condition of patients enrolled in a remote monitoring integrated standard care involve: 1) monitoring of physiological measurement data (self-measured by the patients) and questionnaire answers; 2) consulting patients or patients’ family carer via 24/7 telephone hotline, and 3) visiting the COPD patients (either scheduled or unplanned visit). These assessment tasks must be performed by the clinical care team daily, without interruption. Consequently, based on the information obtained from the abovementioned activities, the clinical care team provided a recommendation on the course of action; either for the patient to manage their condition at home, or they would refer them for acute clinical assessment.

Whenever a home telehealth system is used to enhance a standard care model, the captured physiological measurement data need to be monitored by the ambulatory care team on a daily basis. Subsequently, if more patients enrol in the home management service, the workload associated with manual data processing will likely increase. Consequently, this increase workload could decrease the responsiveness of the care team, and more time will be needed for the monitoring tasks (Rochford et al., 2009, Roberts and Robinson, 2011, Vitacca et al., 2010). Moreover, manual data processing could introduce mistakes associated with human error (Sintchenko and Coiera, 2008).

Thus to assist with the monitoring task, the complexity of the available data should be simplified to a set of distinct concepts (Clifford et al., 2009) to make it easier for the ambulatory care team to make choices concerning patient care based on this information. Moreover,
recommendations should also be provided to the ambulatory care team for them to manage the consultation time with patients or patients’ family carer (i.e., limiting the consultation time use for low priority cases). Therefore, a DSS which can assist with routine tasks has been proposed (Sintchenko and Coiera, 2008).

The DSS should play an assistive role in clinical decision-making. The specific functionality of the DSS should include the provision of patient-specific recommendations, based on patient physiological measurement data and questionnaire answers (Finkelstein, 2003), providing the ambulatory care team with the ability to prioritise monitoring tasks and manage consultation times.
4. Perform an analysis of the effect of home telehealth data quality on DSS accuracy

Measured home telehealth vital sign data, such as saturation of peripheral oxygen (SpO₂), heart rate and blood pressure (systolic and diastolic) are derived automatically from acquired signals. Currently, signal quality measures linked with pulse oximetry and blood pressure vital sign data are not stored in the TMC-Home telehealth system. This is because there are no standard approaches in handling the contaminated raw signals. Therefore, if an extracted measurement parameter is of concern because it is excessively high or low, it is difficult to determine the credibility of the given parameter (Clifford et al., 2009).

Unless DSS implementations can efficiently deal with such data quality problems, it is uncertain whether the guidance provided by these systems can be reliable. Subsequently, data quality needs to be assessed, because erroneous recommendations based on inaccurate data could influence DSS acceptance and may raise liability issues (Aronsky and Haug, 2000).

Three studies have examined the impact of data quality on DSS performance. However, two of the studies analysed medical data recorded in supervised clinical environments and the third study used simulated data (Berner et al., 2005, Hasan and Padman, 2006, Aronsky and Haug, 2000). It can easily be argued that home telehealth data quality is going to affect the performance of a DSS, as the measurement is performed in an unsupervised environment. Therefore, it is important to extend the existing research base by analysing the effect of home telehealth data quality on DSS accuracy.
3.3 Objectives

1. To design a DSS for a home telehealth application

2. To apply CART to home telehealth physiological measurement data for early identification of patients, with a background of COPD, who appear to be at a high risk of exacerbation events

3. To develop a DSS which can make referral recommendations using routinely recorded data from home telehealth recordings

4. To study the effect of unsupervised home telehealth data quality on DSS accuracy.
Chapter 4 : Decision Support System Design

4.1 Introduction

This chapter discusses the proposed design of a DSS for supporting telehealth management of COPD patients; from hereon the DSS is called a Home Management DSS (HM-DSS). First, the management process is outlined, then the BPM concept is used to understand the specific activities in the process and visualise the process flow. Next, the design architecture is proposed to meet the process requirements. Finally, the technological solution used to implement the design is presented.

4.2 Background

Telehealth management systems for chronic disease patients ideally combine information technologies with clinical management functions to administer clinical guidelines. The principal approach of a BPM is to use processes to improve the performance of such a system. In a broad sense the BPM processes can manage tasks performed by humans, systems, applications as well as the associated data flow. It is a set of methods, tools, and technologies used to design, perform, analyse, and control the associated processes. The BPM system is a general software system that is determined by explicit process representations to organise the execution of business processes, which is a set of activities that are performed in a sequence (Weske, 2007).

In a telehealth management context, the nurse carer or the patients can perform business process activities manually or by the help of the information system. In addition, there are also business process activities that do not need any human involvement, these types of activities are accomplished automatically by the information system.
4.3 Design concept

The design has used business processes to incorporate all the components involved in home management using a representative telehealth system. These include human users, databases, and a home monitoring unit. The integration allows rules (as developed in chapter six and chapter seven) to be used as part of the business process tasks. The execution of these rules uses information gathered from the abovementioned humans and systems. Thus, it is important for both these elements, including the rules engine, to be managed in one process.

The process aims to perform systematic actions specifically to achieve the goal, which is to manage COPD patients. Stated informally, the process for the home management of COPD patients using telehealth and HM-DSS is the following:

1- The patients perform the measurements and answer questionnaires using the home monitoring unit
   a. If the patient is feeling unwell, the patient or the patient’s family carer could contact the clinical carer via phone for consultation
2- The HM-DSS processes the data and provides a recommendation and/or alert to the clinical carer
3- The clinical carer views the measurement data remotely and reviews the HM-DSS recommendation and decides on the course of action (whenever an alert is received or at the normal scheduled time or whenever the patient or patient’s family carer are contacted for consultation)
4- The decision is conveyed to the patient or the patient’s family carer

In this chapter, we have used the business processes to design and execute the process that encapsulates the above workflow taking into consideration (Havey, 2005):

1- Who are the persons and what are the elements involved (such as database, home monitoring unit or web interface) in the process (Q1)
2- Which activities are executed by which components (Q2)
3- What are the inputs and outputs of the activities (Q3)
4- What are the sequence of the activities (Q4)
Table 4-1 (a-d) shows the answers for Q1, Q2 and Q3 for each step (1-4) described in the above process. Figure 4-1 shows the sequence of activities (Q4) and the interaction between components in the flow of information.
Table 4-1: Circumstances in the home management of COPD patients using telehealth and HM-DSS.

(a)

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Patients</td>
<td></td>
<td>a) Carers (family or clinical) <em>create</em> a measurement schedule for the patients</td>
<td>a) <strong>Input</strong>: manual entry by the carer on the web interface; <strong>Output</strong>: HM-DSS send reminder to the patients</td>
</tr>
<tr>
<td>b) Home monitoring unit</td>
<td></td>
<td>b) Patients <strong>perform</strong> the measurement and answer the questionnaires as scheduled</td>
<td>b) <strong>Input</strong>: physical manoeuvre and answering questionnaires by the patients; <strong>Output</strong>: measurement data and questionnaires response</td>
</tr>
<tr>
<td>c) Clinical carers*</td>
<td></td>
<td>c) Home monitoring unit <strong>stores</strong> the measurement data</td>
<td>c) <strong>Input</strong>: Phone call by the patient and patient’s family carer; <strong>Output</strong>: course of action*</td>
</tr>
<tr>
<td>d) Phone*</td>
<td></td>
<td>d) Remote database <strong>stores</strong> the measurement data from the home monitoring unit</td>
<td></td>
</tr>
<tr>
<td>e) Clinical carers consult the patients*</td>
<td></td>
<td>e) Clinical carers consult the patients*</td>
<td></td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) HM-DSS</td>
<td></td>
<td>a) HM-DSS <strong>recalls</strong> the data from the remote database</td>
<td>a) <strong>Input</strong>: Data from remote database; <strong>Output</strong>: Emails</td>
</tr>
<tr>
<td>b) Clinical carers</td>
<td></td>
<td>b) HM-DSS <strong>processes</strong> the data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) HM-DSS <strong>sends</strong> the data to the clinical carers via email</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Remote database <strong>stores</strong> the output from the HM-DSS</td>
<td></td>
</tr>
</tbody>
</table>
(c)

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
</table>
| Answers   | a) Clinical carers  
b) Remote database  
c) Web interface | a) Web interface to **visualise** the measurement data  
b) Clinical carers **monitor** the measurement data and questionnaires answers  
c) Clinical carers **review** the recommendations provided by the HM-DSS | a) **Input**: data from remote database; **Output**: visual presentation of data on the web interface  
b) **Input**: visual presentation of data and emails; **Output**: course of action |

(d)

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
</table>
| Answers   | a) Clinical carers  
b) Patients  
c) Phone/email | a) Clinical carers **consult** the patient users via phone or email | a) **Input**: decision and recommendation from clinical carers; **Output**: patient’s response |

* Circumstances whenever consultation is initiated by the patient or the patient’s family carer
The ellipse and rectangular shapes represent the manual and automated activities respectively. The home monitoring unit is the measurement device located at the patient’s home and authorised humans can access the web interface. Telehealth data is stored in a database. The carer and patient can assume two possible user roles in the HM-DSS. Steps 2, 5, 6 are performed by the system. The carer and patient user respectively perform step 1, 3, 6 and 7 manually and the home telehealth unit automatically performs step 4.
There are different human users and different systems involved in each process stage. Based on Table 4-1, we can distinguish the manual and automated activities. The human users perform manual activities, while the HM-DSS executes the automated activities. The behaviour of the manual and automated activities are different from each other. For the manual activities, we have to wait for the human users to respond, while for the automated activities the behaviour is to execute as fast as possible.

The sequence of activities is documented in seven steps (see Figure 4-1). There are two groups of human users; patient users and carer users and three types of systems involved. Steps 2, 5 and 6 indicate the main functionalities of the HM-DSS. The system workflow is triggered by a reminder created by the carer at step 1 at a predefined time. The reminder is to notify the patient (step 2) to perform their routine measurements and to answer their questionnaires (step 3). To create the reminder, the carer uses the web interface (the schedule is automatically stored in the database). The HM-DSS sends the notification that is eventually displayed on the home telehealth console. After a patient has performed the measurement, the data are routinely stored in the database (the home telehealth unit synchronises with the database on an hourly basis) (step 4). Subsequently, the system periodically checks the database (step 5). The data availability is compared with the schedule and if some data are missing, a notification is re-sent (step 2) to advise the patient of the missing measurements. Then, the HM-DSS analyses the data and sends a report containing recommendations to the carer (using email) and/or generates an alarm (step 6) depending on the urgency of the clinical scenario. The clinical carer reviews the data via the web interface and the recommendations from the HM-DSS and decides on the course of action. Step 7, which is the consultation between the clinical carers and the patient, could also happen whenever the need arose.

In our design, the home monitoring system used is the TMC-Home telehealth system, earlier versions of which having been developed previously by our research group. The telehealth system has two main components, the home monitoring unit that is usually placed at the patient’s home and the web interface, which can be accessed using any web browser by user authorisation. The TMC-Home telehealth system could send the data to the remote database via phone line or internet connection. The overall system architecture for the TMC-Home is illustrated in Figure 4-2.
Figure 4-2: Schematic of overall workflow for the TMC-Home.

Vital signs and other data including questionnaires are remotely collected and transferred via the Internet to a central hosting service. Care providers can access these data by way of a secure Web browser and change measurement schedules or medications reminders. The ambulatory care team and the patients could also communicate via phone.
4.4 Design architecture

Three main components used in this design: are the data access layer; the logic layer; and the front-end layer. The data access layer provides the connection to the database. The logic layer holds the steps to perform the activities (the rule engine developed in next chapters is encapsulated within this layer); and the front-end layer facilitates the communication with the user. Figure 4-3 shows the overall design of the system.

The data access layer is the layer that utilises the data from the database thus allowing the system to create, read, update and delete measurement data. This lower level layer connects the higher levels with the database where all the required measurement and questionnaire data are stored. The connection with the database is achieved by mapping the database schema into a form that can be used within the system.

The logic layer holds the system’s core component, namely the process. Each activity in the process is visualised as a node. The final output from the process is entered in an email, which is later sent to the users.

The front end layer is the component that interacts with the patient and carer. Email templates are implemented in the system for notification and recommendation reports. The email is dynamically generated according to the output provided by the process in the logic layer. Therefore, properties such as recipients, attachments do not need to be hard-coded but are also extracted from the database.
Figure 4-3: HM-DSS overall design.

The data access layer is composed of data objects, home and query components (the database figure is the same as mentioned in Figure 4-2). The logic layer is made up of processes, and the mailing components. Each process contains the logic to accomplish a task. A front-end comprises an email template component that is connected directly to the logic layer, the data access layer and the user (patients and carers).
4.5 Technological solution

The HM-DSS has been implemented using the technology developed by JBoss Seam, which is a software technology that knits together a set of code and services into an environment that makes it easier to write a software application. This Java-based framework provides integration with other products such as Hibernate, a Java language library for the database connection, JBoss Business Process Management (jBPM), a software application that relies on the execution of a business process definition, production rule system (Drools) which is a forward chaining inference based rules engine and JavaServer Faces (JSF) technology for the email templates. All these HM-DSS components are run on a server (Salter, 2009, Farley, 2007).

The HM-DSS uses Hibernate as the solution to map the data in the database into data objects in the data access layer. This mapping allows the data object to be used in the logic and the front-end layers. Additionally, the system uses Seam's JSF to link the back-end components that implement the mailing facility. The jBPM Process Definition Language (jPDL), a Java-based process-oriented programming language that creates the process models, by converting the workflow into models that can be implemented in a Java development environment. The rules are translated using the Drools software to enable execution within the business process (Bali, 2009, Browne, 2009).
4.6 Laboratory testing

The proposed system had been tested in a laboratory environment. The design functionality was tested by integrating an existing database (a replication of an actual database used in a home telehealth trial) with the HM-DSS. The database was obtained from TeleMedcare Pty. Ltd. Sydney, Australia and was stored in a Microsoft SQL server version 2005.

4.6.1 Data access layer

For supporting the HM-DSS’s rule engine, the tables from the database as listed in Table 4-2 have been mapped to data objects in the data access layer.

Table 4-2: The description of the tables that have been mapped from the TeleMedcare Pty. Ltd. Sydney, Australia database.

<table>
<thead>
<tr>
<th>Table name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClinicalData</td>
<td>This table provides the definitions of the coded measurements and questionnaires. For example, the code for SpO₂ is ‘11’.</td>
</tr>
<tr>
<td>UserClinicalData</td>
<td>This table stores the actual measurement and questionnaires data values. This table is linked to the User, CodeDictionary and UserClinicalDataRaw tables.</td>
</tr>
<tr>
<td>User</td>
<td>This primary table defined the user. User in the system may be: a) Patient users who are being monitored b) Carer users I. Clinical workers who are looking at the monitored data and devising care plans II. Carers who are assisting users with collecting data, and administering medications</td>
</tr>
<tr>
<td>UserContact</td>
<td>This table contains the list of contacts of the users. For example, their email address and phone numbers.</td>
</tr>
<tr>
<td>UserClinicalDataRaw</td>
<td>This table stores the raw data for the measurements.</td>
</tr>
<tr>
<td>CodeDictionary</td>
<td>This table provide the code for all string names used in the database this table is linked to the ClinicalData table.</td>
</tr>
<tr>
<td>UserHM-DSSReminder</td>
<td>This table stores the scheduled measurements date and time and is linked to the User table.</td>
</tr>
</tbody>
</table>
4.6.2 Logic layer

Figure 4-4 shows the process designed in jPDL. The process starts using the date and time information stored in the UserDSSReminder table. The decision node is used to decide whether the process should go forward or backward. For example if measurement data are missing, the notification node is re-activated. The rules are embedded in the Data Evaluation node. Table 4-3 shows an example of rules translated using Drools. The *WHEN* part in the rules describe the condition, and the *THEN* part show the action (consequent).

![Diagram](image)

**Figure 4-4: A process in jPDL.**
Table 4-3: Example of rules embedded in the *Data evaluation* node.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description*</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>rule &quot;One&quot;</td>
</tr>
<tr>
<td></td>
<td>when</td>
</tr>
<tr>
<td></td>
<td>d : Data(Feature 1&lt;98.00, Feature 2&lt;0.12, Feature 3&lt;1.33)</td>
</tr>
<tr>
<td></td>
<td>then</td>
</tr>
<tr>
<td></td>
<td>System.out.println(&quot;Rule One&quot;);</td>
</tr>
<tr>
<td></td>
<td>System.out.println(&quot;Referral recommendation&quot;);</td>
</tr>
<tr>
<td></td>
<td>End</td>
</tr>
<tr>
<td>Two</td>
<td>rule &quot;Two&quot;</td>
</tr>
<tr>
<td></td>
<td>when</td>
</tr>
<tr>
<td></td>
<td>d : Data(Feature 1&gt;=98.00, Feature 4&lt;2.44, Feature 5&lt;76.42, Feature 6&lt;2.91)</td>
</tr>
<tr>
<td></td>
<td>then</td>
</tr>
<tr>
<td></td>
<td>System.out.println(&quot;Rule Two&quot;);</td>
</tr>
<tr>
<td></td>
<td>System.out.println(&quot;Home management recommendation&quot;);</td>
</tr>
<tr>
<td></td>
<td>End</td>
</tr>
</tbody>
</table>

* The type of features are described in next chapter

4.6.3 *Front end layer*

Table 4-4 shows the JSF elements in the email templates and Figure 4-5 shows the email template.

Table 4-4: Important JSF elements in the email templates.

<table>
<thead>
<tr>
<th>JSF Tag</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;m:to&gt;</td>
<td>JSF message tag for email template. It defines the receiver's address.</td>
</tr>
<tr>
<td>&lt;m:body&gt;</td>
<td>This tag defines the content body of the email. It allows standard HyperText Markup Language or another template to be included in the body.</td>
</tr>
</tbody>
</table>
Figure 4-5: The email template for generating the report.

The email address is obtained from the information in the UserContact table and the measurement data are obtained from the UserClinicalData table. The values are visualised in graphs (see example in Figure 4-6).
Figure 4-6: The measurement data visualised in a time series graph.
4.7 Discussion

A BPM system has been used by Flender et al. to design an architecture for a home telecare model to be used for an enterprise-wide home telecare deployment. However, the design is not meant to assist the clinical carer in making decisions on a patient’s health condition. The model is developed with an objective to improve the healthcare delivery system in general, while our HM-DSS is specifically devised for individual patient management, to convey recommendations from the rules engine to the clinical carer (Flender, 2008).

Adaptation of the BPM concept in a HM-DSS for home management of COPD patients using telehealth has allowed us to understand and formalise the process. For example, a patient user performs a measurement at 8:00 AM in the morning, the clinical carer is scheduled to monitor the measurement at 5:00 PM in the afternoon. However, the HM-DSS detects a condition that requires urgent review from the clinical carers, it sends an alert, and the clinical carer reviews the patient’s information immediately after receiving the alert.

The TMC-Home telehealth system used in the laboratory testing has been pre-programmed to synchronise with the remote database on an hourly basis. However, in a real life situation, the TMC-Home is usually connected using a phone line. Therefore, it is impractical to perform the data synchronisation hourly as it would monopolise the phone line usage. To overcome this limitation, it is suggested that the TMC-Home telehealth service include a broadband connection to avoid phone line interference.

The HM-DSS could send information to the human users via the home monitoring unit or the web interface. This is done by storing the message in the remote database, causing the information to be displayed on the TMC-Home unit console after the next synchronisation. The message would also appear on the web interface, for perusal by the ambulatory care team.

The three-tier HM-DSS approach has the ability to pass data along the process (from the data access to the front-end layers). Moreover, the components that have been proposed in the workflow process use general concepts that will allow re-use in other applications using the same framework and code base.

The abovementioned reason is the rationale of choosing JBoss Seam as the tool for implementation. Seam is a conversational framework as opposed to other frameworks such as
Spring or Struts (Salter, 2009). Seam provides a unified development model, where the data can be accessed from within all three layers of the HM-DSS architecture.

Even though the HM-DSS is a functional system, some aspects can be improved. Nonetheless, when compared to the existing DSS (c.f. Table 2-10) on the necessary criteria for a DSS (Wright and Sittig, 2008), the HM-DSS is the only system that allows the evaluation of multiple users at one time.

The design has used the OSS Seam framework for implementation, to provide a proof of concept of the benefits of HM-DSS approaches for improving aspects of workload, communication and information exchange amongst carers and patients. The HM-DSS needs to be deployed and further evaluated in a real world home telehealth environments.

Link to next chapter

This chapter has presented the HM-DSS architecture design. The most crucial aspect of designing a HM-DSS is to bring together information on patient measurement data and a related set of labelled patient conditions, from which a set of rules might be derived, to be used later in the interpretation of previously unseen patient data. These derived rules must handle data that has inherent uncertainty, commonly resulting from data being generated by unsupervised patients in their own homes. The next chapter will describe the methodology behind the development of the knowledge base rule engine. The method is subsequently used in chapters six and seven.
Chapter 5: Rules Engine Development Methodology

5.1 Introduction

The main component in a clinical HM-DSS design is the algorithms that provide the knowledge. This chapter describes the methodology to develop the rules and associated rules engine that provides this intelligence. The rules were constructed using data from retrospective home telehealth studies. This chapter presents the materials and the general methodology that is used in chapters five and six.

5.2 Materials

5.2.1 Database

Data used in this study came from two different collections but had in common the fact that they were derived from randomised controlled trials comparing standard care from the ambulatory care nurse or the RACS team with remote monitoring integrated with standard care.

The participant inclusion criteria were:

1) English fluency;
2) Finger dexterity to use a keyboard and mouse;
3) Willingness to use a computer in health self-management;
4) Living independently;
5) No major motor deficit that could prevent use of the home monitoring device;
6) Able to give informed consent;
7) Confirmed moderate/severe COPD; and
8) At least one hospital presentation in the last 12 months.

Exclusion criteria were:

1) Significant co-morbidities (e.g., cancer, renal failure);
2) No documented cognitive impairment; and
3) Participation in another trial (Rochford et al., 2009).

Each patient was asked to perform their physiological measurements and complete questionnaires on a daily basis using a TMC-Home (TeleMedcare Pty. Ltd. Sydney, Australia) device.
Consent for the execution of the aforementioned data collection studies were granted by the Austin Health Human Research Ethics Committee (Database I) and Sydney West Area Health Service (SWAHS) Ethics Research Governance Committees (Database II).

### 5.2.2 Data

#### 5.2.2.1 Physiological Measurements Data Types

There are two types of vital sign measurement data, one type is measured automatically, while in the other type the patients enter data manually. Figure 5-1 shows the automatic and manual measurements and Table 5-1 lists the parameters extracted from the measurement types.

![Figure 5-1: TMC Home automatic and manual measurements.](image-url)
### Table 5-1: List of the measurement parameters and the units.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit/category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spirometer</td>
<td></td>
</tr>
<tr>
<td><em>Forced spirometry</em></td>
<td></td>
</tr>
</tbody>
</table>
| 1  
FEV₁                                                                     | Litre                 |
| 2  
FVC                                                                     | Litre                 |
| 3  
Peak expiratory flow (PEF)                                              | Litre/minute          |
| 4  
Forced expired volume in six seconds (FEV₆)                              | Litre                 |
| Relaxed spirometry                                                       |                       |
| 5  
Breathing rate (BR)                                                     | Breaths per minute    |
| 6  
Inspiratory capacity (IC)                                               | Litre                 |
| 7  
Minute ventilation (VE)                                                | Litre/minute          |
| 8  
Tidal volume (VT)                                                       | Litre                 |
| 9  
The time of inspiration to total breath duration (Ti/Ttot)              | Unitless              |
| Pulse oximeter                                                           |                       |
| 10  
SpO₂                                                                    | Percentage            |
| 11  
Heart rate*                                                             | Beats per minute      |
| Weight scale                                                             |                       |
| 12  
Body weight                                                             | Kilogram              |
| Thermometer                                                              |                       |
| 13  
Body temperature                                                       | Celsius (°C)          |
| Blood pressure                                                           |                       |
| 14  
Systolic pressure                                                      | millimetres of mercury (mmHg) |
| 15  
Diastolic pressure                                                     | millimetres of mercury (mmHg) |

* Heart rate value could also be obtained from ECG measurement.

The measurement most commonly employed to evaluate lung function is spirometry. There are two ways to perform spirometry: 1) using forced inspiration or expiration to assess extreme dynamic lung volumes and flow rates; and, 2) using relaxed breathing to assess the limits of normal resting lung volumes and flow rates (Booker, 2006).

In the case of the TMC-Home unit, both types of spirometry measurements are performed using a pneumotachometer respiratory flow sensor. The sensor converts the flow of gases through it into a proportional signal of pressure difference on either side of a central mesh whose design ensures a signal linearity over a range of flow rates. It is connected to a detachable and replaceable mouthpiece through which the patients blow.

Parameters extracted from forced spirometry are FEV₁, FVC, PEF and FEV₆, and parameters derived from the relaxed spirometry manoeuvre are, IC, VE, VT, BR and Ti/Ttot.
Pulse oximetry is based on the principle of photoplethysmography, utilising the effect whereby light is absorbed as it passes through haemoglobin. The pulse oximeter apparatus monitors heart rate and SpO₂ level (Tremper, 1989).

Patients performed the pulse oximetry measurement using an oximeter finger clip known as a ‘pulse-ox probe’ (MCP-NN-OXY, TeleMedcare Pty. Ltd. Sydney, Australia). The probe is connected to the TMC-Home monitor using a workstation connector. To execute the measurement, a patient placed his/her index finger in the probe and pressed a button on the TMC-Home monitor screen.

Naturally, heart rate could also be obtained from the ECG measurement, which is used to monitor the electrical signals associated with the heart muscles. The ECG electrodes are integrated into the moulding of the TMC-Home monitor. Patients performed the ECG measurement by 1) pressing the “Take Measurement” (ECG) button on the main screen and; 2) holding the sides of the monitor (the location of the monitoring electrodes). The system will measure the ECG and derive an average heart rate while displaying the ECG signal on the TMC-Home screen.

The TMC-Home non-invasive blood pressure measurement (NIBPM) module includes a blood pressure cuff with air hose and a auscultatory microphone acting as a stethoscope. During unsupervised measurements, patients have to wrap the blood pressure cuff around their upper arm and press the “Take Measurement” button on the TMC-Home monitor. The cuff is inflated to a preset pressure level (approximately 175 mmHg). After reaching the preset value, the cuff deflates slowly and the microphone records the Korotkoff sounds. The TMC-Home automatically identifies the systolic and diastolic pressure values.

The weight is measured using a weight scale powered by batteries and the body temperature is measured using a thermometer.
5.2.2.2 Questionnaires Data Types

There were in total eight different questions asked during the data collection period. Some minor differences existed between the two databases. Database I recorded five questions while Database II acquired answers from seven questions. Table 5-2 shows the questions and answers.

Table 5-2: Questions and answers asked during the trials.

<table>
<thead>
<tr>
<th>Availability (Database)</th>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
</table>
| I, II                   | Describe your breathing today | 1 – Better than usual  
2 – Same as usual  
3 – Worse than usual  
4 – Much worse than usual |
| I, II                   | How do you feel in general today | 1 – Better than usual  
2 – Same as usual  
3 – Worse than usual  
4 – Much worse than usual |
| I                       | The amount of sputum you produced today | 1 – None  
2 – Some (up to a teaspoonful)  
3 – A little (up to a tablespoon)  
4 – Moderate (up to an egg cupful)  
5 – A lot (more than an egg cupful) |
| II                      | The amount of sputum you produced today | 0 – None  
1 – ½ teaspoon/day  
2 – 1 teaspoon/day  
3 – 1 tablespoon/day  
4 – ½ cup full/day  
5 – more than ½ cup full/day |
| I                       | What is the colour of your sputum today? | 1 – Cream coloured  
2 – Pale yellow coloured  
3 – Dark yellow coloured  
4 – Green coloured  
5 – Other |
| II                      | What is the colour of your sputum today? | 0 – Clear/nil  
1 – Cream coloured  
2 – Pale yellow coloured  
3 – Dark yellow coloured  
4 – Green coloured  
5 – Other |
| I, II                   | How is your sputum consistency today | 0 – watery  
1 – creamy  
2 – thick  
3 – semi solid  
4 – solid (plugs) |
| II  | Cough                      | 0 – nil                        |
|     |                           | 1 – once or twice a day        |
|     |                           | 2 – 3-5 times a day            |
|     |                           | 3 – frequently (5 – 10 times a day) |
|     |                           | 4 – persistently               |
| II  | Nighttime symptoms        | 0 – nil disturbances          |
|     |                           | 1 – wakes occasionally but able to get back to sleep easily |
|     |                           | 2 – wakes regularly due to breathlessness and/or cough     |
|     |                           | 3 – sleep only 1-2 hours total due to breathlessness and/or cough |
|     |                           | 4 – unable to sleep due to breathlessness and/or cough     |
| I   | Do you have cold or flu today | 1-No                         |
|     |                           | 2-Yes                         |
5.3 Methodology

5.3.1 Data Analysis

Figure 5-2: General methodology for the development of the rules engine.

Figure 5-2 shows the general methodology for the development of the rules engine. First, the longitudinally recorded parameters are collected from the retrospective home telehealth databases. Then, health status reference standards are defined, either based on available clinical diagnoses, health service utilisation (i.e. patient’s visits to the clinic and emergency department visits), an informal carer journal, or symptom and medication records data. Additionally, the processes applied to the physiological measurement data are outlier detection and removal, and feature extraction. Finally, using the defined health status standard and the processed data, CART classifiers are trained and tested using a cross-validation method. Subsequent sections will discuss the methods used to process the measurement data.
Since the home telehealth data acquisition protocol was performed in an unsupervised environment, data quality may affect the classifier performance. Therefore, the data collected during the intervention period were manually inspected for outliers; defined as data points that lie very far from the median of the corresponding variable for each patient's data. To determine the allowable range; first, some statistical characteristics of the data (illustrated in Figure 5-3) were obtained; namely, the median (50th percentile), and the lower (25th percentile) and the upper (75th percentile) quartiles, and secondly, the lower and upper thresholds were calculated as the data value on the 25th percentile minus the 1.5 (interquartile range value) and the data value on the 75th percentile plus the 1.5 (interquartile range value) respectively. Any data value outside these limits were considered as possible outliers (Walfish, 2006).

Next, if the possible outlier point was derived from a raw waveform signal (such as SpO₂ from a photoplethysmograph signal, for example), the signal was retrieved and examined visually. Outlier data which was deemed by a human observer to have come from signals corrupted by movement artifact and noise were excluded from further analysis. Alternatively, if the possible outlier value was not derived from a waveform (such as weight or temperature), the values were physiologically validated; that is, only potential outliers still falling within the range of what is physiologically possible were included in the analysis.
Figure 5-3: A box plot to illustrate the characteristic of a patient’s heart rate data.

The dark line in the centre of the box is the median value. The base and top of the box indicates the 25th percentile and the 75th percentile respectively. The T-bars that extend to 1.5 times the height of the box are called inner fences. The circles represent possible outliers. The asterisks are extreme outliers (representing cases that have values more than three times the inter quartile range above (below) the 75th (25th) percentile).

5.3.1.2 Feature Extraction

The baseline value for each parameter varied significantly from subject to subject. Thus, four types of values were calculated from the longitudinal records for the 15 mentioned physiological parameters (Table 5-1): the distribution mean (with the mean calculated using recent data); the percentage change of the parameter from the distribution mean; the distribution standard deviation; and the standard score (also known as the z-score), which is the deviation of the parameter from the distribution mean as a fraction of the distribution standard deviation.

The calculation of the distribution mean and standard deviation of each physiological parameter value was evaluated from a one month epoch of data (or one week if there were not sufficient data) prior to the current day under investigation. The percentage changes and the z-score were then calculated using the physiological parameter values on the day prior to the day in question, relative to the baseline mean and standard deviation over the previous month; that is, use today’s measurement relative to the previous month to estimate tomorrow’s condition. All these values (the actual parameter value, the distribution mean, the standard deviation, the percentage changes and the z-score) are later passed as input to the classifier model to predict incipient exacerbations.

84
An illustrative example of the timeline used for the calculation of the feature values is shown in Figure 5-4.

**Figure 5-4:** The timeline shows an example of how the mean, the standard deviation, the percentage changes and the z-score are calculated.

Tomorrow (1 October) is the day on which the patient is being evaluated as being in either a ‘stable’ or ‘poor’ condition. For the \(i\)th parameter, \(i \in \{1, ..., 15\}\), the mean (\(\mu_i\)) and standard deviation (\(\sigma_i\)) values are calculated using the date range from 31 August to 29 September, inclusive. The measurement data percentage change is calculated using the data value for the \(i\)th parameter on the 30 September (feature vector, \(x_i\)) and the mean, \(\mu_i\). The z-score is calculated as \((x_i - \mu_i)/\sigma_i\).

### 5.3.1.3 Classification

At this stage, five types of features are extracted from each physiological parameter. The five types of features are: the parameter value; the distribution mean from the last month; the standard deviation from the last month, the percentage change of the previous days data relative to the last month, and the z-score value for the previous day relative to the last month. There are 75 features extracted from the 15 parameters listed in Table 5-1.

Two types of multiple features (multivariate) classification are performed:

1) to investigate the individual predictive power of each measurement type, and;
2) to evaluate the collective predictive power of all features.

### 5.3.1.4 CART

From the available training set, let us say that it contains \(N\) observations of the variable \(y\), where \(y \in \{0,1\}\) represents the categories of ‘stable’ and ‘poor’ conditions, respectively. An observation represents a particular day under investigation for which it must be decided if a patient’s ‘poor’ condition is imminent.

The multivariate classification evaluates the pooled predictive power of a vector of features, \(x\). Thus, each of the \(N\) instances is denoted by the 2-tuple, \((y, x)\), where \(x = (x_1, x_2, ..., x_i)\) is the vector of feature variables used, containing;
a) all the features from each individual measurement type, with the minimum value of \( i \) is five

b) all the features from all the parameters, \( i \) is 75 (maximum value)

The training of the CART begins with a single root node containing all \( N \) instances. The node is then split using a single feature to create two leaf nodes; all features are evaluated at various threshold values to find which feature (and corresponding threshold) should be used, so as to reduce a chosen global measure of impurity for the tree by the greatest margin (Lemon et al., 2003); a completely pure node contains only instances from one class. The splitting process is repeated for all existing leaf nodes, until all leaf nodes reside no greater than some predefined depth from the root node (Breiman et al., 1984, Deconinck et al., 2006). The tree depth is defined as the maximum number of branches (a branch joins two nodes) on the path from any leaf node to the root node.

5.3.1.5 Cross Validation

Once the tree has been constructed, its performance is evaluated using cross validation. The cross validation process, used to estimate the expected generalised classification performance, is described as follows. Using data from \( n - 1 \) subjects (where \( n \) is the total number of patients), the CART is trained. The data from the remaining subject is later introduced for testing. This is repeated \( n \) times, removing a different subject from the training set each time. The reported results are the combined results from each of the \( n \) cross validation runs. By using the cross validation method, \( n \) sets of rules are obtained from the \( n \) trained CARTs.

5.3.1.6 Estimating the Optimal Tree Depth

To avoid overtraining the CART, which will invariably occur if the tree is grown to its maximal depth, the optimal depth of the CART is first estimated. This optimum tree depth is estimated prior to training the CART within the primary cross validation loop, described above.

To estimate the optimum tree depth, the set of training instance \((y, x)\), (from \( n - 1 \) subjects in the primary cross validation loop) are randomly divided into ten further subgroups. Nine of these subgroups are used to build the CART, using the method outlined above. Once the CART is trained using these nine subgroups, the remaining group is introduced for testing. This is repeated ten times, removing a different subgroup for testing each time. \( \kappa \) value of each of the ten subgroups, given an allowable tree depth, is calculated. This is repeated for all possible tree
depth values, reducing the depth of the tree on each repetition. The tree depth which gives the maximum averaged $\kappa$ value (averaged across the ten subgroup repetitions) is chosen as the optimum depth (Moisen, 2008). Finally, the tree is trained again using this optimal tree depth, using training data from all ten subgroups combined (that is, data from all eleven subjects in the primary cross validation loop). The withheld subject from the primary cross validation loop is now introduced for testing. This process, as described in the previous section, is repeated $n$ times, each time removing a different subject’s data. Obviously, for each cross validation repetition, a different tree structure, with a different optimal depth may be generated; however, only the aggregated performance results are presented later.

5.3.1.8 Performance Measurements

CART models were developed using data from Database I and Database II. The performances of these models were evaluated by comparing the classification output with the health status reference standards. The CART classifiers sensitivity, specificity, PPV, NPV, accuracy and $\kappa$ value (level of agreement between the reference standard and the CART output) were calculated. The interpretation of the $\kappa$ value is shown in Table 5-3 (Cohen, 1960, Viera and Garrett, 2005).

<table>
<thead>
<tr>
<th>$\kappa$ value</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.00</td>
<td>Less than chance agreement</td>
</tr>
<tr>
<td>0.01-0.20</td>
<td>Slight agreement</td>
</tr>
<tr>
<td>0.21-0.40</td>
<td>Fair agreement</td>
</tr>
<tr>
<td>0.41-0.60</td>
<td>Moderate agreement</td>
</tr>
<tr>
<td>0.61-0.80</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>0.81-0.99</td>
<td>Almost perfect agreement</td>
</tr>
</tbody>
</table>
5.4 Data Collection

5.4.1 Databases

5.4.1.1 Database I Description

Database I consists of data from 12 moderate-severe COPD patients. The mean, minimum and maximum ages of the participants were 69, 46, and 84 years, respectively. Data were acquired over the duration of approximately one year, from February 2007 to January 2008, using a TMC-Home unit placed in the patients’ homes. The subject’s baseline characteristic is presented in Table 5-4.

Table 5-4: Database I subjects baseline data.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Gender</th>
<th>Disease severity</th>
<th>FEV₁ (%)¹</th>
<th>Days enrolled in the trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>59</td>
<td>Female</td>
<td>Severe</td>
<td>31</td>
<td>221</td>
</tr>
<tr>
<td>2</td>
<td>76</td>
<td>Male</td>
<td>Moderate</td>
<td>54</td>
<td>312</td>
</tr>
<tr>
<td>3</td>
<td>74</td>
<td>Male</td>
<td>Severe</td>
<td>43</td>
<td>265</td>
</tr>
<tr>
<td>4</td>
<td>66</td>
<td>Female</td>
<td>Very severe</td>
<td>15</td>
<td>317</td>
</tr>
<tr>
<td>5</td>
<td>78</td>
<td>Female</td>
<td>Very severe</td>
<td>29</td>
<td>289</td>
</tr>
<tr>
<td>6</td>
<td>58</td>
<td>Male</td>
<td>Severe</td>
<td>39</td>
<td>332</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>Female</td>
<td>Severe</td>
<td>31</td>
<td>206</td>
</tr>
<tr>
<td>8</td>
<td>76</td>
<td>Female</td>
<td>Very severe</td>
<td>24</td>
<td>219</td>
</tr>
<tr>
<td>9</td>
<td>78</td>
<td>Male</td>
<td>Moderate</td>
<td>59</td>
<td>335</td>
</tr>
<tr>
<td>10</td>
<td>59</td>
<td>Male</td>
<td>Very severe</td>
<td>29</td>
<td>335</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
<td>Female</td>
<td>Moderate</td>
<td>53</td>
<td>210</td>
</tr>
<tr>
<td>12</td>
<td>84</td>
<td>Female</td>
<td>Severe</td>
<td>48</td>
<td>208</td>
</tr>
</tbody>
</table>

¹Staging criteria based on the GOLD guidelines (Rodriguez-Roisin, 2009)

The database was a part of a randomised controlled study comparing standard best practice care (SBPC) with remote monitoring integrated SBPC. The study was conducted within the catchment area of the Austin hospital in the state of Victoria. In this analysis, only data collected from intervention patients were used.

For the period of the study, the patients’ data were reviewed five days a week by the clinical carers to determine the need for further clinical assessment. There were a number of management options available to the clinical carer if a clinical alert condition was detected. In this study, a clinical alert condition occurred when the clinical carer detected a significant change in a variable (i.e., worsening in any symptom). Figure 5-5 shows the triaging of the remote monitoring patients (Rochford et al., 2009). The clinical carer could call the patient and ask for more information related to the patient’s health; or could arrange a visit by an outreach
nurse to the patient’s home; or could notify the healthcare specialist concerning the deterioration of the patient’s health. In any of these cases, the nurse regularly made some short-form notes in a **carer journal** including points of interest describing each patient’s condition.

**Figure 5-5: Triaging of the remote monitoring patient.**

5.4.1.2 **Database I Data Analysis**

The 12 patients enrolled in the trial for a total of 3249 days. In the analysis, data collected from five types of measurement devices were used: spirometry, pulse oximetry, blood pressure, weight, and body temperature. Table 5-5 shows the number of days on which the patients used each measurement device.

**Table 5-5: Total number of days each measurement device was used by all 12 patients.**

<table>
<thead>
<tr>
<th>No</th>
<th>Measurement device</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spirometer</td>
<td>2427</td>
</tr>
<tr>
<td>2</td>
<td>Pulse oximeter</td>
<td>2406</td>
</tr>
<tr>
<td>3</td>
<td>Blood pressure</td>
<td>2439</td>
</tr>
<tr>
<td>4</td>
<td>Weight scale</td>
<td>2406</td>
</tr>
<tr>
<td>5</td>
<td>Thermometer</td>
<td>2425</td>
</tr>
</tbody>
</table>

Table 5-6 shows the total number of instances of each parameter value for all twelve subjects and the number of outlying instances excluded.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Total number of parameter from all 12 subjects</th>
<th>Total number of parameter removed due to outliers from all 12 subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spirometer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Forced spirometry</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 FEV₁</td>
<td>2455</td>
<td>28</td>
</tr>
<tr>
<td>2 FVC</td>
<td>2455</td>
<td>28</td>
</tr>
<tr>
<td>3 PEF</td>
<td>2455</td>
<td>28</td>
</tr>
<tr>
<td>4 FEV₆</td>
<td>1898</td>
<td>24</td>
</tr>
<tr>
<td><strong>Relaxed spirometry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 BR</td>
<td>2285</td>
<td>27</td>
</tr>
<tr>
<td>6 IC</td>
<td>2287</td>
<td>27</td>
</tr>
<tr>
<td>7 VE</td>
<td>2282</td>
<td>27</td>
</tr>
<tr>
<td>8 VT</td>
<td>2287</td>
<td>27</td>
</tr>
<tr>
<td>9 Ti/Ttot</td>
<td>2287</td>
<td>27</td>
</tr>
<tr>
<td><strong>Pulse oximeter</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 SpO₂</td>
<td>2467</td>
<td>61</td>
</tr>
<tr>
<td>11 Heart rate*</td>
<td>1902</td>
<td>50</td>
</tr>
<tr>
<td><strong>Weight scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Body weight</td>
<td>2405</td>
<td>43</td>
</tr>
<tr>
<td><strong>Thermometer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Body temperature</td>
<td>2478</td>
<td>53</td>
</tr>
<tr>
<td><strong>Blood pressure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Systolic</td>
<td>2700</td>
<td>261</td>
</tr>
<tr>
<td>15 Diastolic</td>
<td>2700</td>
<td>261</td>
</tr>
</tbody>
</table>
5.4.1.3 Database II Description

The second database was obtained from 18 COPD patients, aged between 63-87 years residing in the Western Sydney region of Australia. Data were collected by staff at the Blacktown Hospital, Sydney, New South Wales, Australia, from April 2009 to October 2010. Table 5-7 shows the subject’s baseline data.

The database was a part of a randomised controlled trial comparing enhanced care provided by a specialist care team called RACS against enhanced care integrated with remote monitoring. The RACS aims to support patients with COPD living in the area of Blacktown in the state of NSW. Patients are referred to the program by the hospital or other healthcare providers.

Table 5-7: Database II subjects baseline data.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Gender</th>
<th>Disease severity¹</th>
<th>FEV₁ (%)¹</th>
<th>Days enrolled in the trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66</td>
<td>Female</td>
<td>Severe</td>
<td>32</td>
<td>194</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>Female</td>
<td>Very severe</td>
<td>24</td>
<td>210</td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>Male</td>
<td>Very severe</td>
<td>26</td>
<td>228</td>
</tr>
<tr>
<td>4</td>
<td>87</td>
<td>Male</td>
<td>Very severe</td>
<td>20</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>63</td>
<td>Female</td>
<td>Severe</td>
<td>38</td>
<td>509</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>Female</td>
<td>Very severe</td>
<td>19</td>
<td>558</td>
</tr>
<tr>
<td>7</td>
<td>77</td>
<td>Female</td>
<td>Mild</td>
<td>84</td>
<td>450</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
<td>Female</td>
<td>Severe</td>
<td>32</td>
<td>568</td>
</tr>
<tr>
<td>9</td>
<td>74</td>
<td>Female</td>
<td>Very severe</td>
<td>25</td>
<td>448</td>
</tr>
<tr>
<td>10</td>
<td>76</td>
<td>Female</td>
<td>Severe</td>
<td>31</td>
<td>454</td>
</tr>
<tr>
<td>11</td>
<td>73</td>
<td>Male</td>
<td>Very severe</td>
<td>18</td>
<td>383</td>
</tr>
<tr>
<td>12</td>
<td>63</td>
<td>Female</td>
<td>Very severe</td>
<td>29</td>
<td>197</td>
</tr>
<tr>
<td>13</td>
<td>71</td>
<td>Male</td>
<td>Moderate</td>
<td>66</td>
<td>174</td>
</tr>
<tr>
<td>14</td>
<td>66</td>
<td>Male</td>
<td>Very severe</td>
<td>24</td>
<td>568</td>
</tr>
<tr>
<td>15</td>
<td>84</td>
<td>Female</td>
<td>Severe</td>
<td>41</td>
<td>80</td>
</tr>
<tr>
<td>16</td>
<td>80</td>
<td>Male</td>
<td>Very severe</td>
<td>24</td>
<td>68</td>
</tr>
<tr>
<td>17</td>
<td>71</td>
<td>Male</td>
<td>Severe</td>
<td>30</td>
<td>440</td>
</tr>
<tr>
<td>18</td>
<td>83</td>
<td>Male</td>
<td>Very severe</td>
<td>28</td>
<td>26</td>
</tr>
</tbody>
</table>

¹Staging criteria based on the GOLD guidelines (Rodriguez-Roisin, 2009)

The control arm of the study continued to receive enhanced care as per the usual RACS protocol, while an intervention group received home remote monitoring in addition to that. Each intervention subject was supplied with a TMC-Home system to be used in their home; training in operating the system was provided to patients, while clinicians underwent technical training for reviewing the data on a web interface. Throughout the data collection period, patients were scheduled to perform their measurements and answer questionnaires every day. A remote central database collated data received from patient homes and visualised the longitudinal
records in graphs that were available to the RACS team for daily viewing. The analysis presented herein focuses only on patients who were treated as part of this intervention arm; meaning those who were provided home monitoring data via the TMC-Home on a regular basis. The RACS team occasionally called or visited these patients, whether as part of the normal RACS routine or if the home monitoring indicated a need to do so. Similarly, the patients or the patients’ carers could also consult the RACS team whenever the need arose. During these consultations, the RACS team queried the severity of the patients’ symptoms, presence of co-morbid conditions, risk factors, availability of carer(s), and the ability of the patient or carer to implement the action plan at home. Based on the patient’s responses to these questions, the RACS team provided (Roberts et al., 2008b, Shany et al., 2010):

(a) a provisional diagnosis on the patient’s condition, which was that either
   (i) the patient was having an exacerbation, or
   (ii) the patient was in a stable condition

(b) a recommendation on the course of action which should be pursued, which was that either
   (i) the patient should attempt to continue managing their condition at home, or
   (ii) the patient should be referred for further clinical assessment.

5.4.1.4 Database II Data Analysis

All 18 intervention patients in Database II had enrolled in the trial for a total of 5591 days. Table 5-8 shows the total number of days on which the patients used each measurement device.
Table 5-8: Total number of days each measurement device was used by 18 patients.

<table>
<thead>
<tr>
<th>No</th>
<th>Measurement device</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spirometer</td>
<td>809</td>
</tr>
<tr>
<td>2</td>
<td>Pulse oximeter</td>
<td>4259</td>
</tr>
<tr>
<td>3</td>
<td>Blood pressure meter</td>
<td>4028</td>
</tr>
<tr>
<td>4</td>
<td>Weight scale</td>
<td>2827</td>
</tr>
<tr>
<td>5</td>
<td>Thermometer</td>
<td>3925</td>
</tr>
</tbody>
</table>

Table 5-9 shows the total number of instances of each parameter value for all 18 subjects and the number of outlying instances excluded.

Table 5-9: Total number of available and removed parameter values from all 18 subjects.

<table>
<thead>
<tr>
<th>Parameters*</th>
<th>Total number of parameter from all 18 subjects</th>
<th>Total number of parameter removed due to outliers from all 18 subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximeter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 SpO₂</td>
<td>5780</td>
<td>85</td>
</tr>
<tr>
<td>2 Heart rate*</td>
<td>5780</td>
<td>85</td>
</tr>
<tr>
<td>Weight scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Body weight</td>
<td>3392</td>
<td>27</td>
</tr>
<tr>
<td>Thermometer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Body temperature</td>
<td>5001</td>
<td>31</td>
</tr>
<tr>
<td>Blood pressure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Systolic</td>
<td>5911</td>
<td>469</td>
</tr>
<tr>
<td>6 Diastolic</td>
<td>5911</td>
<td>469</td>
</tr>
</tbody>
</table>

* Due to limited number of instances, the spirometric parameters were not pre-processed

5.4.2 Discussion

Figure 5-6 and Figure 5-7 show the gender and severity distribution for patients in Database I and Database II. The numbers of female patients enrolled in the intervention arm of the trials in both databases are higher than the numbers of male patients. There are four and ten very severe patients in Database I and Database II, respectively. Referring to Table 5-4, three out of four very severe patients in Database I are females and Table 5-7 shows that five out of ten very severe patients in Database II are males.
Combined, the twelve patients in Database I used the home monitoring measurement unit for a total of 2439 days. The blood pressure module is the most used device and the pulse oximeter and weight modules the least used devices (2406 days each). Nevertheless, all measurement devices had been used by the patients for more than 70% of the total days they were enrolled in the trials.

The patients in Database II used pulse oximetry, blood pressure and thermometer devices for more than 3900 days (70.0%) during enrolment in the trial. However, the spirometry device was used for only 809 days (14.5%) during the trial. Therefore, only six out of the 15 measurement parameters (see Table 5-1) were used for the analysis using Database II.

Spirometric measurements were generally performed under direct guidance of trained medical personnel (Finkelstein et al., 2000), thus, the RACS team during the data collection for
Database II did not encourage patients to perform spirometry in an unsupervised environment. The available spirometry measurements most probably had been recorded under the ambulatory care team supervision during the visits to the patient’s home.

Even though the total number of days with data for all twelve patients in Database I was 2439, patients had performed some of the measurements more than once in a day. For example, based on Table 5-6 all twelve patients had performed a temperature measurement 2478 times. Nevertheless, parameters from 53 temperature measurements were deemed to be outliers.

Figure 5-8 and Table 5-10 show the percentages of the removed parameters due to outliers. The percentage of parameters removed is higher in Database I compared to Database II. For example, 2.47% of SpO\textsubscript{2} parameter values were removed from Database I as compared to 1.47% removed from Database II. In both databases, the highest percentage of outliers removed were blood pressure derived parameters (systolic and diastolic).

![Figure 5-8: The percentage of parameters values removed for each parameter type for Database I and Database II.](image-url)
Table 5-10: The percentage of parameters values removed for each parameter type for Database I and Database II.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Percentage of number of parameter values removed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Database I</td>
</tr>
<tr>
<td>FEV₁</td>
<td>1.14</td>
</tr>
<tr>
<td>FVC</td>
<td>1.14</td>
</tr>
<tr>
<td>PEF</td>
<td>1.14</td>
</tr>
<tr>
<td>FEV₆</td>
<td>1.26</td>
</tr>
<tr>
<td>BR</td>
<td>1.18</td>
</tr>
<tr>
<td>IC</td>
<td>1.18</td>
</tr>
<tr>
<td>VE</td>
<td>1.18</td>
</tr>
<tr>
<td>VT</td>
<td>1.18</td>
</tr>
<tr>
<td>Ti/Ttot</td>
<td>1.18</td>
</tr>
<tr>
<td>SpO₂</td>
<td>2.47</td>
</tr>
<tr>
<td>Heart rate</td>
<td>2.63</td>
</tr>
<tr>
<td>Body weight</td>
<td>1.79</td>
</tr>
<tr>
<td>Body temperature</td>
<td>2.14</td>
</tr>
<tr>
<td>Systolic</td>
<td>9.67</td>
</tr>
<tr>
<td>Diastolic</td>
<td>9.67</td>
</tr>
</tbody>
</table>

The patients with COPD in both databases were only recruited from northeast Melbourne, and West Sydney, Australia, and hence may not be representative of the general COPD population. The initial number of patients agreeing to perform daily home telehealth measurement in Database I was 22, of which two died, four withdrew and another four did not perform their physiological measurement recordings. Therefore, only data from 12 patients were available from Database I for analysis. Meanwhile, the original number of patients in Database II was 21, with two patients opting to withdraw and one patient dying, thus only 18 patients’ data were available.

The issue regarding small sample sizes in respiratory related telehealth research does not only occur in this study; there are several home telehealth studies that also report small numbers of recruited patients with COPD (Vontetsianos et al., 2005), due to the difficultly of executing such trials. Vitacca et al. mentioned that the number of patients interested in a respiratory home telehealth study might increase if the study was conducted for a longer period of time. Such trends are likely due to the increasing technological awareness of patients with respiratory disease (Vitacca et al., 2010).
A limitation of Database I was that the diagnosis provided by the ambulatory care nurse was not formally recorded. On the other hand, the data collection for Database II included a formal record of expert opinions and recommendations from the RACS team on the occurrence of exacerbation events and the action taken whenever COPD exacerbation occurred.

Link to chapter six and seven

This chapter presents the methods and the materials that are used in chapters six and seven. Each chapter used different reference standards. Chapter six discusses the construction of algorithms to detect patients at high risk of exacerbations and chapter seven describes the developments of the rules to provide referral recommendations for high-risk patients.
Chapter 6 : Predicting the Risk of Exacerbation in Patients with COPD using Home Telehealth Data

6.1 Introduction

This chapter outlines the development of an algorithm to identify patients with COPD who appear to be at a high risk of suffering an exacerbation event. The ability of home telehealth physiological measurement data to classify patients into low and high exacerbation risk groups is assessed using a CART technique. The data employed in the validation of this algorithm were collected from retrospective home telehealth studies as described in previous chapters.

The physiological measurement data are pre-processed to remove possible outliers, then features are extracted from the measurement data parameters. The predictive power of features from each measurement type and all features are evaluated using multiple feature classification.

6.2 Background

An important feature of the forced spirometry manoeuvre is the FEV\textsubscript{1} value. This parameter has been used to categorise confirmed COPD patient severities into different disease stages (Mannino et al., 2003, Probst-Hensch et al., 2010). Patients with a higher degree of disease severity have a higher risk of suffering an exacerbation episode (Faganello et al., 2010). Moreover, a decline in the FEV\textsubscript{1} value, in patients with moderate COPD, below 1.0 L or less than 40% of the predicted value, is considered an indicator of a severe exacerbation (McKenzie et al., 2009).

Moreover, a worsening of expiratory flow occurs during onset of exacerbation, resulting in dynamic hyperinflation (very high lung inflation). Consequentially, the IC value is reduced and VE, VT and BR values increase (O’Donnell and Parker, 2006, O'Donnell and Laveneziana, 2006, Yetkin and Gunen, 2008).

Hurst et al. showed that SpO\textsubscript{2} and heart rate facilitated the identification of the onset of COPD exacerbation (Hurst et al., 2010). Furthermore, SpO\textsubscript{2} is useful in screening insufficient blood oxygenation saturation in patients with COPD (Guryay et al., 2007), and a value of SpO\textsubscript{2} less than 88% is considered a sign of instability in a COPD patient’s condition; the level indicating that the patients is in a minimal oxygenation reserve condition (McKenzie et al., 2009).
Likewise, heart rate is reported to increase as COPD exacerbations worsen, with a heart rate value of more than 110 bpm or 20% above baseline being one of the clinical assessment guidelines used in determining exacerbation severity (Rodriguez-Roisin, 2009, Franciosi et al., 2006, Postma et al., 1999, Calverley, 2005). The heart rate value can also be obtained from ECG measurements recorded by the TMC-Home; however, in this analysis we only used heart rate value derived from pulse oximetry signals.

Another physical sign of COPD exacerbation is hypotension, a term for low blood pressure (Schumaker and Epstein, 2004). The significance of low blood pressure with COPD exacerbation in-hospital mortality had been reported by Edwards et al. (Edwards et al., 2011). Since low blood pressure is identified as a systolic blood pressure less than 90 mmHg or a diastolic pressure less than 60 mmHg (Chang et al., 2011), both parameters were included in the predictive model development.

Temperature is a useful parameter when monitoring a chronic disease patient’s daily wellbeing. One particular study has shown that more than one-third of exacerbations requiring hospitalisation had fever as an indicator (Lieberman et al., 2003). Patients performed their body temperature measurement (oral or axillary) manually using the provided thermometers and entered the value using the TMC-Home console.

Similarly, in patients with advanced COPD, a decrease in weight regularly occurs as a consequence of increased resting energy expenditure due to exacerbations (Barnett, 2006). However, obese (excess body weight) COPD patients could have respiratory complications such as increased respiratory effort, respiratory muscle atrophy and diminished respiratory compliance. As a result, COPD patients with obesity commonly have hypoxaemia and pulmonary hypertension, which are risk factors for exacerbations (Cote et al., 2007, Chaouat et al., 2008).
6.3 Methodology

6.3.1 Exacerbation Risk Health Status Reference Standard

6.3.1.1 Database I

In order to construct a classification model, to identify the prodromal stage of an exacerbation event, a reference standard containing each subject’s health status on days at risk of the study was constructed. This reference standard was used in the training phase of the classifier design to fine-tune the performance of the classifier, before independent validation was performed. The process of generating the reference standard is described below.

The reference standard development was adapted from the definition of COPD exacerbation by the GOLD guidelines (Rodriguez-Roisin, 2009, Rodriguez-Roisin, 2000). In the symptom questionnaires, patients were asked if there were changes in their symptoms, namely: increase in sputum amount; and worsening in breathing condition or if patients were experiencing cold or influenza. Meanwhile, in the daily medication questionnaires, patients were asked if they had increased their dosage, or had started the use of a respiratory medication on that day.

Worsening in any one of the symptoms that warranted an increased dosage, or initiating the usage of a respiratory medication, indicated that the patient may have undergone an exacerbation episode on that particular day (McKenzie et al., 2009, Rodriguez-Roisin, 2009). Henceforth, one of two categorisation labels was assigned to the patient’s health condition on each day: high risk of developing exacerbation, or low risk. The ‘high risk’ category was assigned if they confirmed a worsening of any one of the symptoms concomitant with an increase in medication usage. Otherwise the patient was deemed to be ‘low risk’. This health status reference standard was later used in the algorithm development and validation processes.
A gold standard is defined as an authoritative diagnosis of a particular patient’s condition (Petrie and Sabin, 2009). Database II has a gold standard reference for each subject’s health status, obtained from the diagnosis of the RACS nurse. The collected data in Database II were sent automatically to a remote database, where a RACS nurse accessed the data via a website. The RACS nurse called or visited the patients when necessary, and/or the patients or the patients’ carers could consult the RACS nurse whenever they felt the need. Whenever the RACS nurse or a patient or carer called for a consultation, the nurse filled in an evaluation form. The form’s major purpose was to assess the patient’s condition in terms of ability to cope at home, the patient’s symptoms, level of activity the patients could perform, and the type of medications or remedial actions taken. Based on this information, the RACS nurse provided a provisional diagnosis on the patient’s condition. The types of diagnosis were:

a) Stable condition

b) AECOPD (infective) – if the patient exhibited signs and symptoms suggestive of an infective AECOPD, such as increased amount or change in the colour of sputum outside the patient’s normal range

c) AECOPD (non infective) – if the patient exhibited worsening symptoms suggestive of an AECOPD, such as increased cough and breathlessness outside the patients normal range, but nil change in sputum

Using the RACS nurse’s provisional diagnosis, one of two categorisation labels was assigned to the patient’s health condition on each day: ‘high risk’, or ‘low risk’. The ‘high risk’ category was assigned if the diagnosis was either (b) or (c); else, the patient’s health condition was categorised as ‘low risk’.
6.4 Results

6.4.1 Exacerbation Risk Health Status Reference Standard

6.4.1.1 Database I

Only 680 out of 3249 days from all the 12 patients with COPD data contained symptoms and concurrent respiratory medication records. Of these 680 records, approximately 36 days did not have any physiological measurement data at all, and 23 days of measurement data were deemed to contain outliers and were thus excluded from the analysis, leaving a total of 621 useable training instances. Of these 621 days, 314 days were labelled as ‘high risk’ and the remaining 307 days were categorised as ‘low risk’ based on the reference standard.

Figure 6-1: Distribution of 621 days containing both respiratory medication records and measurement data, for the 12 patients in Database I.

The COPD risk rating derived from the medications record is also indicated for each subject using light and dark shading.

Figure 6-1 shows the data distribution among the 12 patients. The days that contained patient measurement data, symptoms, and medication records, were distributed unevenly, with patient 2 recording the highest number of days with data (193 days) and patients 9 and 11 the fewest number of days (9 days).
6.4.1.2 Database II

Consultation between the patients and the RACS team occurred for 801 days within the trial. However, home telehealth data were available for analysis for 348 out of 801 days. The 348 days came from 16 patients records. Patients were diagnosed as stable (low risk) for 245 days and labelled as high risk for 103 days (80 days were diagnosed as AECOPD infective and 23 days were diagnosed as AECOPD non-infective). Figure 6-2 shows the distribution of the low and high risk cases for each of the 16 patients.

![Figure 6-2: Distribution of 348 consultation days containing both RACS diagnosis and measurement data, for the 16 patients in Database II.](image)

6.4.2 Data Analysis

6.4.2.1 Exacerbation Classification Performance for Database I

The multivariate classification performances, estimated using cross validation, are presented in the following tables. Table 6-1 shows the $\kappa$ value of the CART classifiers constructed using all features for each measurement type, and Table 6-2 shows the confusion matrix of the CART trained and tested using all features from all parameters. The rules obtained from the cross validation training-testing processes are presented in Appendix A.
Table 6-1: Multiple feature classification using all features from each measurement type with $\kappa$ value estimated using cross validation.

<table>
<thead>
<tr>
<th>Measurement technique</th>
<th>Parameters</th>
<th>Number of extracted features</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximetry</td>
<td>SpO$_2$, Heart rate</td>
<td>10</td>
<td>0.51</td>
</tr>
<tr>
<td>Forced spirometry</td>
<td>FEV$<em>1$, FVC, PEF, FEV$</em>{5}$</td>
<td>20</td>
<td>0.54</td>
</tr>
<tr>
<td>Relaxed spirometry</td>
<td>BR, IC, VE, VT, Ti/Ttot</td>
<td>25</td>
<td>0.20</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Systolic, Diastolic</td>
<td>10</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight</td>
<td>Weight</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Temperature measurement</td>
<td>Body temperature</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6-2: Confusion matrix for CART classifiers estimated using cross validation using data from all 12 patients from Database I.

<table>
<thead>
<tr>
<th></th>
<th>Reference standard</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>CART</td>
<td>222 (TP)</td>
<td>37 (FP)</td>
</tr>
<tr>
<td></td>
<td>92 (FN)</td>
<td>270 (TN)</td>
</tr>
<tr>
<td>Total</td>
<td>314</td>
<td>307</td>
</tr>
</tbody>
</table>

Accuracy = 79.00%, Specificity = 87.90%, Sensitivity = 70.70%, PPV = 85.71%, NPV = 74.59%, $\kappa = 0.59$

---

2 The type of features have been discussed in Chapter five
6.4.2.2 Agreement between Exacerbation Risk Health Status Reference Standard I from Database I and Reference Standard II from Database II

The availability of symptoms and medication questionnaires in Database II allowed us to assess the agreement between the two exacerbation risk reference standards. However, the comparison and the calculation of kappa statistics could only be performed on 311 cases. This is because reference standard I could not be generated for 27 out of 348 days for which reference standard II was calculable; due to the absence of symptoms or medications questionnaires records. Table 6-3 shows the confusion matrix.

<table>
<thead>
<tr>
<th>Reference standard I</th>
<th>High risk</th>
<th>Low risk</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference standard II</td>
<td>High risk</td>
<td>72</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>24</td>
<td>196</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
<td>215</td>
<td>311</td>
</tr>
</tbody>
</table>

\[\kappa = 0.67\]

Reference standard I versus reference standard II: Accuracy = 86.17\%, Specificity = 89.09\%, Sensitivity = 79.12\%

Reference standard II versus reference standard I: Accuracy = 86.17\%, Specificity = 91.16\%, Sensitivity = 75.00\%

6.4.2.3 Exacerbation classification performance for Database II

Table 6-4 shows the \(\kappa\) value of the CART classifiers constructed using all features for each measurement type. The results also include the performance of the CART that used only questionnaires answers. Table 6-5 shows the confusion matrix of the CART trained and tested using all features from all parameters and questionnaires answers and Appendix B shows the rules.
Table 6-4: Multiple feature classification using all features from each measurement type with \( \kappa \) value estimated using cross validation.

<table>
<thead>
<tr>
<th>Measurement technique</th>
<th>Parameters/questions</th>
<th>Number of features(^3)</th>
<th>( \kappa )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximetry</td>
<td>( \text{SpO}_2 ), Heart rate</td>
<td>10</td>
<td>0.14</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Systolic, Diastolic</td>
<td>10</td>
<td>0.14</td>
</tr>
<tr>
<td>Weight</td>
<td>Weight</td>
<td>5</td>
<td>0.09</td>
</tr>
<tr>
<td>Temperature measurement</td>
<td>Body temperature</td>
<td>5</td>
<td>0.09</td>
</tr>
<tr>
<td>Questionnaires</td>
<td>Breathing, general feeling, sputum amount, colour and consistency, cough and night time symptoms</td>
<td>7</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 6-5: Confusion matrix for CART developed using data from Database II.

<table>
<thead>
<tr>
<th>Reference standard</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High risk</td>
</tr>
<tr>
<td>CART</td>
<td></td>
</tr>
<tr>
<td>High risk</td>
<td>AECOPD (infective)</td>
</tr>
<tr>
<td>36 (TP)</td>
<td>7 (TP)</td>
</tr>
<tr>
<td>Low risk</td>
<td>44 (FN)</td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
</tr>
</tbody>
</table>

Accuracy=76.72%, specificity=91.43%, sensitivity=41.74%, PPV=67.19%, NPV =78.87%, \( \kappa =0.37 \)

6.5 Discussion

Referring to Table 6-1 the two most useful measurement types that gave \( \kappa \) values of more than 0.5 (moderate agreement with the reference standard) are the pulse oximetry and the forced spirometry measurements.

Examining the results in Table 6-2, from 61 days, 307 days were labelled as ‘low risk’ and the remaining 314 days were labelled as ‘high risk’. This classifier identified 222 true positive cases out of a possible 314 ‘high risk’ cases, but misclassified 9 cases as ‘low risk’, when in fact they were considered ‘high risk’ by the reference standard. Of 307 cases categorised as ‘low risk’ by the CARTs classifier, 270 were correctly recognised.

\(^3\) The type of features have been discussed in Chapter five

106
Due to the limited number of spirometry measurements in Database II, the parameters were not included in the analysis. Table 6-4 shows that questionnaire answers alone demonstrate a fair agreement with the RACS team COPD exacerbation diagnosis. This finding is in alignment with the study conducted by Trappenburg et al., stating that clinical COPD questionnaires can be used as a tool to detect exacerbation events (Calverley, 2005). In addition, the CART classifiers constructed using the features from pulse oximetry and the blood pressure parameters demonstrate a slight agreement with the reference standard.

The multivariate CART classifiers constructed using all features from all 16 subjects gives an accuracy of 76.7%, and \( \kappa \) value of 0.37. The performance of the CART is lower compared to the performance of the CART constructed using data from Database I. Possible reasons that may have affected the CART performance is the absence of features from spirometry measurements in Database II.

The exacerbation reference standard for Database I was generated using the definition provided in the GOLD guidelines. While the reference standard for the analysis performed using data from Database II employed the diagnosis from the RACS team. Even though the agreement between the two reference standards is categorised as substantial, with a \( \kappa \) value of 0.67 (Cohen, 1960, Viera and Garrett, 2005), in general it is not a trivial exercise to define the risk of exacerbation for COPD patients using their self-reported symptoms and questionnaires, this is because; 1) self data recording could introduce inconsistencies, as each individual patient interprets his/her symptom severity differently (Calverley, 2005); and 2) the subjects were elderly people and concerns were raised that the tendency to misreport information about their health was high (Bush et al., 1989, Tisnado et al., 2006). Thus, Database II CARTs were built only using the reference standard II definition that is based on clinical expert diagnosis.

Marin et al. used a COX regression model and the BODE index to predict the risk of exacerbation (Marin et al., 2009). However, the technique does not generate decision rules, which are important at the point of care in a home telehealth scenario; a decision tree classifier makes the HM-DSS decision easy for a health practitioner to interpret (Ayer et al., 2010, Núñez et al., 2002). An example of a decision tree, constructed during one of the cross validation runs, using reference standard I and Database I data, is presented in Figure 6-3.

In addition, Miravitlles et al. used logistic regression to estimate the probabilities of frequent exacerbations among COPD patients. The analysis suggested that FEV<sub>1</sub> impairment explains the risk of frequent exacerbations. Nevertheless, the study was conducted in a general practice
environment and does not include patients who may not be able to perform forced spirometry measurements (Miravitlles et al., 2000). The CART models proposed here minimise this limitation by showing that other measurements, such as pulse oximetry, could also further assist the prediction of exacerbations (Table 6-1 and Table 6-4).

Recently, Jensen et al. used linear regression in predicting COPD exacerbations. The analysis was performed using data collected from telecare patients enrolled in a tele-rehabilitation program. The algorithm classifies exacerbation and non-exacerbation cases, with seven out of ten exacerbation cases correctly identified (Jensen et al., 2012). Table 6-6 shows the comparison between the approach taken by (Jensen et al., 2012) with the approach described in this chapter.

**Table 6-6: Comparisons between the approach taken by (Jensen et al., 2012) and the approach described in chapter six.**

<table>
<thead>
<tr>
<th></th>
<th>Jensen et al.</th>
<th>Chapter six</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference standard</td>
<td>Determined from a combination of hospital admissions and patient’s medication record and reported symptoms</td>
<td>Based on: I) clinical guidelines (patient’s symptoms and medication records) II) provisional diagnosis provided by the RACS team</td>
</tr>
<tr>
<td>Home monitoring device</td>
<td>1) Digital body weight scale (UC-321 pbt, A&amp;D Medical)</td>
<td>Integrated TMC-Home system (integrates all types of vital signs measurements)</td>
</tr>
<tr>
<td></td>
<td>2) Blood pressure meter (UA-767 plus BT, A&amp;D Medical)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3) Pulse oximeter (Onyx II, Nonin)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4) Spirometer (Lung Monitor S, Vitalograph)</td>
<td></td>
</tr>
</tbody>
</table>

**Link to next chapter**

Not all exacerbation events need hospital admission or specialist intervention. Some ‘high risk’ cases could be managed from the patient’s home. Thus, the next chapter will describe the development of an algorithm for supporting the care team with patient telehealth management by generating referral recommendations for individuals having a high risk of exacerbation, to prioritise cases that potentially require hospital or specialist intervention.
Figure 6-3: Illustration of a CART constructed using data from 11 subjects (one of the 12 cross validation run) in Database I.

The circle shape symbolizes a leaf node which can represent either a ‘low risk’ (white circle) or ‘high risk’ (black circle) health condition (L=number of low risk cases, H=number of high risk cases). The alphabetic labelling of each decision node represents the corresponding feature used and the decision threshold. The splits go to the left if less than the threshold value.

A=  SpO$_2$ distribution mean (%), B= IC distribution mean (litre), C= Heart rate standard deviation (bpm), D= Weight percentage changes (%), E= Temperature distribution mean (0C), F= VT standard deviation (litre), G= FEV$_6$ standard deviation (litre), H= Heart rate percentage changes (%), I= Weight data value (kg), J= FEV$_6$ percentage changes (%), K= Ti/Ttot distribution mean (unitless), L= BR distribution mean (breaths per minute), M= Diastolic distribution mean (mmHg), N= FVC standard deviation (litre), O= VE standard deviation (litre), P= Heart rate data value (bpm), Q= FEV1 distribution mean (litre), R= SpO$_2$ standard deviation (%), S= FEV$_1$ standard deviation (litre), T= Systolic zScore (unitless), U= SpO$_2$ data value (%), V= Heart rate distribution mean (bpm), W= Temperature zScore (unitless)
Chapter 7: Generating Clinical Assessment Referral Recommendations from Home Telehealth Data

7.1 Introduction

The previous chapter described the construction of a set of rules, which used home telehealth data to identify COPD patients at high risk of imminent exacerbation. This chapter extends this analysis to further stratify those high risk cases into two groups; one group that needed to be referred for clinical assessment, and a second group that were safe to continue to self-manage their condition at home.

The methods and materials used in this chapter have been presented in chapter five. A health status reference standard was constructed, against which the CART was trained and tested. As in chapter six, two databases were used, the Austin database (Database I) and the Blacktown database (Database II).

7.2 Background

Remote monitoring integrated with a standard care scheme is an alternative to repeat hospitalisation and involves the support from an ambulatory care team or a RACS team providing care at home for early discharged patients. Enrolled patients are initially referred into the program by the hospital or by the community nurse.

When using a home telehealth system, ambulatory care teams have to monitor the patients’ physiological measurement data (self-measured by patients) and questionnaire responses on a regular basis. While this additional task can keep track of a patient’s condition continuously, there is much larger volume of data to be monitored, consequently increasing the clinical carer’s workload (Darkins and Cary, 2000).

Thus, to assist the ambulatory care team in monitoring patients, it is proposed herein that a knowledge base of patient measurement data and a corresponding set of labelled patient conditions be compiled.
7.2 Methodology

7.2.1 Referral Status Reference Standard

7.2.1.1 Database I

Two types of indicators are employed to generate a health status reference standard, against which the CART is later trained and compared. The primary indicator leveraged was the medications and symptoms questionnaire data, entered by the patient; specifically, the symptoms and medications questionnaires related to the use of respiratory therapy. A worsening of symptoms resulting in an increase in the use of respiratory medications is considered a predictor of worsening health status among COPD sufferers. It is then established whether these patients were managed at home, or had been referred to a GP or respiratory specialist, or had been visited by an outreach nurse, or had been admitted to hospital soon after this increase in their medications. This is done by checking the carer journal entry (c.f. Chapter five, section 5.2.1.1) for any record of a visit to a clinic, or for a hospital admission, or an unscheduled visit by a nurse on that day. Henceforth, one of two classification labels is assigned to the patient’s health status for that day:

1) **Home management:** If a worsening of any one of the symptoms concomitant with an increase in respiratory medication usage was observed, but the patient was not visited by a nurse or was not referred to the GP or admitted to hospital soon afterward.

2) **Referral:** If, again, an increase in medications was observed, but this time the patient was visited by a nurse or was referred to the GP or admitted to hospital soon afterward.

If there is no entry in the carer journal on the date which the increase in medications occurred, the date is excluded from the analysis.
7.2.1.2 Database II

During the data collection period, the RACS team occasionally called or visited these patients and similarly, the patients or the patients’ carers could also consult the RACS team whenever the need arose. During these consultations, the RACS team provides five types of recommendations on the course of action that should be pursued, which is that either:

a) Reassurance and advice and patient should continue with home management
b) Reassurance and advice, patient should continue with home management and the RACS team will perform a follow-up call
c) Reassurance and advice, patient should continue with home management and the RACS team will perform a follow-up visit
d) Patient’s condition should be reviewed by GP
e) Patients should be sent to the hospital

Using the RACS nurse’s recommendations, one of two categorisation labels was assigned to the patient’s health condition on that day: ‘referral’, or ‘home management’. If the patient for that day fell into the (d) or (e) categories he/she was assigned to the ‘referral’ state.
7.3 Results

7.3.1 Referral Status Reference Standard

7.3.1.1 Database I

There were 371 days with data from all the 12 COPD patients that were used in this analysis. Of these 371 records, 82 days were labelled as referred cases and the remaining 289 days were categorised as home management cases based on the reference standard.

Figure 7-1 shows the distribution of the 82 referred cases. Sixty-five percent (53 cases) were referred to a GP or respiratory specialist, 19 (23%) cases were labelled as ‘referral’ due to visits by an outreach nurse to the patients home and the remaining 10 (12%) cases were hospital admissions or visits and emergency department (ED) presentations cases.

![Figure 7-1: Distribution of the 82 referred cases.](image-url)
7.3.1.2 Database II

The analysis was performed on the 103 cases that had been diagnosed as COPD exacerbation cases by the RACS team during the consultation. The RACS team had given a home management recommendation to 12 patients involved in 41 different consultations. The remaining 6 cases comprising 14 patients were labelled as ‘referral’ cases. Figure 7-2 shows the distribution of the types of home management and referred cases.

![Figure 7-2: Distribution of the home management and the referral cases.](image)

A = Home management cases, B = Home management cases with follow-up call or visit planned, C = GP review cases, D = Hospital cases
7.3.2 Data Analysis

7.3.2.1 Database I

The CART performances, estimated using cross validation, are presented in Table 7-1 and Table 7-2 and the 12 sets of rules are presented in Appendix C.

Table 7-1: $\kappa$ value for classification using data from each measurement technique, estimated using cross validation.

<table>
<thead>
<tr>
<th>Measurement technique</th>
<th>Parameters</th>
<th>Number of extracted features $^4$</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximetry</td>
<td>$\text{SpO}_2$, Heart rate</td>
<td>10</td>
<td>0.43</td>
</tr>
<tr>
<td>Forced spirometry</td>
<td>$\text{FEV}_1$, $\text{FVC}$, $\text{PEF}$, $\text{FEV}_6$</td>
<td>20</td>
<td>0.19</td>
</tr>
<tr>
<td>Relaxed spirometry</td>
<td>$\text{BR}$, $\text{IC}$, $\text{VE}$, $\text{VT}$, Ti/Ttot</td>
<td>25</td>
<td>0.48</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Systolic, Diastolic</td>
<td>10</td>
<td>0.14</td>
</tr>
<tr>
<td>Weight measurement</td>
<td>Weight</td>
<td>5</td>
<td>0.00</td>
</tr>
<tr>
<td>Temperature measurement</td>
<td>Body temperature</td>
<td>5</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

Table 7-2: Confusion matrix for HM-DSS developed using data from Database I.

<table>
<thead>
<tr>
<th>Reference standard</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Referral</td>
</tr>
<tr>
<td><strong>HM-DSS</strong></td>
<td></td>
</tr>
<tr>
<td>Hospital admission, visit or ED presentation</td>
<td>48 (FP)</td>
</tr>
<tr>
<td>Visit to GP or Respiratory specialist clinic</td>
<td>35 (TP)</td>
</tr>
<tr>
<td>Visited by outreach nurse</td>
<td>18 (TP)</td>
</tr>
<tr>
<td>Referral</td>
<td></td>
</tr>
<tr>
<td>8 (TP)</td>
<td></td>
</tr>
<tr>
<td>Home management</td>
<td></td>
</tr>
<tr>
<td>2 (FN)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
</tr>
</tbody>
</table>

Accuracy=81.40%, specificity=83.39%, sensitivity=74.39%, PPV=55.96%, NPV = 91.98%, $\kappa = 0.52$

$^4$ The type of features have been discussed in Chapter four
7.3.2.2 Database II

The $\kappa$ value for a multivariate CART using features from each measurement is presented in Table 7-3. In addition, Table 7-4 shows the confusion matrix comparing the classification output by the CART in comparison to the reference standard. Both tables show aggregated results estimated using a cross validation technique. Appendix D shows the 14 sets of rules obtained.

**Table 7-3: $\kappa$ value for classification using features from each measurement technique, estimated using cross validation.**

<table>
<thead>
<tr>
<th>Measurement technique</th>
<th>Parameters</th>
<th>Number of extracted features $^5$</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximetry</td>
<td>$\text{SpO}_2$, Heart rate</td>
<td>10</td>
<td>0.29</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Systolic, Diastolic</td>
<td>10</td>
<td>0.34</td>
</tr>
<tr>
<td>Weight</td>
<td>Weight</td>
<td>5</td>
<td>0.12</td>
</tr>
<tr>
<td>Temperature</td>
<td>Body temperature</td>
<td>5</td>
<td>0.24</td>
</tr>
<tr>
<td>Questionnaires</td>
<td>Breathing, general feeling, sputum amount, colour and consistency, cough and night time symptoms</td>
<td>7</td>
<td>0.19</td>
</tr>
</tbody>
</table>

**Table 7-4: Confusion matrix for HM-DSS developed using data from Database II.**

<table>
<thead>
<tr>
<th>Reference standard</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Referral</td>
</tr>
<tr>
<td>HM-DSS</td>
<td>GP review</td>
</tr>
<tr>
<td></td>
<td>44 (TP)</td>
</tr>
<tr>
<td>Home management</td>
<td>15 (FN)</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
</tr>
</tbody>
</table>

Accuracy = 72.82%, specificity = 73.17%, sensitivity = 72.58%, PPV = 80.35%, NPV = 63.82%, $\kappa = 0.45$

$^5$ The type of features have been discussed in Chapter four
7.4 Discussion

The reference standard for Database I was generated based on the carer journal entry. On the other hand, the referral reference standard for Database II was produced based on the RACS team action plans, which were recommendations provided to the patients.

A visit by the outreach nurse to a patient’s home in Database I was categorised as a referral case. This is because visiting events that had been recorded in the Database I carer journal were unscheduled visits and occurred due to a patient emergency. While visit event in Database II were planned (scheduled) by the RACS team.

To identify the measurement technique that is most useful in assisting referral case detection, we have performed classification using features extracted from parameters derived from each measurement technique.

For analysis using Database I, the classification results in Table 7-1 show that relaxed spirometry is the most useful technique with a $\kappa$ value of 0.48 followed by pulse oximetry with a $\kappa$ value of 0.43. $\kappa$ values in Table 7-3, denoting classification using data from Database II shows that blood pressure ($\kappa=0.34$), pulse oximetry ($\kappa=0.29$) and temperature ($\kappa=0.24$) are important measures in recommending a referral to patients.

Referring to Table 7-2, the CART has achieved a classification accuracy of 81.40%, sensitivity of 74.39% and specificity of 83.39%. However, the CART has misclassified 21 cases as home management cases when they were actually labelled as referral cases. Eighteen misclassified referral cases were visits to a GP or respiratory specialist, two were hospital admissions cases and the remaining case was an outreach nurse visit. The CART correctly identified all ten ED presentation cases.

The CART for generating referral recommendations for Database II produced an accuracy of 72.82%. The CART misclassified 15 GP review cases and two cases recommended for hospitalisation. The CART correctly identified 45 out of 62 referral cases and 30 out of 41 home management cases.

In addition, the results in this chapter show that there is an alternative way for patients to perform their remote monitoring measurement. Namely that COPD patients may just record
pulse oximetry measurements alone. This option certainly increases for the convenience and usability of the system for COPD patients. This is because they can choose to perform only the easier measurement, such as pulse oximetry (Wiltshire et al., 2001), without greatly compromising the way a HM-DSS generates referral decisions (κ value of 0.43 for Database I and κ value of 0.29 for Database II CARTs; constructed using features from pulse oximetry derived parameters alone).

The reference standards used in generating the CARTs are based on information available in the informal carer journal (Database I) or recommendations provided by the RACS nurses. This information might be less comprehensive when compared to information available in the clinical free-text notes (Trafton et al., 2010). However, constructing a reference standard from the free-text notes is a complex task. This is because the contained information is often vague or partially articulate (Luger and Stubblefield, 1990).

The CART analyses have been performed using home telehealth data, self-measured by patients. However, the nature of unsupervised recording is also a significant challenge facing telehealth monitoring. Since motion artifact (caused by unidentified poor measurement technique in the home environment) was expected to be a confounding factor in the analysis (Piramuthu, 2008), outliers produced by motion artifact were manually excluded from the dataset; however, the existence of erroneous parameter values which were not identified as outliers could still have influenced the classification performances quoted here, as they would not have been flagged for inspection and removed if necessary, and manual inspection of the raw waveforms underlying all data points is impractical. While the validation of signal quality was performed manually in this instance, recent methods have been developed to automatically identify movement artifact in physiological signals, specifically those recorded in the telehealth environment (Redmond et al., 2008, Abdul Sukor et al., 2011). Chapter eight will discuss methods to overcome the challenge posed by motion artifact in these telehealth recordings.
Chapter 8 : Effect of Home Telehealth Data Quality on Decision Support System Performance

8.1 Introduction

In normal operation, a patient performs their vital sign measurements using the TMC-Home integrated measurement apparatus and parameter data are automatically derived from the raw signals.

In this chapter, the quality of the automatically derived measurement data are evaluated using available signal quality algorithms. The analyses are performed on raw signals from two types of measurements; the pulse oximetry and the blood pressure.

To demonstrate the effect of data quality on the HM-DSS, the performance of the HM-DSS is compared by using data that originally comes directly from the TMC-Home device and against data extracted from the raw signals after performing signal quality analysis. The results are presented and discussed.

8.2 Background

8.2.1 Pulse Oximetry Signal Quality Analysis

Pulse oximetry is a technique that monitors SpO₂ and heart rate. SpO₂ represents the estimates of the oxygen saturation (SaO₂) value; that is the ratio of oxygenated haemoglobin (HbO₂) to the combined amount of HbO₂ and deoxygenated haemoglobin (Hb) present in arterial blood.

The method of measuring SpO₂ using pulse oximetry is based on the principle of photoplethysmography (PPG). A conventional PPG probe uses two wavelengths (red and infrared, often at 660 nm and 940 nm, respectively) and each wavelength is preferentially absorbed by HbO₂ or Hb in the blood. This enables the derivation of an SpO₂ estimate using a general empirical linear approximation (Equation 8), where \( R \) is the ratio of red to infrared lights absorption (Rusch and Sankar, 1999).

\[
SpO_2 = 110 - 25R
\]
Figure 8-1 illustrates an example of a PPG signal, presented in the form of light absorption intensity that corresponds to the volume change in arterial blood in the finger.

![Pulse Oximetry](image)

**Figure 8-1: A typical PPG signal in the form of light absorption intensity that corresponds to arterial blood volume changes.**

Despite the considerable advantages obtained by employing pulse oximetry for SpO₂ estimation, there are some factors that may degrade the accuracy of the estimation. The introduction of ambient light at the photodetector, wearing nail polish, poor blood perfusion of the peripheral tissues, motion artifact (including the relative movement between the fingertip and the oximetry probe), are all well-known sources of error (Rusch et al., 1996, Lee et al., 2004).

A pulse oximetry signal quality algorithm has been developed by our research group to detect and eliminate noise in noise contaminated pulse oximetry signals. The noise detection method is based on waveform morphology analysis since there are morphological aspects of the PPG that could be used to differentiate between good and corrupted pulses. The output of the algorithm comprises sections of clean pulses free of various forms of artifact. This has the potential to help increase the accuracy of SpO₂ and heart rate derivation (Abdul Sukor et al., 2011).
8.2.2 Blood Pressure Signal Quality Analysis

Blood pressure usually refers to the arterial pressure of the systemic circulation. During each heartbeat, blood pressure varies between a maximum (systolic) and a minimum (diastolic) pressure; “systolic” and “diastolic” are two terms commonly used to summarise the blood pressure profile. Systolic blood pressure refers to the pressure in the systemic arteries as the heart expels blood during each beat, while diastolic blood pressure is the pressure as the heart relaxes (Marshall, 2004).

The TMC-Home NIBPM applies two types of methods; auscultatory and oscillometric methods. The auscultatory method is based on the measurement of Korotkoff sounds while the oscillometric method is based on the oscillation of pressure in the cuff during its gradual deflation from above the systolic to below the diastolic pressure.

Figure 8-2 shows the examples of signals from a NIBPM measurement. The upper graph is the cuff pressure signal during the gradual deflation stage, the middle and the lower graphs are the oscillometric and Korotkoff sound signals respectively.

![Figure 8-2: Examples of cuff pressure, oscillometric and Korotkoff sound signals.](image-url)
The measurement of blood pressure through NIBPM is highly vulnerable to artifact such as patient movement, talking and coughing that subsequently leads to a degradation in the accuracy of the measurement. When the measurement is contaminated with motion artifact, the air flow in the deflating cuff is typically interrupted, causing corruption of the small pressure oscillations from each heartbeat which constitutes the oscillometric waveform, (Charbonnier et al., 2000) and introducing acoustic artifact which will be picked up by the microphone used to detect the Korotkoff sounds (Chittenden and Weaver, 1988, Walloch et al., 1990, Glasser and Ramsey, 1981).

The signal quality algorithm automatically assessed the quality of a blood pressure measurement by determining the feasibility of accurately estimating the diastolic and systolic features from the recorded signal. A detailed description on the development of the blood pressure signal quality algorithm can be found in (Abdul Sukor et al., 2012).

8.3 Methodology

8.3.1 Statistical Analysis

Hypothetically, a parameter extracted from a good quality signal will have the same value before and after the signal quality evaluation. Thus, to test the hypothesis, we have performed a Wilcoxon signed rank test. The test is performed on features extracted from SpO$_2$, heart rate, and systolic and diastolic blood pressure parameter values for each subject. The hypothesis is:

\[ H_0: \text{the median difference in the data obtained before and after the signal quality analysis equals zero} \]

\[ H_1: \text{the median difference in the data obtained before and after the signal quality analysis does not equal zero.} \]

\( H_1 \) is accepted if the significance level (\( p \)-value) is less than 0.05 (Petrie and Sabin, 2009).
8.3.2 HM-DSS Performance

The performance of a number of HM-DSSs were evaluated by reporting the accuracy, sensitivity, specificity, PPV, NPV and $\kappa$ values. The HM-DSS outputs are compared with the reference standard that has been developed in chapter seven. HM-DSSs were constructed using:

1) data originally coming from the TMC-Home device without manually removing outliers (c.f. section 5.3.1.1)
2) data originally coming from the TMC-Home device after manually removing outliers (c.f. section 5.3.1.1)
3) data extracted from the raw signals after performing the above-mentioned signal quality analysis.

The evaluated HM-DSSs were used to categorise whether to either manage COPD patients at home or to refer them for further clinical assessment. The HM-DSSs were developed using data from Database I (comprising 12 COPD patients with the data collection managed from the Austin Hospital, Victoria), as discussed in chapter seven.

The methods of constructing the HM-DSSs are shown in Figure 8-3 with the descriptions in Table 8-1. The comparison are made between: HM-DSS I, HM-DSS II and HM-DSS III, HM-DSS IV, HM-DSS V and HM-DSS VI, HM-DSS VII, HM-DSS VIII and HM-DSS IX.
Figure 8-3: Methods to construct nine types of HM-DSSs.
Table 8-1: The descriptions of the nine types of HM-DSSs.

<table>
<thead>
<tr>
<th>HM-DSS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>pulse oximetry</strong> derived parameters obtained from the database</td>
</tr>
<tr>
<td>II</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>pulse oximetry</strong> derived parameters obtained from the database after excluding outliers manually</td>
</tr>
<tr>
<td>III</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>pulse oximetry</strong> derived parameters <strong>after</strong> performing <strong>signal quality analysis</strong> on the PPG signals which led to the recalculation(^a) of SpO(_2) and heart rate values</td>
</tr>
<tr>
<td>IV</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>blood pressure</strong> signals obtained from the database</td>
</tr>
<tr>
<td>V</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>blood pressure</strong> signals obtained from the database after excluding outliers manually</td>
</tr>
<tr>
<td>VI</td>
<td>Analysis performed using <strong>only</strong> the features extracted from <strong>blood pressure</strong> signals <strong>after</strong> performing <strong>signal quality analysis</strong> which led to the removal of parameter values categorised to originate from ‘bad’ quality signals</td>
</tr>
<tr>
<td>VII</td>
<td>Analysis performed using features extracted from <strong>pulse oximetry</strong> parameters, <strong>blood pressure</strong> parameters and <strong>other(^b)</strong> parameters obtained from the database</td>
</tr>
<tr>
<td>VIII</td>
<td>Analysis performed using features extracted from <strong>pulse oximetry</strong> parameters, <strong>blood pressure</strong> parameters and <strong>other(^b)</strong> parameters obtained from the database after excluding outliers manually</td>
</tr>
<tr>
<td>IX</td>
<td>Analysis performed using features extracted from <strong>pulse oximetry</strong> parameters, <strong>blood pressure</strong> parameters and <strong>other(^b)</strong> parameters <strong>after</strong> performing <strong>signal quality analyses</strong> on the PPG and the blood pressure signals</td>
</tr>
</tbody>
</table>

\(^a\) The recalculation of SpO\(_2\) is using Equation 8. The heart rate value is estimated by averaging the time of arrival of PPG pulses (measured from peak-to-peak) as taken from the sections of good PPG signals (Abdul Sukor et al., 2011).

\(^b\) The other parameters were derived from spirometry (relaxed and forced), weight and temperature measurements.
8.4 Results

8.4.1 Examples of Corrupted Signals

The pulse oximetry signal quality algorithm was designed to use a waveform morphology analysis technique to identify good quality pulses and concurrently reject artifact in a PPG signal. Figure 8-4 shows one example of a PPG signal contaminated with noise. When the signal is analysed using the pulse oximetry signal quality algorithm, the section with bad pulses is removed and the SpO$_2$ and heart rate values are recalculated based only on the segments of the signal containing acceptable quality beats.

![Figure 8-4: Example of a PPG signal with sections of acceptable and poor quality waveform segments.](image)

In the blood pressure signal quality algorithm, the distortions of the Korotkoff sound signals are the fundamental guide to estimate the reliability of systolic and diastolic blood pressure values. In the case of the Korotkoff signal shown in Figure 8-5, the diastolic blood pressure is unable to be determined due to the signal corruption.
Figure 8-5: Example of a raw Korotkoff sound signal affected with noise and thus potential yielding unreliable diastolic blood pressure estimation.

8.4.2 Statistical Analysis

Table 8-2 shows the \( p \)-values obtained using the Wilcoxon signed rank test performed on each subject’s data before (without manual data pre-processing) and after the signal quality analysis. The \( \text{SpO}_2 \) and heart rate parameter value, distribution mean and distribution standard deviation for all 12 subjects have changed significantly after the signal quality analysis. However, \( \text{SpO}_2 \) percentage changes and z-score from most subjects have not been affected by the signal quality analysis. A different scenario is observed on the features extracted from the blood pressure parameters. Significant changes are noticed on most subject’s distribution mean, distribution standard deviation, percentage changes and z-score data.
Table 8-2: *p*-values obtained using the Wilcoxon signed rank test performed on each feature from each subject’s data without* and with signal quality analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Feature</th>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpO₂</td>
<td>N</td>
<td></td>
<td>144</td>
<td>281</td>
<td>199</td>
<td>249</td>
<td>188</td>
<td>294</td>
<td>144</td>
<td>97</td>
<td>232</td>
<td>313</td>
<td>173</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.80</td>
<td>0.28</td>
<td>0.04</td>
<td>0.265</td>
<td>0.27</td>
<td>0.05</td>
<td>0.48</td>
<td>0.14</td>
<td>0.51</td>
<td>0.60</td>
<td>0.97</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.90</td>
<td>0.33</td>
<td>0.06</td>
<td>0.215</td>
<td>0.69</td>
<td>0.56</td>
<td>0.54</td>
<td>0.04</td>
<td>0.85</td>
<td>0.99</td>
<td>0.83</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Heart rate</td>
<td>N</td>
<td></td>
<td>144</td>
<td>281</td>
<td>199</td>
<td>249</td>
<td>188</td>
<td>294</td>
<td>144</td>
<td>97</td>
<td>232</td>
<td>313</td>
<td>173</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.00</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.71</td>
<td>0.00</td>
<td>0.15</td>
<td>0.55</td>
<td>0.00</td>
<td>0.00</td>
<td>0.57</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.46</td>
<td>0.00</td>
<td>0.31</td>
<td>0.05</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.73</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.39</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.65</td>
<td>0.85</td>
<td>0.19</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.99</td>
<td>0.11</td>
<td>0.00</td>
<td>0.10</td>
<td>0.09</td>
<td>0.49</td>
<td>0.00</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Diastolic</td>
<td>N</td>
<td></td>
<td>219</td>
<td>276</td>
<td>187</td>
<td>241</td>
<td>155</td>
<td>176</td>
<td>145</td>
<td>97</td>
<td>199</td>
<td>270</td>
<td>162</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.78</td>
<td>0.68</td>
<td>0.21</td>
<td>1.00</td>
<td>1.00</td>
<td>0.84</td>
<td>0.89</td>
<td>1.00</td>
<td>0.94</td>
<td>1.00</td>
<td>0.53</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.46</td>
<td>0.94</td>
<td>0.35</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.00</td>
<td>0.51</td>
<td>0.00</td>
<td>0.25</td>
<td>0.36</td>
<td>0.00</td>
<td>0.28</td>
<td>0.52</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.00</td>
<td>0.05</td>
<td>0.16</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.78</td>
<td>0.63</td>
<td>0.16</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.66</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.76</td>
<td>0.80</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>N</td>
<td></td>
<td>211</td>
<td>308</td>
<td>187</td>
<td>243</td>
<td>158</td>
<td>286</td>
<td>145</td>
<td>98</td>
<td>217</td>
<td>295</td>
<td>172</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.76</td>
<td>0.45</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>0.64</td>
<td>0.62</td>
<td>1.00</td>
<td>0.81</td>
<td>0.71</td>
<td>0.48</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.00</td>
<td>0.19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.48</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.00</td>
<td>0.28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.52</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.01</td>
<td>0.21</td>
<td>0.00</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

* Features are extracted without manual outlier removal
1= Parameter value, 2 = Distribution mean, 3= Distribution standard deviation, 4= Percentage changes, 5= z-score
8.4.3 Analysis Using Only Pulse Oximetry Features

The 12 COPD patients used in this study recorded 2467 pulse oximetry measurements in 2406 days (including 61 days with more than one measurement performed in a day). Out of these 2467 pulse oximetry recordings, 44 were removed due to being corrupted. Figure 8-6 and Table 8-3 show the results of the developed HM-DSSs.

![Bar charts represent HM-DSS I, HM-DSS II and HM-DSS III performance.](image)

Table 8-3: The $\kappa$ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS I, HM-DSS II and HM-DSS III.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>$\kappa$</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM-DSS I</td>
<td>369</td>
<td>0.41</td>
<td>78.53</td>
<td>60.26</td>
<td>83.45</td>
<td>49.47</td>
<td>88.64</td>
</tr>
<tr>
<td>HM-DSS II</td>
<td>366</td>
<td>0.43</td>
<td>78.71</td>
<td>64.63</td>
<td>82.70</td>
<td>51.46</td>
<td>89.18</td>
</tr>
<tr>
<td>HM-DSS III</td>
<td>365</td>
<td>0.45</td>
<td>81.69</td>
<td>56.41</td>
<td>88.54</td>
<td>57.14</td>
<td>88.24</td>
</tr>
</tbody>
</table>
8.4.4 Analysis Using Only Blood Pressure Features

A total of 2703 blood pressure parameter values measured over 2439 days were obtained from the database. Approximately 17.13% or 463 blood pressure parameter values were removed due to ‘bad’ signals as suggested by the signal quality indicator. This resulted in 079 days remaining with blood pressure measurement data and 332 cases (comprising 261 home management cases and 71 referral cases) for HM-DSS VI construction. Figure 8-7 and Table 8-4 show the results.

Figure 8-7: The bar charts detail the performance of HM-DSS IV, HM-DSS V and HM-DSS VI.
Table 8-4: The $\kappa$ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS III and HM-DSS IV.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>$\kappa$</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM-DSS IV</td>
<td>373</td>
<td>0.14</td>
<td>65.15</td>
<td>46.34</td>
<td>70.45</td>
<td>30.65</td>
<td>82.33</td>
</tr>
<tr>
<td>HM-DSS V</td>
<td>371</td>
<td>0.14</td>
<td>66.31</td>
<td>43.90</td>
<td>72.66</td>
<td>31.30</td>
<td>82.03</td>
</tr>
<tr>
<td>HM-DSS VI</td>
<td>332</td>
<td>0.27</td>
<td>75.07</td>
<td>44.16</td>
<td>83.45</td>
<td>41.98</td>
<td>84.64</td>
</tr>
</tbody>
</table>

8.4.5 Analysis Using All Features

The analysis for HM-DSS VII, HM-DSS VIII and HM-DSS IX used all available physiological measurement data. This included parameters from pulse oximetry, blood pressure, spirometry, weight and temperature measurements. Figure 8-8 and Table 8-5 show the relative performance of the HM-DSSs.

![Figure 8-8: The bar charts detail the performance of HM-DSS VII, HM-DSS VIII and HM-DSS IX.](image-url)
Table 8-5: The $\kappa$ value, accuracy, sensitivity, specificity, PPV, and NPV for HM-DSS VII, HM-DSS VIII and HM-DSS IX

<table>
<thead>
<tr>
<th></th>
<th>$n$</th>
<th>$\kappa$</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM-DSS VII</td>
<td>373</td>
<td>0.48</td>
<td>78.71</td>
<td>64.63</td>
<td>82.70</td>
<td>51.46</td>
<td>89.18</td>
</tr>
<tr>
<td>HM-DSS VIII</td>
<td>371</td>
<td>0.52</td>
<td>81.40</td>
<td>74.39</td>
<td>83.39</td>
<td>55.90</td>
<td>91.98</td>
</tr>
<tr>
<td>HM-DSS IX</td>
<td>366</td>
<td>0.51</td>
<td>82.79</td>
<td>65.38</td>
<td>87.50</td>
<td>58.65</td>
<td>90.32</td>
</tr>
</tbody>
</table>

8.5 Discussion

In this study, we only performed signal quality analysis on pulse oximetry and blood pressure signals and not spirometry signals. This was because; the TMC-Home spirometry measurement has adopted the international established standard for its quality assurance. The standard was developed by the ATS and European Respiratory Society (ERS) task force for standardisation of lung function testing. The TMC-Home spirometry quality module assures that the spirometry parameter values come from correctly acquired signals. For example, if patients were using wrong techniques in performing the spirometry measurements, the device will reject the measurements and prompt a notification on the TMC-Home monitor. The notification will advise the patient for the reasons of failing the measurement and request the patient to start the measurement again. The reasons that a failed spirometry measurement may occur include; the patient hesitates at the start of the measurement, the patient takes an extra breath during the measurement, the patient does not exhale fully or the patient exhales too slowly (Miller et al., 2005).

The technique performed in analysing the pulse oximetry signal quality is similar to the approach taken by Suzuki et al. and Ogawa et al.. These researchers equipped a system called Welfare Techno House with a quality indicator algorithm for the ECG signal. The recorded ECG waveforms were categorised into three classes; ‘excellent’, ‘good’ and ‘not good’ (Kawarada et al., 1998, Ogawa et al., 2002, Suzuki et al., 2001).

Figure 8-6 and Table 8-3 show the results for HM-DSS I, HM-DSS II and HM-DSS III. HM-DSS II and HM-DSS III demonstrate an improvement in terms of $\kappa$ value, accuracy, specificity and PPV when compared to HM-DSS I. However, comparing HM-DSS I and HM-DSS III, the sensitivity has decreased from 60.26% to 56.41% ($p>0.05$).
The removal of blood pressure parameter values which come from ‘poor’ quality signals does affect the performance of the HM-DSS constructed using features extracted from blood pressure parameters alone. The $\kappa$ value increased from 0.14 (HM-DSS IV) to 0.27 (HM-DSS VI), and the accuracy improved from 65.15% to 75.07% ($p<0.05$). This enhanced the agreement between the blood pressure parameters and the health status reference standard from slight to fair agreement.

Overall, the quality of the pulse oximetry and the blood pressure signals does affect the performance of the HM-DSS. The result shows that both HM-DSS VIII and HM-DSS IX have improved the $\kappa$ value of HM-DSS VII, from 0.48 to 0.52 ($p<0.05$) (comparison between HM-DSS VII and HM-DSS VIII) and from 0.48 to 0.51 ($p<0.05$) (comparison between HM-DSS VII and HM-DSS XI). HM-DSS XI has a slightly higher accuracy of 82.79% compared to HM-DSS VIII accuracy of 81.40%.

Based on the results, it was found that both manual data pre-processing and signal quality analysis techniques have improved the usability of the parameters extracted from the pulse oximetry measurements. However, only signal quality analysis improved the blood pressure measurement usability in the HM-DSS analysis.

Conventionally, in other studies, SpO$_2$ data derived from artifact-contaminated signals were ignored (Pilling and Cutaia, 1999). This analysis has provided a possible alternative solution to this approach, where, data from contaminated PPG signals have been ‘corrected’ before use.

Furthermore, the ability of the signal quality algorithms to detect ‘poor’ quality signals can be used as an early notification tool for pulse oximetry and blood pressure measurements in the unsupervised telehealth environment. In this manner, patients can be prompted to retake a measurement if the automated algorithms flag the initial measurement as corrupted.

There are assumptions made in this chapter: 1) there are enough acceptable waveform beats for the recalculation of SpO$_2$ and heart rate in the PPG signal after sections with artifact have been removed; 2) The TMC-Home device uses the same generally accepted equations in deriving the SpO$_2$ and heart rate values from the PPG signal and, 3) the major source of error in the pulse oximetry and the blood pressure measurements are contributed by motion artifacts.

In the future, there are a number of improvements that could be made to improve the outcomes of this study:
1) The pulse oximetry measurements could be performed for a longer period of time, consequently increasing the likelihood of obtaining longer sections of acceptable pulses;
2) The analysis could be enhanced by using signal quality algorithms that could detect all sources of poor signal quality, not only motion artifacts.

Unfortunately, it was not possible to perform the data quality analysis on data obtained from Database II. This is because of a technical issue in accessing the raw signal data due to issues with data synchronisation.

Nevertheless, the analyses performed in this chapter has confirmed that data quality issues do affect the reliability of the recommendations generated by the HM-DSS, in the case of the particular algorithms and data sets from Database I.
Chapter 9 : Summary

This thesis has centred on the design of a DSS to assist with the management of COPD patients using telehealth systems. The scope of work presented included the description and analysis of clinical measurement data contained in retrospective databases; development of algorithms to facilitate COPD management; and the proposal of a HM-DSS framework for a prospective application. This chapter summarises these findings and presents possible future work.

Chapter two highlighted the conventional way of managing COPD patients, which involved the investment of substantial resources and associated costs. These factors were drivers for the establishment of standard care schemes to manage suitable COPD patients in their home environment. The scheme involved clinical ambulatory carers monitoring COPD patients on a regular basis and performing vital intervention when these COPD patients experienced exacerbations.

Consequently, to assist with COPD management, telehealth technologies have been integrated with standard care. The maturation of the telehealth technologies to remotely monitor patients has enabled proactive management by the ambulatory care teams with the COPD patients.

However, by introducing home telehealth to be used by patients enrolled in standard home care, we have to consider the workload imposition on the ambulatory care team. This is because the personnel will have to monitor a much more comprehensive and voluminous data set as captured by the home monitoring device. Thus, this thesis has introduced a customised HM-DSS to assist the ambulatory care team in the task of monitoring and assessing the patient’s telehealth measurement data.

Chapter three listed the problems that we are trying to solve, which are: 1) to design a HM-DSS framework that can be integrated with available home telehealth systems; 2) to detect patients who appear to be at a high risk of exacerbation; 3) to generate referral recommendations; and 4) to assess the effect of home telehealth data quality on the HM-DSS performance.

Consequently, in chapter four, a customised HM-DSS design using the BPM system was proposed. In the implementation, we considered the process for the home management of COPD patients using telehealth with a HM-DSS, and introduced an informational flow diagram. The flow diagram identified the possible users and systems involved and activities conducted by the various parties. A three-tier architecture was created. The components involved were the
data access layer; the logic layer; and the front-end layer. These components have been implemented using an open source technology approach to minimise cost and maximise interoperability. In the future, the proposed design will be deployed in a prospective trial.

Nonetheless, the HM-DSS could improve the likelihood of obtaining a more complete data set in a real-time data collection; by automatically prompting the patient user to perform his/her scheduled measurements and proactively sending reminders to the patient to enter missing data. The automatic reminder will also replace the manual aide memoire from the clinical carer thus reducing the carer workload.

The main component in the HM-DSS is the rules engine. Chapter five presented the construction of the rules engine which began with the utilisation of data from two types of databases. The home telehealth data were collected from two groups of COPD patients enrolled as intervention groups in two randomised controlled trials. The patients in Database I were monitored by ambulatory care nurses, while patients in Database II were monitored by a RACS team.

The major limitation in this thesis was the data used in the analysis came from two retrospective studies that were conducted for different purposes. Thus, as the data came from two different studies performed at two different locations: 1) around the Austin hospital in the state of Victoria (Database I) and, 2) around the Blacktown hospital in the state of NSW (Database II); there were inconsistencies in the types of information obtained. Data from Database I did not contain formal diagnosis from the ambulatory care teams while data from Database II did not have sufficient spirometry measurements. These differences and deficiencies, which had occurred due to different trial management and diverse clinical opinions, have affected the analyses. The effects are; 1) data from Database I and II had to be analysed separately with a specific reference standard for each database; 2) the reference standard in Database I had been generated using symptoms and medications questionnaires, thus, only features extracted from physiological measurement parameters were used as predictors; and 3) analyses performed using data from Database II had only included features extracted from pulse oximetry, blood pressure, weight and temperature measurements and questionnaire answers as predictors. In the future, it is recommended the differences in terms of trial administration and non-standardised clinical opinion in managing COPD patients should be addressed in order to facilitate meta-analysis of collected data.
Nevertheless, the limitation has added to the distinctiveness of this thesis within this field of study, where to our knowledge, this thesis has presented the initial attempt of performing a multidisciplinary work with multiple types of analysis using different types of retrospective data.

In this thesis, a CART technique was used to generate the rules, which is the core component of the HM-DSS. The home telehealth data were pre-processed and features were extracted from each acquired physiological signal. The CART technique examined all the extracted features as a group, and found those features that were most predictive by applying specific cut thresholds to each feature to split the data into two or more classes.

In view of the fact that the main condition that affects the health condition of a COPD patient is the exacerbation of their symptoms, the initial algorithm in Chapter six was constructed to predict patients who were at high risk of these exacerbation events. The development was based on a reference standard generated from guidelines for Database I and the RACS provisional diagnosis for Database II. The performance of the CART produced from the data in Database I were: accuracy of 79.00%, specificity of 87.90%, sensitivity of 70.70%, PPV of 85.71%, NPV of 74.59% and \( \kappa \) value of 0.59, and the performance of the CART from Database II were: accuracy of 76.72%, specificity of 91.43%, sensitivity of 41.74%, PPV of 67.19%, NPV of 90.32% and \( \kappa \) value of 0.37. The absences of the lung functions measurements in Database II decreased the \( \kappa \) value as compared to CART from Database I (from 0.59 to 0.37).

Nevertheless, chapter six has presented CART classifier models that were used to retrospectively identify exacerbation events in COPD subjects, undergoing home telehealth monitoring, using simple statistical features. The CART constructed from Database I data had used a reference standard defined by the guidelines (generated from patient’s symptoms and medications information). The advantage of the constructed CART is that it can automatically detect a possible risk of exacerbations that are not identified by the patients themselves, or their carer or the ambulatory health-care workers. On the other hand, the CART model built using Database II data was based on diagnosis provided by the RACS team. The diagnoses were only obtained if communication occurred between patients or patient’s carer and the RACS team. Therefore, the Database II CART is useful in assisting the RACS team with the consultation, but not in detecting unreported COPD exacerbations. As a conclusion, in chapter six we have obtained two set of knowledge (rules) that can be used to detect both unreported and reported COPD exacerbations.
The subsequent CART algorithms discussed in chapter seven are meant to support the ambulatory care team with the management of the telehealth patient. This is because not all exacerbation events need hospital or GP interventions. Thus, the algorithms in this chapter generate referral recommendations for individuals who are experiencing an exacerbation. We have used two types of reference standards: 1) generated from a carer journal for Database I; and 2) constructed using RACS recommendations for Database II. The CART’s $\kappa$ values were 0.52 and 0.45 obtained from analysis using Database I and Database II respectively. In addition to the CARTs developed using all features from all measurement types, multivariate CART classifiers were also constructed using features from each individual measurement type. Analyses in both databases have shown that features from pulse oximetry measurement alone are also useful in classifying patients who need to be referred.

The resulting rules from the work done in chapter seven has made the HM-DSS more comprehensive. As well as detecting possible COPD exacerbation events, the HM-DSS could also assist the ambulatory care team with referral decisions. However, the health status reference standards could be further improved by using the information in the free-text medical records, once all personal information is de-identified. If the free-text notes are made available (an opportunity that may be available in the future) an advanced text mining algorithms will be performed to dissect the information contained in the free-text notes in order to generate more comprehensive reference standards.

Furthermore, in chapter eight, with an objective to demonstrate the effect of data quality on the HM-DSS, the CART, which was developed from Database I (in chapter seven) was evaluated using available quality tools. The data quality analyses were performed on raw signals from two types of measurements; the pulse oximetry and the blood pressure. The results of this chapter have confirmed that pulse oximetry and blood pressure data quality issues do affect the reliability of the HM-DSS for the particular algorithms and data sets used herein. Both techniques (the manual outlier removal and the automated signal quality analysis) have improved the performance of the HM-DSS. Therefore, these automated signal quality tools are considered useful and will be included in the HM-DSS for the purpose of data quality assurance. This finding has also provided an additional method that can reduce the workload imposed when performing signal recording verification manually.

Overall, in this thesis, we have developed the rules engine and presented a HM-DSS design to assist in the management of COPD patients and we have also performed a preliminary investigation of the effect of data quality issues on HM-DSS operation. The HM-DSS is
designed in such a way to promote the usage of telehealth systems in a standard care environment and to facilitate the monitoring of additional COPD patients without impacting on the ambulatory care team’s workload. We believe the future success of the HM-DSS design relies on improvements to the rules. Future larger scale prospective trials of COPD patients will be required to perform these tasks; (i) to refine and validate such rules and (ii) to assess the impact of the CART algorithms misclassification on longer-term patient outcomes. Moreover, in the future it is recommended to include more types of data that could be measured at the patient’s home and included in the CART analysis. Examples are the parameters that are associated with activity of daily living of a COPD patient and the air quality index within the patient’s home environment.
References


surveillance system for patient at home: The MEDIVILLE system. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).


BARNETT, M. 2006. Chronic Obstructive Pulmonary Disease in Primary Care, West Sussex, England, John Wiley and Sons, Ltd.


KOIZUMI, T., TAKIZAWA, M., NAKAI, K., YAMAMOTO, Y., MURASE, S., FUJII, T., KOBAYASHI, T., HATAYAMA, O., FUJIMOTO, K. & KUBO, K. 2005. Trial of remote telemedicine support for patients with chronic respiratory failure at home


WALTERS, J. 2010. COPD - Diagnosis, management and the role of the GP. *Australian Family Physician*, 39, 100-103.


YANG, X., REDMOND, S. J., BASILAKIS, J. & LOVELL, N. H. Year. Effect of ECG quality measures on piecewise-linear trend detection for telehealth decision support systems.


Appendices
Appendix A: Chapter six 12 sets of rules obtained from the 12 trained CARTs using data from Database I.

Rules set 1

1  if StdHR<8.19317 then node 2 else node 3
2  if DataSpO2<97.495 then node 4 else node 5
3  if StdHR<8.57141 then node 6 else node 7
4  if StdSystolic<11.4983 then node 8 else node 9
5  if MeanSystolic<147.161 then node 10 else node 11
6  if MeanIC<1.56111 then node 12 else node 13
7  if StdHR<11.7061 then node 14 else node 15
8  if StdDiastolic<7.13261 then node 16 else node 17
9  if StdVT<0.109615 then node 18 else node 19
10 if DataTemp<34.985 then node 20 else node 21
11 class = Low risk
12 if MeanSpO2<98.6303 then node 22 else node 23
13 if DataIC<1.64716 then node 24 else node 25
14 class = Low risk
15 if MeanHR<101.413 then node 26 else node 27
16 class = High risk
17 class = Low risk
18 if MeanHR<91.3661 then node 28 else node 29
19 if DataFEV1<0.817624 then node 30 else node 31
20 class = Low risk
21 if DataBR<14.8251 then node 32 else node 33
22 class = Low risk
23 class = High risk
24 class = High risk
25 if StdWeight<0.517997 then node 34 else node 35
26 class = High risk
27 if StdFEV6<0.614112 then node 36 else node 37
28 class = High risk
29 class = Low risk
30 class = High risk
31 class = Low risk
32 class = Low risk
33 class = High risk
34 class = High risk
35 class = Low risk
36 if ChangesFEV1<-19.9017 then node 38 else node 39
37 if MeanSpO2<98.4407 then node 40 else node 41
38 class = High risk
39 class = Low risk
40 if MeanDiastolic<70.5484 then node 42 else node 43
41 if StdFEV1<0.0538929 then node 44 else node 45
42 class = Low risk
43 class = High risk
44 class = Low risk
45 if MeanDiastolic<69.5806 then node 46 else node 47
46 if zScoreTemp<-1.92834 then node 48 else node 49
class = Low risk
47 class = Low risk
48 class = High risk

Rules set 2

1 if MeanHR<74.6224 then node 2 else node 3
2 if StdVT<0.125527 then node 4 else node 5
3 if MeanSpO2<98.8911 then node 6 else node 7
4 if StdDiastolic<15.0061 then node 8 else node 9
5 if MeanBR<16.9676 then node 10 else node 11
6 if MeanIC<1.53484 then node 12 else node 13
7 if StdFEV1<0.0876794 then node 14 else node 15
8 if MeanHR<67.1718 then node 16 else node 17
9 class = Low risk
10 class = High risk
11 if MeanTemp<35.6103 then node 18 else node 19
12 if ChangesWeight<4.17557 then node 20 else node 21
13 if MeanTemp<35.7661 then node 22 else node 23
14 class = High risk
15 if ChangesFEV1<-18.8199 then node 24 else node 25
16 class = Low risk
17 class = High risk
18 class = High risk
19 if zScoreHR<-0.960114 then node 26 else node 27
20 if zScoreHR<2.65938 then node 28 else node 29
21 if DataWeight<47.55 then node 30 else node 31
22 if ChangesFEV6<-0.69778 then node 32 else node 33
23 if zScoreTi2Tot<-0.963361 then node 34 else node 35
24 class = High risk
25 class = Low risk
26 class = High risk
27 class = Low risk
28 class = Low risk
29 if DataHR<111.49 then node 36 else node 37
30 class = High risk
31 if StdFVC<0.05635 then node 38 else node 39
32 if StdVE<1.2805 then node 40 else node 41
33 if DataHR<75.285 then node 42 else node 43
34 class = High risk
35 if StdFEV1<0.0189823 then node 44 else node 45
36 class = High risk
37 if MeanDiastolic<74.9032 then node 46 else node 47
38 class = High risk
39 if zScoreSystolic<-2.6053 then node 48 else node 49
40 class = Low risk
41 if DataSpO2<97.665 then node 50 else node 51
42 class = High risk
43 if DataSpO2<93.555 then node 52 else node 53
44 class = High risk
45 class = Low risk
46 class = Low risk
47 class = High risk
48 class = High risk
49  class = Low risk
50  class = High risk
51  class = Low risk
52  class = High risk
53  class = Low risk

Rules set 3

1 if MeanSpO2<97.9848 then node 2 else node 3
2 if MeanTemp<35.5919 then node 4 else node 5
3 if StdHR<8.57141 then node 6 else node 7
4 if ChangesFEV6<-1.36564 then node 8 else node 9
5 if MeanWeight<45.1508 then node 10 else node 11
6 if StdVT<0.125527 then node 12 else node 13
7 if StdFEV6<0.614112 then node 14 else node 15
8 if MeanBR<12.5956 then node 16 else node 17
9 if DataHR<76.87 then node 18 else node 19
10  class = High risk
11  if zScoreWeight<-0.679188 then node 20 else node 21
12  if StdSpO2<1.40825 then node 22 else node 23
13  if MeanBR<16.9676 then node 24 else node 25
14  if ChangesIC<-23.4101 then node 26 else node 27
15  if DataFEV1<1.08734 then node 28 else node 29
16  class = High risk
17  if DataBR<12.3313 then node 30 else node 31
18  class = High risk
19  if StdHR<7.33742 then node 32 else node 33
20  if zScoreFVC<-2.28069 then node 34 else node 35
21  if StdWeight<1.31471 then node 36 else node 37
22  if MeanTi2Tot<0.315281 then node 38 else node 39
23  class = Low risk
24  class = High risk
25  if MeanTemp<35.6103 then node 40 else node 41
26  class = High risk
27  class = Low risk
28  class = High risk
29  if StdWeight<0.393133 then node 42 else node 43
30  class = High risk
31  if DataHR<92.77 then node 44 else node 45
32  class = High risk
33  if DataSpO2<93.555 then node 46 else node 47
34  class = High risk
35  if StdSpO2<0.70813 then node 48 else node 49
36  class = Low risk
37  if StdWeight<1.3879 then node 50 else node 51
38  class = Low risk
39  if MeanHR<67.1718 then node 52 else node 53
40  class = High risk
41  if MeanHR<70.1121 then node 54 else node 55
42  class = Low risk
43  if DataSystolic<149.5 then node 56 else node 57
44  class = Low risk
45  class = High risk
46  class = High risk
class = High risk
47 class = Low risk
48 class = Low risk
49 class = High risk
50 class = High risk
51 if zScoreSystolic<-0.456058 then node 58 else node 59
52 class = Low risk
53 class = High risk
54 class = Low risk
55 class = High risk
56 class = High risk
57 class = Low risk
58 class = High risk
59 if ChangesBR<-29.0037 then node 60 else node 61
60 class = High risk
61 class = Low risk

Rules set 4

Decision tree for classification
1 if MeanSpO2<97.9848 then node 2 else node 3
2 if MeanBR<12.5938 then node 4 else node 5
3 if StdHR<8.57141 then node 6 else node 7
4 if StdDiastolic<14.0487 then node 8 else node 9
5 if MeanIC<1.7371 then node 10 else node 11
6 if StdVT<0.125527 then node 12 else node 13
7 if StdFEV<0.614112 then node 14 else node 15
8 if MeanSpO2<96.2148 then node 16 else node 17
9 class = Low risk
10 if ChangesWeight<-3.5044 then node 18 else node 19
11 if StdWeight<0.567234 then node 20 else node 21
12 if StdSpO2<1.40825 then node 22 else node 23
13 if MeanBR<-16.9676 then node 24 else node 25
14 if ChangesIC<-23.4101 then node 26 else node 27
15 if DataFEV1<1.08734 then node 28 else node 29
16 class = High risk
17 class = Low risk
18 if DataBR<12.8304 then node 30 else node 31
19 if DataWeight<47.55 then node 32 else node 33
20 class = High risk
21 if ChangesTemp<2.56106 then node 34 else node 35
22 if MeanTi2Tot<0.315281 then node 36 else node 37
23 class = Low risk
24 class = High risk
25 if MeanTemp<35.6103 then node 38 else node 39
26 class = High risk
27 class = Low risk
28 class = High risk
29 if StdWeight<0.393133 then node 40 else node 41
30 if StdDiastolic<8.77111 then node 42 else node 43
31 if MeanTi2Tot<0.400875 then node 44 else node 45
32 class = High risk
33 if DataWeight<57.75 then node 46 else node 47
34 if StdFEV1<0.0189823 then node 48 else node 49
35 class = High risk

161
class = Low risk
if MeanHR<67.1718 then node 50 else node 51
class = High risk
if MeanHR<70.1121 then node 52 else node 53
class = Low risk
if DataSystolic<149.5 then node 54 else node 55
class = High risk
if ChangesSpO2<2.18064 then node 56 else node 57
if MeanFEV6<2.23256 then node 58 else node 59
class = High risk
class = Low risk
if DataHR<88.565 then node 60 else node 61
class = High risk
class = Low risk
class = Low risk
class = High risk
class = High risk
class = Low risk
if MeanFVC<3.11769 then node 62 else node 63
class = High risk
class = Low risk
if MeanFEV6<2.23366 then node 64 else node 65
class = High risk
if DataFEV1<0.560827 then node 66 else node 67
class = High risk
class = High risk
if StdFVC<0.0715293 then node 68 else node 69
class = High risk
if DataBR<13.2331 then node 70 else node 71
class = High risk
class = Low risk
class = Low risk

Rules set 5
1 if StdIC<0.0713701 then node 2 else node 3
2 if MeanFEV6<1.61291 then node 4 else node 5
3 if MeanTemp<35.5919 then node 6 else node 7
4 class = Low risk
5 if DataFEV6<1.90473 then node 8 else node 9
6 if ChangesFEV6<-4.14107 then node 10 else node 11
7 if MeanWeight<45.1508 then node 12 else node 13
8 class = High risk
9 if DataHR<94.76 then node 14 else node 15
10 if DataPEF<3.01452 then node 16 else node 17
11 if MeanFEV6<1.24833 then node 18 else node 19
12 class = High risk
13 if StdWeight<0.562757 then node 20 else node 21
14 class = High risk
15 if StdFEV1<0.0538929 then node 22 else node 23
16  if ChangesPEF<-21.3365 then node 24 else node 25
17  class = Low risk
18  class = High risk
19  if MeanFEV6<2.2293 then node 26 else node 27
20  if StdHR<8.05839 then node 28 else node 29
21  if DataWeight<55.475 then node 30 else node 31
22  class = Low risk
23  if DataFVC<1.97199 then node 32 else node 33
24  class = Low risk
25  class = High risk
26  if MeanIC<1.59189 then node 34 else node 35
27  if DataHR<78.915 then node 36 else node 37
28  class = High risk
29  class = Low risk
30  if zScoreFVC<-2.34639 then node 38 else node 39
31  if ChangesBR<-32.7854 then node 40 else node 41
32  class = High risk
33  class = Low risk
34  if DataSpO2<93.155 then node 42 else node 43
35  class = High risk
36  class = Low risk
37  class = High risk
38  class = High risk
39  if StdTemp<0.209362 then node 44 else node 45
40  class = High risk
41  if MeanIC<2.18888 then node 46 else node 47
42  class = High risk
43  class = Low risk
44  class = High risk
45  class = Low risk
46  if ChangesSystolic<-21.7431 then node 48 else node 49
47  class = Low risk
48  class = High risk
49  class = Low risk

Rules set 6

1  if MeanSpO2<97.9848 then node 2 else node 3
2  if MeanIC<1.55639 then node 4 else node 5
3  if StdHR<8.57141 then node 6 else node 7
4  if ChangesWeight<4.17557 then node 8 else node 9
5  if MeanTemp<35.6202 then node 10 else node 11
6  if StdVT<0.125527 then node 12 else node 13
7  if StdFEV6<0.614112 then node 14 else node 15
8  if ChangesDiastolic<27.0837 then node 16 else node 17
9  if DataWeight<47.55 then node 18 else node 19
10  if DataFVC<2.85271 then node 20 else node 21
11  if StdFEV1<0.0189823 then node 22 else node 23
12  if StdSpO2<1.40825 then node 24 else node 25
13  if MeanBR<16.9676 then node 26 else node 27
14  if ChangesIC<-23.4101 then node 28 else node 29
15  if DataFEV1<1.08734 then node 30 else node 31
16  if DataFEV1<1.28645 then node 32 else node 33
17  class = Low risk
Rules set 7

Decision tree for classification

18  class = High risk
19  if StdFVC<0.05635 then node 34 else node 35
20  class = High risk
21  if DataVT<0.921639 then node 36 else node 37
22  class = High risk
23  class = Low risk
24  if MeanTi2Tot<0.315281 then node 38 else node 39
25  class = Low risk
26  class = High risk
27  if MeanTemp<35.6103 then node 40 else node 41
28  class = High risk
29  class = Low risk
30  class = High risk
31  if StdWeight<0.393133 then node 42 else node 43
32  if MeanFEV6<2.13195 then node 44 else node 45
33  class = Low risk
34  class = High risk
35  if zScoreSystolic<-2.6053 then node 46 else node 47
36  if DataIC<1.56218 then node 48 else node 49
37  class = Low risk
38  class = Low risk
39  if MeanHR<67.1718 then node 50 else node 51
40  class = High risk
41  if MeanHR<70.1121 then node 52 else node 53
42  class = Low risk
43  class = Low risk
44  class = High risk
45  if DataSystolic<149.5 then node 54 else node 55
46  class = High risk
47  class = Low risk
48  class = Low risk
49  class = High risk
50  class = Low risk
51  class = High risk
52  class = Low risk
53  class = High risk
54  class = High risk
55  class = Low risk
56  class = High risk
57  class = Low risk

Rules set 7

Decision tree for classification

1  if MeanSpO2<97.9848 then node 2 else node 3
2  if StdWeight<1.96067 then node 4 else node 5
3  if StdHR<8.57141 then node 6 else node 7
4  if StdSpO2<0.694132 then node 8 else node 9
5  if MeanWeight<45.1508 then node 10 else node 11
6  if StdVT<0.125527 then node 12 else node 13
7  if StdFEV6<0.614112 then node 14 else node 15
8  class = High risk
9  if DataFEV6<2.70583 then node 16 else node 17
10 class = High risk

164
11 if ChangesSystolic<=-5.1089 then node 18 else node 19
12 if StdSpO2<1.40825 then node 20 else node 21
13 if MeanBR<16.9676 then node 22 else node 23
14 if ChangesIC<=-23.4101 then node 24 else node 25
15 if DataFEV1<1.08734 then node 26 else node 27
16 class = Low risk
17 if MeanHR<=81.6336 then node 28 else node 29
18 class = High risk
19 if DataBR<13.2331 then node 30 else node 31
20 if MeanTi2Tot<0.315281 then node 32 else node 33
21 class = Low risk
22 class = High risk
23 if MeanTemp<35.6103 then node 34 else node 35
24 class = High risk
25 class = Low risk
26 class = High risk
27 if StdWeight<0.393133 then node 36 else node 37
28 class = High risk
29 if StdFVC<0.0584864 then node 38 else node 39
30 class = High risk
31 class = Low risk
32 class = Low risk
33 if MeanHR<=67.1718 then node 40 else node 41
34 class = High risk
35 if MeanHR<=70.1121 then node 42 else node 43
36 class = Low risk
37 if DataSystolic<=149.5 then node 44 else node 45
38 class = High risk
39 class = Low risk
40 class = Low risk
41 class = High risk
42 class = Low risk
43 class = High risk
44 class = High risk
45 class = Low risk

Rules set 8

1 if MeanSpO2<=97.9593 then node 2 else node 3
2 if MeanFVC<=2.99471 then node 4 else node 5
3 if zScoreSpO2<-3.11978 then node 6 else node 7
4 if StdWeight<1.96067 then node 8 else node 9
5 if ChangesFEV1<=-0.0251613 then node 10 else node 11
6 class = Low risk
7 if StdHR<8.29982 then node 12 else node 13
8 if StdSpO2<=0.694132 then node 14 else node 15
9 if MeanWeight<=45.1508 then node 16 else node 17
10 if StdVE<1.2805 then node 18 else node 19
11 if ChangesHR<=14.8571 then node 20 else node 21
12 if StdVT<0.127346 then node 22 else node 23
13 if DataFEV6<=2.16881 then node 24 else node 25
14 class = High risk
15 if DataFEV6<=2.67423 then node 26 else node 27
16 class = High risk
if StdFVC<0.0561835 then node 28 else node 29
class = Low risk
if MeanSpO2<96.5418 then node 30 else node 31
class = High risk
if MeanBR<11.9826 then node 32 else node 33
if StdSpO2<1.40825 then node 34 else node 35
if MeanBR<16.9676 then node 36 else node 37
class = High risk
if StdSystolic<11.0437 then node 38 else node 39
if zScoreHR<2.91029 then node 40 else node 41
if MeanFVC<2.96204 then node 54 else node 55
class = Low risk
Rules set 9
1 if MeanSpO2<97.9848 then node 2 else node 3
2 if MeanIC<1.53484 then node 4 else node 5
3 if StdHR<8.18823 then node 6 else node 7
4 if ChangesWeight<4.17557 then node 8 else node 9
5 if MeanTemp<35.6202 then node 10 else node 11
if StdVT<0.125527 then node 12 else node 13
if StdFEV6<0.614112 then node 14 else node 15
if ChangesHR<24.3553 then node 16 else node 17
if DataWeight<47.55 then node 18 else node 19
if ChangesFEV6<-0.69778 then node 20 else node 21
if MeanIC<1.7371 then node 22 else node 23
if MeanTi2Tot<0.315281 then node 24 else node 25
if ChangesIC<-23.4101 then node 28 else node 29
if DataFEV1<1.08734 then node 30 else node 31
class = Low risk
if MeanDiastolic<74.9032 then node 32 else node 33
class = High risk
if StdFVC<0.05635 then node 34 else node 35
class = Low risk
if StdVE<1.2805 then node 36 else node 37
class = Low risk
class = High risk
if MeanIC<1.7371 then node 22 else node 23
class = Low risk
if MeanTi2Tot<0.315281 then node 24 else node 25
class = Low risk
class = High risk
if MeanHR<67.1718 then node 42 else node 43
class = Low risk
if MeanHR<67.1718 then node 42 else node 43
class = High risk
if MeanDiastolic<74.9032 then node 32 else node 33
class = High risk
if MeanHR<67.1718 then node 42 else node 43
class = High risk
Rules set 10

1  if MeanSpO2<97.9848 then node 2 else node 3
2  if MeanIC<1.53484 then node 4 else node 5
3  if StdHR<8.57141 then node 6 else node 7
4  if DataWeight<55.475 then node 8 else node 9
5  if MeanTemp<35.6202 then node 10 else node 11
6  if StdVT<0.125527 then node 12 else node 13
7  if StdFEV6<0.614112 then node 14 else node 15
8  if zScoreFVC<-2.53536 then node 16 else node 17
9  class = Low risk
10 if ChangesFEV6<-0.69778 then node 18 else node 19
11 if MeanIC<1.7371 then node 20 else node 21
12 if MeanTi2Tot<0.315281 then node 22 else node 23
13 if MeanBR<16.9676 then node 24 else node 25
14 if ChangesIC<-23.4101 then node 26 else node 27
15 if DataFEV1<1.08734 then node 28 else node 29
16 class = High risk
17 if DataVT<0.329737 then node 30 else node 31
18 if StdVE<1.2805 then node 32 else node 33
19 if DataHR<75.285 then node 34 else node 35
20 class = Low risk
21 if StdWeight<0.566028 then node 36 else node 37
22 class = Low risk
23 if StdSpO2<1.40825 then node 38 else node 39
24 class = High risk
25 if MeanTemp<35.6103 then node 40 else node 41
26 class = High risk
27 class = Low risk
28 class = High risk
29 if StdWeight<0.393133 then node 42 else node 43
30 class = High risk
31 if StdWeight<1.2702 then node 44 else node 45
32 class = Low risk
33 if DataSpO2<97.665 then node 46 else node 47
34 class = High risk
35 if DataSpO2<93.555 then node 48 else node 49
36 class = High risk
37 if StdFEV1<0.0189823 then node 50 else node 51
38 if MeanHR<67.1718 then node 52 else node 53
39 class = Low risk
40 class = High risk
41 if MeanHR<70.1121 then node 54 else node 55
42 class = Low risk
43 if DataSystolic<149.5 then node 56 else node 57
44 class = Low risk
45 class = High risk
46 class = High risk
47 class = Low risk
48 class = High risk
49 class = Low risk
50 class = High risk
51 class = Low risk

168
169

class = Low risk
class = High risk
class = Low risk
class = High risk
class = High risk
class = Low risk

Rules set 11

1 if MeanSpO2<97.9848 then node 2 else node 3
2 if MeanIC<1.53484 then node 4 else node 5
3 if MeanSystolic<147.161 then node 6 else node 7
4 if ChangesWeight<4.17557 then node 8 else node 9
5 if MeanTemp<35.7661 then node 10 else node 11
6 if MeanTi2Tot<0.315281 then node 12 else node 13
7 if DataWeight<49.875 then node 14 else node 15
8 if ChangesHR<24.3553 then node 16 else node 17
9 if DataWeight<47.55 then node 18 else node 19
10 if ChangesFEV6<0.69778 then node 20 else node 21
11 if MeanIC<1.7371 then node 22 else node 23
class = Low risk
12 if MeanHR<66.9335 then node 24 else node 25
class = High risk
class = Low risk
class = Low risk
13 if MeanDiastolic<74.9032 then node 26 else node 27
class = High risk
14 if StdFVC<0.05635 then node 28 else node 29
15 if StdVE<1.2805 then node 30 else node 31
16 if DataHR<75.285 then node 32 else node 33
class = Low risk
17 if zScoreHR<1.46661 then node 34 else node 35
class = Low risk
class = High risk
class = High risk
18 if zScoreSystolic<-2.6053 then node 38 else node 39
class = Low risk
19 if DataSpO2<97.665 then node 40 else node 41
class = High risk
20 if DataSpO2<93.555 then node 42 else node 43
class = High risk
21 if DataHR<74.755 then node 46 else node 47
class = High risk
class = Low risk
22 if StdFEV1<0.0189823 then node 44 else node 45
class = High risk
23 if DataHR<74.755 then node 46 else node 47
class = High risk
class = Low risk
24 if DataSpO2<97.665 then node 40 else node 41
class = High risk
class = Low risk
class = High risk
25 if DataSpO2<93.555 then node 42 else node 43
class = High risk
class = Low risk
class = Low risk
26 if DataSpO2<97.665 then node 40 else node 41
class = High risk
class = Low risk
class = High risk
27 if DataSpO2<93.555 then node 42 else node 43
class = High risk
class = Low risk
class = Low risk
28 if DataSpO2<97.665 then node 40 else node 41
class = High risk
class = Low risk
class = High risk
class = Low risk
47 if zScoreFEV6<1.70679 then node 48 else node 49
48 class = High risk
49 if DataHR<94.76 then node 50 else node 51
50 class = High risk
51 if StdFEV1<0.0538929 then node 52 else node 53
52 class = Low risk
53 if MeanDiastolic<69.5806 then node 54 else node 55
54 if zScoreTemp<-1.92834 then node 56 else node 57
55 class = Low risk
56 class = Low risk
57 class = High risk

Rules set 12
1 if MeanSpO2<97.9848 then node 2 else node 3
2 if MeanIC<1.53484 then node 4 else node 5
3 if StdHR<8.57141 then node 6 else node 7
4 if ChangesWeight<4.17557 then node 8 else node 9
5 if MeanTemp<35.6202 then node 10 else node 11
6 if StdVT<0.125527 then node 12 else node 13
7 if StdFEV6<0.614112 then node 14 else node 15
8 if ChangesHR<24.3553 then node 16 else node 17
9 if DataWeight<47.55 then node 18 else node 19
10 if ChangesFEV6<-0.69778 then node 20 else node 21
11 if MeanIC<1.7371 then node 22 else node 23
12 if MeanTi2Tot<0.315281 then node 24 else node 25
13 if MeanBR<16.9676 then node 26 else node 27
14 class = Low risk
15 if MeanDiastolic<69.5806 then node 28 else node 29
16 class = Low risk
17 if MeanDiastolic<74.9032 then node 30 else node 31
18 class = High risk
19 if StdFVC<0.05635 then node 32 else node 33
20 if StdVE<1.2805 then node 34 else node 35
21 if DataHR<75.285 then node 36 else node 37
22 class = Low risk
23 if MeanFEV1<1.31194 then node 38 else node 39
24 class = Low risk
25 if StdSpO2<1.40825 then node 40 else node 41
26 class = High risk
27 if MeanTemp<35.6103 then node 42 else node 43
28 if StdFEV1<0.0538929 then node 44 else node 45
29 class = Low risk
30 class = Low risk
31 class = High risk
32 class = High risk
33 if zScoreSystolic<-2.6053 then node 46 else node 47
34 class = Low risk
35 if DataSpO2<97.665 then node 48 else node 49
36 class = High risk
37 if DataSpO2<93.555 then node 50 else node 51
38 if StdFEV1<0.0189823 then node 52 else node 53
39 class = Low risk
40 if MeanHR<67.1718 then node 54 else node 55
41 class = Low risk
42 class = High risk
43 if MeanHR<70.1121 then node 56 else node 57
44 class = Low risk
45 if zScoreTemp<-1.92834 then node 58 else node 59
46 class = High risk
47 class = Low risk
48 class = High risk
49 class = Low risk
50 class = High risk
51 class = Low risk
52 class = High risk
53 class = Low risk
54 class = Low risk
55 class = High risk
56 class = Low risk
57 class = High risk
58 class = Low risk
59 class = High risk
Appendix B: Chapter six 16 sets of rules obtained from the 16 trained CARTs using data from Database II.

Rules set 1

1. if BreathQ<2.25 then node 2 else node 3
2. if zScoreSpO2<-0.38339 then node 4 else node 5
3. if zScoreDiastolic<0.907988 then node 6 else node 7
4. if zScoreSpO2<-0.444293 then node 8 else node 9
5. if DataHR<75.3872 then node 10 else node 11
6. if MeanWeight<55.5532 then node 12 else node 13
7. if MeanDiastolic<83.3871 then node 14 else node 15
8. if DataDiastolic<79.75 then node 16 else node 17
   class = High risk
9. if SColorQ<1.5 then node 18 else node 19
10. if MeanHR<60.4573 then node 20 else node 21
11. if StdHR<7.02049 then node 22 else node 23
12. if ChangesWeight<1.67064 then node 24 else node 25
13. if zScoreTemp<-1.48994 then node 26 else node 27
   class = Low risk
14. if zScoreHR<0.935775 then node 28 else node 29
   class = High risk
15. class = Low risk
16. if zScoreHR<0.935775 then node 28 else node 29
17. class = High risk
18. class = Low risk
19. if StdHR<6.27545 then node 30 else node 31
20. class = High risk
21. if StdWeight<2.40006 then node 32 else node 33
22. class = Low risk
23. class = High risk
24. if zScoreSystolic<-1.19499 then node 34 else node 35
25. if StdWeight<0.663353 then node 36 else node 37
26. class = Low risk
27. if StdSpO2<0.294167 then node 38 else node 39
28. if StdDiastolic<15.9413 then node 40 else node 41
29. if DataHR<105.934 then node 42 else node 43
30. class = Low risk
31. class = High risk
32. class = Low risk
33. if DataTemp<34.55 then node 44 else node 45
34. class = High risk
35. if zScoreSystolic<-2.17176 then node 46 else node 47
36. class = Low risk
37. if SColorQ<0.5 then node 48 else node 49
38. class = Low risk
39. class = High risk
40. if GFeelQ<2.5 then node 50 else node 51
41. class = High risk
42. class = High risk
43. class = Low risk
44. class = High risk
45. class = Low risk
46. if DataSpO2<84.855 then node 52 else node 53
47. class = High risk
48. class = Low risk
49  if SColorQ<1.75 then node 54 else node 55
50  if SColorQ<2.5 then node 56 else node 57
51  class = Low risk
52  class = High risk
53  if MeanSpO2<84.979 then node 58 else node 59
54  class = High risk
55  if MeanTemp<35.76 then node 60 else node 61
56  class = Low risk
57  if StdHR<12.4147 then node 62 else node 63
58  class = High risk
59  if DataTemp<37.45 then node 64 else node 65
60  class = Low risk
61  if DataSpO2<97.225 then node 66 else node 67
62  class = Low risk
63  class = High risk
64  if CoughQ<1 then node 68 else node 69
65  class = Low risk
66  if StdDiastolic<17.3798 then node 70 else node 71
67  class = Low risk
68  class = High risk
69  class = Low risk
70  class = High risk
71  class = Low risk

Rules set 2

1  if BreathQ<2.25 then node 2 else node 3
2  if zScoreSpO2<-0.383339 then node 4 else node 5
3  if zScoreDiastolic<0.918153 then node 6 else node 7
4  if zScoreSpO2<-0.444293 then node 8 else node 9
5  if DataHR<75.3872 then node 10 else node 11
6  if MeanWeight<55.5532 then node 12 else node 13
7  if DataSystolic<162.5 then node 14 else node 15
8  if GFeelQ<2.5 then node 16 else node 17
9  class = High risk
10  if SColorQ<1.5 then node 18 else node 19
11  if MeanHR<60.4573 then node 20 else node 21
12  if StdHR<7.02049 then node 22 else node 23
13  if ChangesWeight<1.67064 then node 24 else node 25
14  if zScoreTemp<-1.48994 then node 26 else node 27
15  class = Low risk
16  if StdDiastolic<14.0194 then node 28 else node 29
17  class = High risk
18  class = Low risk
19  if StdHR<6.27545 then node 30 else node 31
20  class = High risk
21  if StdWeight<2.40006 then node 32 else node 33
22  class = Low risk
23  class = High risk
24  if zScoreSystolic<-0.549396 then node 34 else node 35
25  if StdWeight<0.663353 then node 36 else node 37
26  class = Low risk
27  if StdSpO2<0.294167 then node 38 else node 39
28  if zScoreHR<-0.935775 then node 40 else node 41
29 class = Low risk
30 class = Low risk
31 class = High risk
32 class = Low risk
33 if zScoreTemp<-1.80105 then node 42 else node 43
34 if DataSystolic<115 then node 44 else node 45
35 if zScoreSystolic<-2.17176 then node 46 else node 47
36 class = Low risk
37 if SColorQ<0.5 then node 48 else node 49
38 class = Low risk
39 class = High risk
40 if SColorQ<2.5 then node 50 else node 51
41 if zScoreSpO2<-1.40465 then node 52 else node 53
42 class = High risk
43 class = Low risk
44 class = Low risk
45 class = High risk
46 if MeanSpO2<86.3663 then node 54 else node 55
47 class = High risk
48 class = Low risk
49 if SColorQ<1.75 then node 56 else node 57
50 class = Low risk
51 if StdHR<12.4147 then node 58 else node 59
52 class = High risk
53 class = Low risk
54 class = High risk
55 class = Low risk
56 class = High risk
57 if MeanTemp<35.76 then node 60 else node 61
58 if DataDiastolic<79.75 then node 62 else node 63
59 class = High risk
60 class = Low risk
61 if DataSpO2<97.225 then node 64 else node 65
62 class = Low risk
63 class = High risk
64 if MeanHR<81.7052 then node 66 else node 67
65 class = Low risk
66 class = Low risk
67 class = High risk

Rules set 3

1 if BreathQ<2.25 then node 2 else node 3
2 if MeanDiastolic<68.4785 then node 4 else node 5
3 if zScoreDiastolic<0.918153 then node 6 else node 7
4 if DataWeight<39.2 then node 8 else node 9
5 if zScoreSpO2<-0.382963 then node 10 else node 11
6 if MeanWeight<55.5532 then node 12 else node 13
7 if MeanDiastolic<83.3871 then node 14 else node 15
8 class = High risk
9 if zScoreHR<51.1325 then node 16 else node 17
10 if DataWeight<84.7 then node 18 else node 19
11 if DataHR<75.59 then node 20 else node 21
12 if StdHR<7.02049 then node 22 else node 23
if ChangesWeight<1.67064 then node 24 else node 25
if zScoreTemp<-1.48994 then node 26 else node 27
class = Low risk
if DataTemp<34.45 then node 28 else node 29
class = High risk
if DataSpO2<93.545 then node 30 else node 31
if MeanHR<100.244 then node 32 else node 33
if SColorQ<1.5 then node 34 else node 35
if MeanTemp<36.9823 then node 36 else node 37
class = Low risk
class = High risk
if DataHR<103.345 then node 38 else node 39
if StdWeight<0.663353 then node 40 else node 41
class = High risk
if StdSpO2<0.294167 then node 42 else node 43
class = High risk
class = Low risk
if MeanDiastolic<68.828 then node 44 else node 45
class = High risk
class = High risk
class = Low risk
class = Low risk
if StdHR<6.27545 then node 46 else node 47
if StdHR<4.5932 then node 48 else node 49
class = High risk
if zScoreSystolic<-1.19499 then node 50 else node 51
class = High risk
class = Low risk
if SColorQ<0.5 then node 52 else node 53
class = Low risk
class = High risk
class = High risk
class = Low risk
class = Low risk
class = High risk
class = High risk
if SColorQ<3.5 then node 54 else node 55
class = High risk
if DataSpO2<84.855 then node 56 else node 57
class = Low risk
class = Low risk
class = Low risk
if ChangesWeight<-1.56734 then node 58 else node 59
class = High risk
if ChangesTemp<1.83569 then node 60 else node 61
class = High risk
if MeanTemp<35.76 then node 62 else node 63
class = High risk
if zScoreTemp<-1.69836 then node 64 else node 65
if MeanSpO2<84.979 then node 66 else node 67
class = Low risk
class = Low risk
if DataSpO2<97.225 then node 68 else node 69
class = Low risk
class = High risk
class = Low risk
class = High risk
if CoughQ<1 then node 72 else node 73
if StdDiastolic<17.3798 then node 74 else node 75
class = Low risk
class = High risk
class = Low risk
class = High risk
class = Low risk

Rules set 4

1 if BreathQ<2.25 then node 2 else node 3
2 if zScoreSpO2<-0.383339 then node 4 else node 5
3 if zScoreHR<0.611673 then node 6 else node 7
4 if zScoreSpO2<-0.444293 then node 8 else node 9
5 if CoughQ<1.75 then node 10 else node 11
6 if zScoreWeight<1.32731 then node 12 else node 13
7 if zScoreDiastolic<-0.167513 then node 14 else node 15
8 if GFeelQ<2.5 then node 16 else node 17
9 class = High risk
class = Low risk
11 if SColorQ<3.5 then node 18 else node 19
12 if ChangesSpO2<4.58344 then node 20 else node 21
13 if ChangesWeight<1.92298 then node 22 else node 23
14 if StdTemp<0.23839 then node 24 else node 25
15 if MeanHR<104.746 then node 26 else node 27
16 if DataTemp<37.05 then node 28 else node 29
class = High risk
18 if MeanHR<60.4573 then node 30 else node 31
19 if DataHR<81.51 then node 32 else node 33
20 if SAmountQ<0.5 then node 34 else node 35
class = High risk
class = High risk
23 if DataWeight<65.875 then node 36 else node 37
class = Low risk
25 if DataHR<93.8313 then node 38 else node 39
26 if ChangesWeight<0.872787 then node 40 else node 41
class = Low risk
28 if SAmountQ<3.5 then node 42 else node 43
class = High risk
class = Low risk
31 if DataTemp<34.2 then node 44 else node 45
class = High risk
class = Low risk
class = High risk
35 if zScoreSystolic<-0.770246 then node 46 else node 47
class = High risk
37 if SColorQ<1.25 then node 48 else node 49
class = Low risk
39 if StdHR<15.4704 then node 50 else node 51
class = High risk
41 if StdSpO2<4.28222 then node 52 else node 53
class = High risk
class = Low risk
class = High risk
42 if StdWeight<1.7181 then node 54 else node 55
class = High risk
class = High risk
if MeanTemp < 36.9823 then node 56 else node 57
class = High risk
class = Low risk
class = Low risk
if SColorQ < 3.5 then node 58 else node 59
class = High risk
class = Low risk
class = Low risk
class = Low risk
if MeanHR < 99.6156 then node 60 else node 61
if DataDiastolic < 69.4167 then node 62 else node 63
class = Low risk
class = High risk
if zScoreTemp < -0.199075 then node 64 else node 65
class = High risk
class = Low risk
class = Low risk
if ChangesDiastolic < 4.23712 then node 66 else node 67
class = High risk
class = Low risk
if ChangesTemp < -0.631231 then node 68 else node 69
class = High risk
class = Low risk
if StdSystolic < 8.54442 then node 70 else node 71
class = High risk
class = Low risk

Rules set 5

1 if BreathQ < 2.25 then node 2 else node 3
2 if zScoreSpO2 < -0.383339 then node 4 else node 5
3 if SColorQ < 0.5 then node 6 else node 7
4 if zScoreSpO2 < -0.444293 then node 8 else node 9
5 if zScoreHR < 51.1325 then node 10 else node 11
6 if zScoreTemp < -0.142035 then node 12 else node 13
7 if zScoreDiastolic < 0.851746 then node 14 else node 15
8 if GFeelQ < 2.5 then node 16 else node 17
9 class = High risk
10 if DataTemp < 34.2 then node 18 else node 19
11 class = Low risk
12 class = High risk
13 class = Low risk
14 if MeanTemp < 35.7241 then node 20 else node 21
15 if SColorQ < 1.5 then node 22 else node 23
16 if ChangesTemp < 1.05503 then node 24 else node 25
17 class = High risk
18 class = High risk
19 if MeanTemp < 36.9823 then node 26 else node 27
20 class = Low risk
21 if ChangesSystolic < 2.49363 then node 28 else node 29
22 class = Low risk
23 if DataSystolic<106.5 then node 30 else node 31
24 if DataHR<86.39 then node 32 else node 33
25 class = Low risk
26 if DataHR<74.325 then node 34 else node 35
27 class = Low risk
28 if SConsQ<0.5 then node 36 else node 37
29 if StdWeight<0.472966 then node 38 else node 39
30 class = Low risk
31 if zScoreTemp<-0.466162 then node 40 else node 41
32 if DataHR<81.525 then node 42 else node 43
33 class = Low risk
34 if StdSpO2<1.02169 then node 44 else node 45
35 if ChangesSystolic<-18.0755 then node 46 else node 47
36 class = High risk
37 if DataSystolic<112.5 then node 48 else node 49
38 class = High risk
39 if DataDiastolic<67 then node 50 else node 51
40 class = Low risk
41 if DataWeight<52 then node 52 else node 53
42 class = Low risk
43 class = High risk
44 class = High risk
45 class = Low risk
46 class = Low risk
47 if MeanDiastolic<72.1452 then node 54 else node 55
48 class = Low risk
49 class = High risk
50 class = High risk
51 class = Low risk
52 class = Low risk
53 if ChangesSpO2<10.8099 then node 56 else node 57
54 class = Low risk
55 if StdDiastolic<5.33422 then node 58 else node 59
56 class = High risk
57 class = Low risk
58 class = High risk
59 if MeanDiastolic<72.2419 then node 60 else node 61
60 class = High risk
61 class = Low risk

Rules set 6

1 if BreathQ<2.75 then node 2 else node 3
2 if zScoreSpO2<-0.383339 then node 4 else node 5
3 if zScoreDiastolic<0.918153 then node 6 else node 7
4 if zScoreSpO2<-0.444293 then node 8 else node 9
5 if DataHR<75.3872 then node 10 else node 11
6 if MeanWeight<55.5532 then node 12 else node 13
7 if MeanDiastolic<83.3871 then node 14 else node 15
8 if GFeelQ<2.25 then node 16 else node 17
9 class = High risk
10 if SColorQ<1.5 then node 18 else node 19
11 if MeanHR<60.4573 then node 20 else node 21
12 class = High risk
13 if ChangesWeight<1.67064 then node 22 else node 23
14 if zScoreTemp<-1.48994 then node 24 else node 25
15 class = Low risk
16 if StdDiastolic<14.0194 then node 26 else node 27
17 class = High risk
18 class = Low risk
19 if StdHR<6.27545 then node 28 else node 29
20 class = High risk
21 if DataTemp<34.2 then node 30 else node 31
22 if StdHR<9.2516 then node 32 else node 33
23 if StdWeight<0.663353 then node 34 else node 35
24 class = Low risk
25 if StdSpO2<0.294167 then node 36 else node 37
26 if SColorQ<1.5 then node 38 else node 39
27 class = Low risk
28 class = Low risk
29 class = High risk
30 class = High risk
31 if MeanTemp>36.9823 then node 40 else node 41
32 if StdHR<8.35928 then node 42 else node 43
33 if ChangesWeight<-0.578512 then node 44 else node 45
34 class = Low risk
35 if SColorQ<0.5 then node 46 else node 47
36 class = Low risk
37 class = High risk
38 class = Low risk
39 if MeanHR<76.9471 then node 48 else node 49
40 if zScoreSystolic<-2.19298 then node 50 else node 51
41 class = Low risk
42 if DataHR<94.605 then node 52 else node 53
43 class = High risk
44 class = High risk
45 class = Low risk
46 class = Low risk
47 if SColorQ<1.5 then node 54 else node 55
48 class = High risk
49 if DataWeight<47.525 then node 56 else node 57
50 class = Low risk
51 if SColorQ<3.5 then node 58 else node 59
52 class = Low risk
53 class = High risk
54 class = High risk
55 if MeanTemp<35.76 then node 60 else node 61
56 class = High risk
57 class = Low risk
58 if StdSystolic<8.46516 then node 62 else node 63
59 if ChangesWeight<-1.56734 then node 64 else node 65
60 class = Low risk
61 if DataSpO2<97.225 then node 66 else node 67
62 if StdSpO2<2.57186 then node 68 else node 69
63 class = Low risk
64 class = High risk
65 class = Low risk
66 if StdDiastolic<17.3798 then node 70 else node 71
class = Low risk
class = Low risk
class = High risk
class = High risk
class = Low risk

Rules set 7

1 if GFeedQ<2.25 then node 2 else node 3
2 if DataHR<75.5047 then node 4 else node 5
3 if SColorQ<1.25 then node 6 else node 7
4 if SColorQ<1.5 then node 8 else node 9
5 if StdDiastolic<14.4148 then node 10 else node 11
6 if ChangesTemp<0.00978081 then node 12 else node 13
7 if zScoreDiastolic<0.794889 then node 14 else node 15
8 class = Low risk
9 if StdHR<6.27545 then node 16 else node 17
10 if StdHR<7.9141 then node 18 else node 19
11 class = Low risk
12 class = High risk
13 if DataWeight<40.425 then node 20 else node 21
14 if DataHR<93.025 then node 22 else node 23
15 if zScoreTemp<0.466162 then node 24 else node 25
16 class = Low risk
17 if DataSystolic<118 then node 26 else node 27
18 if zScoreSpO2<-0.356676 then node 28 else node 29
19 if SColorQ<2.5 then node 30 else node 31
20 class = High risk
21 class = Low risk
22 if ChangesDiastolic<-15.9312 then node 32 else node 33
23 if DataHR<99.9275 then node 34 else node 35
24 class = High risk
25 if DataWeight<51.175 then node 36 else node 37
26 class = Low risk
27 if DataSystolic<146.25 then node 38 else node 39
28 if zScoreSpO2<-0.526692 then node 40 else node 41
29 if StdWeight<2.40006 then node 42 else node 43
30 class = Low risk
31 if ChangesWeight<-1.52507 then node 44 else node 45
32 class = High risk
33 if ChangesHR<-23.344 then node 46 else node 47
34 if DataSystolic<149.5 then node 48 else node 49
35 if MeanHR<87.6806 then node 50 else node 51
36 class = Low risk
37 if ChangesSpO2<-10.8099 then node 52 else node 53
38 class = High risk
39 class = Low risk
40 if ChangesHR<8.0727 then node 54 else node 55
41 class = High risk
42 if StdSpO2<4.4362 then node 56 else node 57
43 class = Low risk
44 class = High risk
45 if ChangesSpO2<5.65906 then node 58 else node 59
46 class = High risk
181

if SColorQ<1.75 then node 60 else node 61
47  class = High risk
48  class = Low risk
49  class = High risk
50  class = High risk
51  if ChangesDiastolic<4.90832 then node 62 else node 63
52  class = High risk
53  class = Low risk
54  class = Low risk
55  class = High risk
56  if MeanHR<60.4573 then node 64 else node 65
57  class = High risk
58  class = Low risk
59  class = High risk
60  class = High risk
61  if CoughQ<1 then node 66 else node 67
62  class = Low risk
63  class = High risk
64  class = High risk
65  class = Low risk
66  class = High risk
67  class = Low risk

Rules set 8

1  if GFeelQ<2.25 then node 2 else node 3
2  if SConsQ<1.5 then node 4 else node 5
3  if zScoreDiastolic<0.794889 then node 6 else node 7
4  if StdDiastolic<14.4148 then node 8 else node 9
5  if DataHR<76.1325 then node 10 else node 11
6  if DataHR<93.025 then node 12 else node 13
7  if zScoreSystolic<28.3839 then node 14 else node 15
8  if zScoreHR<0.9524 then node 16 else node 17
9  class = Low risk
10  if ChangesSpO2<-1.37916 then node 18 else node 19
11  if StdWeight<0.312664 then node 20 else node 21
12  if ChangesDiastolic<-16.0443 then node 22 else node 23
13  if MeanHR<89.0692 then node 24 else node 25
14  class = High risk
15  class = Low risk
16  if zScoreWeight<2.18047 then node 26 else node 27
17  if zScoreSpO2<-0.345058 then node 28 else node 29
18  class = Low risk
19  if StdHR<7.15387 then node 30 else node 31
20  class = High risk
21  if ChangesWeight<-1.52507 then node 32 else node 33
22  class = High risk
23  if ChangesWeight<1.67064 then node 34 else node 35
24  class = High risk
25  if DataHR<99.8875 then node 36 else node 37
26  if StdSpO2<0.73199 then node 38 else node 39
27  if StdWeight<0.502334 then node 40 else node 41
28  class = High risk
29  if zScoreHR<49.0126 then node 42 else node 43
30  class = Low risk
if DataSystolic<146.25 then node 44 else node 45
if ChangesWeight<0.746066 then node 46 else node 47
if SConsQ<0.5 then node 48 else node 49
class = Low risk
if StdWeight<0.492748 then node 50 else node 51
class = High risk
if StdWeight<2.40006 then node 52 else node 53
class = Low risk
if SColorQ<2.5 then node 54 else node 55
class = High risk
if MeanDiastolic<74.9113 then node 56 else node 57
class = High risk
class = High risk
class = Low risk
class = Low risk
if zScoreWeight<0.814753 then node 58 else node 59
class = High risk
if CoughQ<1 then node 60 else node 61
class = High risk
class = Low risk
class = High risk
class = High risk
if ChangesWeight<1.90976 then node 68 else node 69
class = High risk
class = Low risk
class = Low risk
class = Low risk
if DataWeight<69.2 then node 72 else node 73
class = Low risk
if StdDiastolic<7.61622 then node 74 else node 75
class = Low risk
if BreathQ<2.5 then node 76 else node 77
class = High risk
class = Low risk
class = Low risk
if ChangesTemp<1.09144 then node 78 else node 79
class = High risk
class = Low risk
class = Low risk
Rules set 9
1 if GFeelQ<2.25 then node 2 else node 3
2 if SConsQ<1.5 then node 4 else node 5
3 if SColorQ<1.25 then node 6 else node 7
4 if zScoreHR<0.9524 then node 8 else node 9
5 if DataHR<76.1325 then node 10 else node 11
6 if ChangesTemp<-0.119 then node 12 else node 13
7 if zScoreDiastolic<0.794889 then node 14 else node 15
8 if StdDiastolic<3.05013 then node 16 else node 17
9 if StdSystolic<10.3676 then node 18 else node 19
10 if ChangesSpO2<-1.37916 then node 20 else node 21
11 if ChangesWeight<-1.52507 then node 22 else node 23
12 class = High risk
13 if DataWeight<40.425 then node 24 else node 25
14 if DataHR<93.025 then node 26 else node 27
15 if DataSystolic<95 then node 28 else node 29
16 class = High risk
17 if StdDiastolic<18.1829 then node 30 else node 31
18 class = High risk
19 if ChangesWeight<-2.24786 then node 32 else node 33
20 class = Low risk
21 if StdHR<7.15387 then node 34 else node 35
22 class = High risk
23 if ChangesWeight<0.746066 then node 36 else node 37
24 class = High risk
25 class = Low risk
26 if ChangesDiastolic<-15.9312 then node 38 else node 39
27 if DataHR<99.9275 then node 40 else node 41
28 class = Low risk
29 if StdTemp<0.210227 then node 42 else node 43
30 if SColorQ<3.5 then node 44 else node 45
31 class = High risk
32 class = High risk
33 class = Low risk
34 class = Low risk
35 if DataSystolic<146.25 then node 46 else node 47
36 class = Low risk
37 if zScoreWeight<0.814753 then node 48 else node 49
38 class = High risk
39 if ChangesHR<-23.344 then node 50 else node 51
40 if DataSystolic<149.5 then node 52 else node 53
41 if zScoreHR<0.853529 then node 54 else node 55
42 class = Low risk
43 if DataHR<116.947 then node 56 else node 57
44 if MeanTemp<36.9823 then node 58 else node 59
45 class = Low risk
46 class = High risk
47 class = Low risk
48 class = High risk
49 if ChangesWeight<1.90976 then node 60 else node 61
50 class = High risk
51 if SColorQ<1.75 then node 62 else node 63
52 class = High risk
53 class = Low risk
54 class = Low risk
55 if StdTemp<-0.300155 then node 64 else node 65
56 class = High risk
class = Low risk
if DataHR<73.175 then node 66 else node 67
class = Low risk
if DataWeight<69.2 then node 68 else node 69
class = High risk
if CoughQ<1 then node 70 else node 71
class = Low risk
if DataTemp<36.6 then node 72 else node 73
class = High risk
if DataHR<70.915 then node 74 else node 75
if StdHR<4.5932 then node 76 else node 77
class = Low risk
if BreathQ<2.5 then node 78 else node 79
class = High risk
class = Low risk
class = Low risk
class = Low risk
class = Low risk
class = High risk
class = High risk
class = Low risk
if ChangesTemp<1.09144 then node 80 else node 81
class = High risk
class = Low risk
class = High risk
Rules set 10
1 if GFeelQ<2.25 then node 2 else node 3
2 if SConsQ<1.5 then node 4 else node 5
3 if SColorQ<1.25 then node 6 else node 7
4 if zScoreHR<0.9524 then node 8 else node 9
5 if DataHR<76.1325 then node 10 else node 11
6 if ChangesTemp<-0.119 then node 12 else node 13
7 if zScoreDiastolic<0.794889 then node 14 else node 15
8 if StdDiastolic<3.05013 then node 16 else node 17
9 if StdSystolic<10.3676 then node 18 else node 19
10 if ChangesSpO2<-1.37916 then node 20 else node 21
11 if ChangesWeight<-1.52507 then node 22 else node 23
class = High risk
12 if DataWeight<40.425 then node 24 else node 25
13 if DataHR<93.025 then node 26 else node 27
14 if DataSystolic<95 then node 28 else node 29
class = High risk
15 if StdHR<7.15387 then node 30 else node 31
class = High risk
16 if ChangesDiastolic<18.1829 then node 32 else node 33
class = High risk
17 if ChangesWeight<-2.24786 then node 34 else node 33
class = Low risk
18 if StdHR<-7.15387 then node 34 else node 35
class = High risk
19 if ChangesWeight<0.746066 then node 36 else node 37
class = High risk
class = Low risk
20 if ChangesDiastolic<-15.9312 then node 38 else node 39
27 if DataHR<99.9275 then node 40 else node 41
28 class = Low risk
29 if StdTemp<0.210227 then node 42 else node 43
30 if SColorQ<3.5 then node 44 else node 45
31 class = High risk
32 class = High risk
33 class = Low risk
34 class = Low risk
35 if DataSystolic<146.25 then node 46 else node 47
36 class = Low risk
37 if zScoreWeight<0.814753 then node 48 else node 49
38 class = High risk
39 if ChangesHR<-23.344 then node 50 else node 51
40 if DataSystolic<149.5 then node 52 else node 53
41 if zScoreHR<0.853529 then node 54 else node 55
42 class = Low risk
43 if DataHR<116.947 then node 56 else node 57
44 if MeanTemp<36.9823 then node 58 else node 59
45 class = Low risk
46 class = High risk
47 class = Low risk
48 class = High risk
49 if ChangesWeight<1.90976 then node 60 else node 61
50 class = High risk
51 if SColorQ<1.75 then node 62 else node 63
52 class = High risk
53 class = Low risk
54 class = Low risk
55 if StdTemp<0.300155 then node 64 else node 65
56 class = High risk
57 class = Low risk
58 if DataHR<73.175 then node 66 else node 67
59 class = Low risk
60 class = Low risk
61 if DataWeight<69.2 then node 68 else node 69
62 class = High risk
63 if CoughQ<1 then node 70 else node 71
64 class = Low risk
65 if DataTemp<36.6 then node 72 else node 73
66 if DataHR<70.915 then node 74 else node 75
67 if StdHR<4.5932 then node 76 else node 77
68 class = Low risk
69 if BreathQ<2.5 then node 78 else node 79
70 class = High risk
71 class = Low risk
72 class = High risk
73 class = Low risk
74 class = Low risk
75 class = High risk
76 class = High risk
77 class = Low risk
78 if ChangesTemp<1.09144 then node 80 else node 81
79 class = High risk
80 class = Low risk
81
class = High risk

Rules set II

1. if GFeelQ<2.25 then node 2 else node 3
2. if SConsQ<1.5 then node 4 else node 5
3. if SColorQ<1.25 then node 6 else node 7
4. if zScoreHR<0.9524 then node 8 else node 9
5. if DataHR<76.1325 then node 10 else node 11
6. if ChangesTemp<-0.119 then node 12 else node 13
7. if zScoreDiastolic<0.794889 then node 14 else node 15
8. if zScoreWeight<2.18047 then node 16 else node 17
9. if zScoreSystolic<-1.40883 then node 18 else node 19
10. if ChangesSpO2<-1.37916 then node 20 else node 21
11. if MeanWeight<55.55 then node 22 else node 23
12. class = High risk
13. if DataWeight<40.425 then node 24 else node 25
14. if DataHR<93.025 then node 26 else node 27
15. if DataSystolic<95 then node 28 else node 29
16. if StdSpO2<0.73199 then node 30 else node 31
17. if StdWeight<0.502334 then node 32 else node 33
18. class = High risk
19. if zScoreHR<0.991238 then node 34 else node 35
20. class = Low risk
21. if StdHR<7.15387 then node 36 else node 37
22. class = High risk
23. if ChangesWeight<1.90976 then node 38 else node 39
24. class = High risk
25. class = Low risk
26. if ChangesDiastolic<-15.9312 then node 40 else node 41
27. if DataHR<99.9275 then node 42 else node 43
28. class = Low risk
29. if DataWeight<98.4 then node 44 else node 45
30. class = High risk
31. if StdHR<4.5219 then node 46 else node 47
32. class = Low risk
33. if MeanTemp<35.7286 then node 48 else node 49
34. class = High risk
35. if MeanHR<60.4573 then node 50 else node 51
36. class = Low risk
37. if DataSystolic<146.25 then node 52 else node 53
38. class = Low risk
39. if DataWeight<69.2 then node 54 else node 55
40. class = High risk
41. if ChangesHR<-23.344 then node 56 else node 57
42. if DataSystolic<149.5 then node 58 else node 59
43. if zScoreHR<0.853529 then node 60 else node 61
44. class = High risk
45. if StdTemp<0.210227 then node 62 else node 63
46. class = High risk
47. if StdDiastolic<18.1829 then node 64 else node 65
48. class = High risk
49. if ChangesHR<19.9637 then node 66 else node 67
50. class = High risk
51 class = Low risk
52 class = High risk
53 class = Low risk
54 class = Low risk
55 if BreathQ<2.5 then node 68 else node 69
56 class = High risk
57 if SColorQ<1.75 then node 70 else node 71
58 class = High risk
59 class = Low risk
60 class = Low risk
61 if StdTemp<0.300155 then node 72 else node 73
62 class = Low risk
63 class = High risk
64 if StdWeight<2.40006 then node 74 else node 75
65 class = High risk
66 class = High risk
67 class = Low risk
68 if ChangesTemp<1.09144 then node 76 else node 77
69 class = High risk
70 class = High risk
71 if CoughQ<1 then node 78 else node 79
72 class = Low risk
73 if DataTemp<36.6 then node 80 else node 81
74 class = Low risk
75 class = High risk
76 class = Low risk
77 class = High risk
78 class = High risk
79 class = Low risk
80 class = High risk
81 class = Low risk

Rules set 12
1 if BreathQ<2.25 then node 2 else node 3
2 if DataHR<75.5047 then node 4 else node 5
3 if zScoreDiastolic<0.918153 then node 6 else node 7
4 if SColorQ<1.5 then node 8 else node 9
5 if zScoreSpO2<-0.373407 then node 10 else node 11
6 if MeanWeight<55.5532 then node 12 else node 13
7 if MeanDiastolic<83.3871 then node 14 else node 15
8 class = Low risk
9 if ChangesDiastolic<-2.43868 then node 16 else node 17
10 if zScoreSpO2<-0.444293 then node 18 else node 19
11 if zScoreHR<51.1325 then node 20 else node 21
12 if StdHR<7.02049 then node 22 else node 23
13 if ChangesWeight<1.67064 then node 24 else node 25
14 if zScoreTemp<-1.48994 then node 26 else node 27
15 class = Low risk
16 class = Low risk
17 class = High risk
18 if ChangesTemp<1.05503 then node 28 else node 29
19 class = High risk
20 if MeanTemp<36.9823 then node 30 else node 31
21 class = High risk
22 class = Low risk
23 class = High risk
24 if zScoreSystolic<-0.549396 then node 32 else node 33
25 if StdWeight<0.663353 then node 34 else node 35
26 class = High risk
27 if StdSpO2<0.294167 then node 36 else node 37
28 if StdWeight<0.575645 then node 38 else node 39
29 if StdDiastolic<9.28896 then node 40 else node 41
30 if DataTemp<34.2 then node 42 else node 43
31 class = Low risk
32 if DataSystolic<115 then node 44 else node 45
33 if MeanSpO2<86.3663 then node 46 else node 47
34 class = Low risk
35 if SColorQ<0.5 then node 48 else node 49
36 class = Low risk
37 class = High risk
38 class = High risk
39 class = Low risk
40 class = High risk
41 class = Low risk
42 class = High risk
43 if zScoreSystolic<-2.19298 then node 50 else node 51
44 class = Low risk
45 class = High risk
46 class = High risk
47 if DataSystolic<160 then node 52 else node 53
48 class = Low risk
49 if SColorQ<1.75 then node 54 else node 55
50 class = High risk
51 if SColorQ<3.5 then node 56 else node 57
52 if CoughQ<1 then node 58 else node 59
53 class = High risk
54 class = High risk
55 if MeanTemp<35.76 then node 60 else node 61
56 if StdSystolic<8.46516 then node 62 else node 63
57 if ChangesWeight<-1.56734 then node 64 else node 65
58 class = High risk
59 class = Low risk
60 class = Low risk
61 if DataSpO2<97.225 then node 66 else node 67
62 if StdSpO2<2.57186 then node 68 else node 69
63 class = Low risk
64 class = High risk
65 class = Low risk
66 if StdDiastolic<17.3798 then node 70 else node 71
67 class = Low risk
68 class = Low risk
69 class = High risk
70 class = High risk
71 class = Low risk

Rules set 13
1) if BreathQ < 2.25 then node 2 else node 3
2) if DataHR < 75.5047 then node 4 else node 5
3) if zScoreDiastolic < 0.918153 then node 6 else node 7
4) if SColorQ < 1.5 then node 8 else node 9
5) if GFeelQ < 2.75 then node 10 else node 11
6) if MeanWeight < 55.5532 then node 12 else node 13
7) if MeanDiastolic < 83.3871 then node 14 else node 15
8) class = Low risk
9) if StdHR < 6.27545 then node 16 else node 17
10) if StdWeight < 2.40006 then node 18 else node 19
11) if MeanSpO2 < 93.4552 then node 20 else node 21
12) if StdHR < 7.02049 then node 22 else node 23
13) if ChangesWeight < 1.67064 then node 24 else node 25
14) if zScoreTemp < -1.48994 then node 26 else node 27
15) class = Low risk
16) class = Low risk
17) if ChangesDiastolic < -2.43868 then node 28 else node 29
18) if MeanSpO2 < 97.6102 then node 30 else node 31
19) if SColorQ < 1.5 then node 32 else node 33
20) class = Low risk
21) class = High risk
22) class = Low risk
23) class = High risk
24) if zScoreSystolic < -1.04102 then node 34 else node 35
25) if DataWeight < 65.875 then node 36 else node 37
26) class = Low risk
27) if StdSpO2 < 0.294167 then node 38 else node 39
28) class = Low risk
29) class = High risk
30) if StdDiastolic < 14.4152 then node 40 else node 41
31) class = High risk
32) if ChangesSpO2 < -2.02146 then node 42 else node 43
33) if ChangesTemp < 1.09144 then node 44 else node 45
34) class = High risk
35) if ChangesSpO2 < 4.43792 then node 46 else node 47
36) class = High risk
37) if MeanTemp < 35.76 then node 48 else node 49
38) class = Low risk
39) class = High risk
40) if MeanHR < 60.4573 then node 50 else node 51
41) class = Low risk
42) class = High risk
43) class = Low risk
44) if DataTemp < 34.55 then node 52 else node 53
45) class = High risk
46) if zScoreSystolic < 2.17176 then node 54 else node 55
47) class = High risk
48) class = Low risk
49) if SColorQ < 0.5 then node 56 else node 57
50) class = High risk
51) class = Low risk
52) class = High risk
53) if MeanTemp < 36.9419 then node 58 else node 59
54) if DataSpO2 < 84.855 then node 60 else node 61
class = High risk
class = Low risk
if SColorQ<2.5 then node 62 else node 63
class = Low risk
class = High risk
class = High risk
if DataTemp<37.45 then node 64 else node 65
class = High risk
if DataSpO2<97.225 then node 66 else node 67
class = High risk
class = Low risk
class = High risk
class = Low risk

Rules set 14

1 if BreathQ<2.25 then node 2 else node 3
2 if MeanDiastolic<68.4785 then node 4 else node 5
3 if zScoreDiastolic<-0.918153 then node 6 else node 7
4 if zScoreHR<51.1325 then node 8 else node 9
5 if zScoreSpO2<-0.382963 then node 10 else node 11
6 if MeanWeight<55.5532 then node 12 else node 13
7 if MeanDiastolic<83.3871 then node 14 else node 15
8 if MeanSystolic<142.925 then node 16 else node 17
9 class = High risk
10 if DataWeight<85.975 then node 18 else node 19
11 if DataHR<75.59 then node 20 else node 21
12 if StdHR<7.02049 then node 22 else node 23
13 if ChangesWeight<1.67064 then node 24 else node 25
14 if zScoreTemp<-1.48994 then node 26 else node 27
15 class = Low risk
16 class = Low risk
17 class = High risk
18 if zScoreSpO2<-0.505994 then node 28 else node 29
19 if StdTemp<0.243512 then node 30 else node 31
20 if SColorQ<1.5 then node 32 else node 33
21 if MeanTemp<36.9823 then node 34 else node 35
22 class = Low risk
23 class = High risk
24 if StdHR<9.2516 then node 36 else node 37
25 if StdWeight<0.663353 then node 38 else node 39
26 class = Low risk
27 if StdSpO2<0.294167 then node 40 else node 41
28 if MeanSpO2<98.1774 then node 42 else node 43
29 class = High risk
30 class = Low risk
31 if ChangesDiastolic<-19.1324 then node 44 else node 45
32 class = Low risk
33 if StdHR<6.27545 then node 46 else node 47
34 if StdHR<4.5932 then node 48 else node 49
35 class = High risk
36 if GFeelQ<2.75 then node 50 else node 51
if ChangesWeight<-0.578512 then node 52 else node 53
class = Low risk
if SColorQ<0.5 then node 54 else node 55
class = Low risk
class = High risk
class = High risk
class = Low risk
class = High risk
class = Low risk
if SColorQ<3.5 then node 56 else node 57
class = Low risk
if NTimeSumQ<2.5 then node 58 else node 59
class = High risk
class = Low risk
class = Low risk
if SColorQ<1.75 then node 60 else node 61
class = Low risk
if ChangesWeight<-1.56734 then node 62 else node 63
class = High risk
class = Low risk
class = High risk
class = High risk
if SColorQ<3.5 then node 56 else node 57
class = Low risk
if SColorQ<1.75 then node 60 else node 61
class = Low risk

Rules set 15

if BreathQ<2.25 then node 2 else node 3
if zScoreSpO2<-0.383339 then node 4 else node 5
if zScoreDiastolic<-0.918153 then node 6 else node 7
if SColorQ<-0.444293 then node 8 else node 9
class = High risk
if SColorQ<1.5 then node 10 else node 11
class = Low risk
class = High risk
class = Low risk
if SColorQ<1.5 then node 10 else node 11
class = High risk
if MeanSystolic<116.409 then node 72 else node 73
class = Low risk
class = High risk
15 class = Low risk
16 if StdDiastolic<14.0194 then node 28 else node 29
17 class = High risk
18 class = Low risk
19 if StdHR<6.27545 then node 30 else node 31
20 class = High risk
21 if DataTemp<34.2 then node 32 else node 33
22 class = Low risk
23 class = High risk
24 if zScoreSystolic<-1.04102 then node 34 else node 35
25 if StdWeight<0.663353 then node 36 else node 37
26 class = Low risk
27 if StdSpO2<0.294167 then node 38 else node 39
28 if zScoreHR<0.935775 then node 40 else node 41
29 class = Low risk
30 class = Low risk
31 class = High risk
32 class = High risk
33 if MeanTemp<36.9823 then node 42 else node 43
34 class = High risk
35 if ChangesSpO2<4.43792 then node 44 else node 45
36 class = Low risk
37 if SColorQ<0.5 then node 46 else node 47
38 class = Low risk
39 class = High risk
40 if SColorQ<2.5 then node 48 else node 49
41 if zScoreSpO2<-1.40465 then node 50 else node 51
42 if zScoreSystolic<-2.19298 then node 52 else node 53
43 class = Low risk
44 if zScoreSystolic<-2.17176 then node 54 else node 55
45 class = High risk
46 class = Low risk
47 if SColorQ<1.75 then node 56 else node 57
48 class = Low risk
49 if MeanHR<77.1602 then node 58 else node 59
50 class = Low risk
51 class = High risk
52 class = Low risk
53 if SColorQ<3.5 then node 60 else node 61
54 if DataSpO2<84.855 then node 62 else node 63
55 class = High risk
56 class = High risk
57 if DataSpO2<94.105 then node 64 else node 65
58 class = High risk
59 if DataWeight<47.525 then node 66 else node 67
60 if StdSystolic<8.46516 then node 68 else node 69
61 if ChangesWeight<-1.56734 then node 70 else node 71
62 class = High risk
63 if DataTemp<37.45 then node 72 else node 73
64 class = High risk
65 if StdSystolic<8.82143 then node 74 else node 75
66 class = High risk
67 if StdHR<12.4147 then node 76 else node 77
68 if StdSpO2<2.57186 then node 78 else node 79
class = Low risk
class = High risk
class = Low risk
if CoughQ<1 then node 80 else node 81
class = Low risk
class = High risk
class = Low risk
class = Low risk
class = Low risk
class = High risk
class = High risk
class = Low risk
class = High risk
class = High risk
class = Low risk
Rules set 16

1 if BreathQ<2.25 then node 2 else node 3
2 if zScoreSpO2<-0.383339 then node 4 else node 5
3 if zScoreDiastolic<0.983847 then node 6 else node 7
4 if zScoreSpO2<-0.444293 then node 8 else node 9
5 if DataHR<75.3872 then node 10 else node 11
6 if MeanWeight<55.5532 then node 12 else node 13
7 if SColorQ<1.5 then node 14 else node 15
8 if SColorQ<2.5 then node 16 else node 17
9 class = High risk
10 if SColorQ<1.5 then node 18 else node 19
11 if zScoreHR<51.1325 then node 20 else node 21
12 if StdHR<7.02049 then node 22 else node 23
13 if ChangesWeight<1.67064 then node 24 else node 25
14 class = Low risk
15 if zScoreDiastolic<1.04031 then node 26 else node 27
16 if StdHR<7.82281 then node 28 else node 29
17 if MeanHR<77.1602 then node 30 else node 31
18 class = Low risk
19 if StdHR<6.27545 then node 32 else node 33
20 if DataTemp<34.2 then node 34 else node 35
21 class = Low risk
22 class = Low risk
23 class = High risk
24 if zScoreSystolic<-1.04102 then node 36 else node 37
25 if StdWeight<0.663353 then node 38 else node 39
26 class = High risk
27 if DataSystolic<95 then node 40 else node 41
28 if ChangesHR<8.0727 then node 42 else node 43
29 if ChangesSystolic<-11.5955 then node 44 else node 45
30 class = High risk
31 if DataDiastolic<81 then node 46 else node 47
32 class = Low risk
33 class = High risk
34 class = High risk
35 if zScoreSystolic<-2.19298 then node 48 else node 49
36 class = High risk
37 if ChangesSpO2<4.43792 then node 50 else node 51
38 class = Low risk
39  if SColorQ<0.5 then node 52 else node 53
40  class = Low risk
41  if zScoreTemp<-0.466162 then node 54 else node 55
42  class = Low risk
43  class = High risk
44  class = High risk
45  class = Low risk
46  if ChangesSpO2<-0.467794 then node 56 else node 57
47  class = High risk
48  class = Low risk
49  if SColorQ<3.5 then node 58 else node 59
50  if DataSpO2<84.855 then node 60 else node 61
51  class = High risk
52  class = Low risk
53  if SColorQ<1.75 then node 62 else node 63
54  class = High risk
55  if DataWeight<52.2 then node 64 else node 65
56  if StdHR<12.4147 then node 66 else node 67
57  class = High risk
58  if SdSystolic<8.46516 then node 68 else node 69
59  if ChangesWeight<-1.56734 then node 70 else node 71
60  class = High risk
61  if ChangesHR<29.5639 then node 72 else node 73
62  class = High risk
63  if MeanTemp<35.76 then node 74 else node 75
64  class = Low risk
65  if ChangesSpO2<10.8099 then node 76 else node 77
66  class = Low risk
67  class = High risk
68  if StdSpO2<2.57186 then node 78 else node 79
69  class = Low risk
70  class = High risk
71  class = Low risk
72  if CoughQ<1 then node 80 else node 81
73  class = High risk
74  class = Low risk
75  if DataSpO2<97.225 then node 82 else node 83
76  class = High risk
77  class = Low risk
78  class = Low risk
79  class = High risk
80  class = High risk
81  class = Low risk
82  if StdDiastolic<17.4195 then node 84 else node 85
83  class = Low risk
84  class = High risk
85  class = Low risk
Appendix C: Chapter seven 12 sets of rules obtained from the 12 trained CARTs using data from Database I.

Rules set 1

1. if MeanPEF<98.0073 then node 2 else node 3
2. if StdFVC<0.121921 then node 4 else node 5
3. if MeanBR<2.43713 then node 6 else node 7
4. if DataFEV1<1.32509 then node 8 else node 9
5. if StdFVC<0.156625 then node 10 else node 11
6. if MeanWeight<76.421 then node 12 else node 13
7. if StdSpO2<1.55537 then node 14 else node 15
8. class = Referral
9. class = Home management
10. class = Home management
11. if MeanSpO2<12.0842 then node 16 else node 17
12. if MeanSystolic<2.90749 then node 18 else node 19
13. if ChangesVT<-5.91613 then node 20 else node 21
14. if DataSpO2<16.1465 then node 22 else node 23
15. if ChangesFEV1<1.1895 then node 24 else node 25
16. class = Home management
17. if DataFEV6<9.48762 then node 26 else node 27
18. if MeanHR<43.3387 then node 28 else node 29
19. class = Referral
20. class = Referral
21. if DataSystolic<2.73252 then node 30 else node 31
22. class = Referral
23. class = Home management
24. class = Referral
25. class = Home management
26. class = Home management
27. if MeanHR<81.629 then node 32 else node 33
28. class = Referral
29. if DataIC<6.33849 then node 34 else node 35
30. class = Home management
31. class = Referral
32. if zScoreBR<-1.9437 then node 36 else node 37
33. class = Home management
34. if StdWeight<0.298959 then node 38 else node 39
35. class = Referral
36. class = Home management
37. class = Referral
38. class = Referral
39. if ChangesVT<16.9891 then node 40 else node 41
40. if StdIC<1.41457 then node 42 else node 43
41. if StdHR<14.8032 then node 44 else node 45
42. if zScoreTemp<2.77685 then node 46 else node 47
43. class = Home management
44. class = Home management
45. class = Referral
46. class = Home management
47. class = Referral
Rules set 2

1  if DataPEF<98.54 then node 2 else node 3
2  if MeanFEV6<10.9587 then node 4 else node 5
3  if StdBR<0.169615 then node 6 else node 7
4  if MeanWeight<56.6315 then node 8 else node 9
5  if ChangesSpO2<-9.20521 then node 10 else node 11
6  if MeanFEV1<1.08528 then node 12 else node 13
7  if MeanTi2Tot<140.629 then node 14 else node 15
8  if ChangesWeight<-2.12205 then node 16 else node 17
9  class = Referral
10  class = Home management
11  if DataFVC<1.70583 then node 18 else node 19
12  if DataIC<6.19721 then node 20 else node 21
13  class = Referral
14  class = Home management
15  class = Referral
16  class = Referral
17  if MeanFEV1<1.19641 then node 22 else node 23
18  class = Referral
19  if DataFEV1<0.694648 then node 24 else node 25
20  if MeanWeight<76.421 then node 26 else node 27
21  class = Referral
22  class = Home management
23  class = Referral
24  class = Referral
25  class = Home management
26  if ChangesSpO2<-10.5016 then node 28 else node 29
27  class = Home management
28  class = Referral
29  class = Home management

Rules set 3

1  if MeanPEF<98.0073 then node 2 else node 3
2  if StdFVC<0.120727 then node 4 else node 5
3  if MeanBR<2.43713 then node 6 else node 7
4  if MeanWeight<48.6788 then node 8 else node 9
5  if StdFVC<0.156625 then node 10 else node 11
6  if MeanWeight<76.421 then node 12 else node 13
7  if StdSpO2<1.55537 then node 14 else node 15
8  class = Home management
9  if ChangesSpO2<-9.35194 then node 16 else node 17
10  class = Home management
11  if MeanSpO2<12.0842 then node 18 else node 19
12  if MeanSystolic<2.90749 then node 20 else node 21
13  if ChangesVT<-5.91613 then node 22 else node 23
14  if DataSpO2<16.1465 then node 24 else node 25
15  if ChangesFEV1<1.1895 then node 26 else node 27
16  class = Home management
17  class = Referral
18  class = Home management
19  if DataFEV6<9.48762 then node 28 else node 29
20  if MeanHR<43.3387 then node 30 else node 31
```plaintext
21 class = Referral
22 class = Referral
23 if DataSystolic < 2.73252 then node 32 else node 33
24 class = Referral
25 class = Home management
26 class = Referral
27 class = Home management
28 class = Home management
29 if MeanHR < 81.629 then node 34 else node 35
30 class = Referral
31 if DataIC < 6.33849 then node 36 else node 37
32 class = Home management
33 class = Referral
34 if zScoreVE < 2.13726 then node 38 else node 39
35 class = Home management
36 if StdWeight < 0.298959 then node 40 else node 41
37 class = Referral
38 class = Referral
39 class = Home management
40 class = Referral
41 if MeanVT < 0.313994 then node 42 else node 43
42 class = Referral
43 if StdIC < 1.41457 then node 44 else node 45
44 if zScoreTemp < 2.77685 then node 46 else node 47
45 class = Home management
46 class = Home management
47 if DataHR < 59 then node 48 else node 49
48 class = Home management
49 class = Referral

Rules set 4

1 if MeanPEF < 98.0073 then node 2 else node 3
2 if StdFVC < 0.120727 then node 4 else node 5
3 if MeanBR < 2.43713 then node 6 else node 7
4 if zScoreWeight < 1.01245 then node 8 else node 9
5 if StdFVC < 0.156625 then node 10 else node 11
6 if MeanWeight < 76.421 then node 12 else node 13
7 if StdSpO2 < 1.55537 then node 14 else node 15
8 class = Referral
9 if ChangesDiastolic < -6.90808 then node 16 else node 17
10 class = Home management
11 if MeanVT < 0.406996 then node 18 else node 19
12 if MeanSystolic < 2.90749 then node 20 else node 21
13 if ChangesVT < -5.91613 then node 22 else node 23
14 if DataSpO2 < 16.1465 then node 24 else node 25
15 if ChangesFEV1 < 1.1895 then node 26 else node 27
16 class = Home management
17 class = Referral
18 if StdFEV1 < 0.0501056 then node 28 else node 29
19 if DataFEV1 < 0.694648 then node 30 else node 31
20 if DataIC < 6.33849 then node 32 else node 33
21 class = Referral
22 class = Referral
```

23 if DataSystolic<2.73252 then node 34 else node 35
24 class = Referral
25 class = Home management
26 class = Referral
27 class = Home management
28 class = Home management
29 class = Referral
30 class = Referral
31 class = Home management
32 if MeanHR<43.3387 then node 36 else node 37
33 class = Referral
34 class = Home management
35 class = Referral
36 class = Referral
37 if StdWeight<0.298959 then node 38 else node 39
38 class = Referral
39 if MeanVT<0.313994 then node 40 else node 41
40 class = Referral
41 if StdIC<1.41457 then node 42 else node 43
42 if zScoreTemp<2.77685 then node 44 else node 45
43 class = Home management
44 class = Home management
45 if DataHR<59 then node 46 else node 47
46 class = Home management
47 class = Referral

Rules set 5

1 if StdFVC<0.0835943 then node 2 else node 3
2 if StdDiastolic<8.77522 then node 4 else node 5
3 if MeanFEV6<10.9587 then node 6 else node 7
4 if MeanDiastolic<80.6319 then node 8 else node 9
5 if MeanFEV1<1.06588 then node 10 else node 11
6 if MeanWeight<56.6315 then node 12 else node 13
7 if ChangesTemp<25.0018 then node 14 else node 15
8 class = Referral
9 if StdWeight<0.298959 then node 16 else node 17
10 class = Referral
11 class = Home management
12 if StdPEF<1.40851 then node 18 else node 19
13 class = Referral
14 if ChangesDiastolic<-15.7811 then node 20 else node 21
15 if DataFEV1<0.694648 then node 22 else node 23
16 class = Referral
17 class = Home management
18 class = Referral
19 if StdHR<13.9874 then node 24 else node 25
20 class = Referral
21 if zScoreSystolic<-2.96742 then node 26 else node 27
22 class = Referral
23 class = Home management
24 class = Home management
25 class = Referral
26 class = Home management
27  class = Referral

Rules set 6

1  if MeanPEF<98.0073 then node 2 else node 3
2  if StdFVC<0.120727 then node 4 else node 5
3  if MeanBR<2.437137 then node 6 else node 7
4  if ChangesDiastolic<-10.1219 then node 8 else node 9
5  if StdFVC<0.156625 then node 10 else node 11
6  if MeanWeight<76.421 then node 12 else node 13
7  if StdSpO2<1.55537 then node 14 else node 15
8  class = Home management
9  class = Referral
10  class = Home management
11  if MeanVT<0.406996 then node 16 else node 17
12  if MeanSystolic<2.90749 then node 18 else node 19
13  if ChangesVT<-5.91613 then node 20 else node 21
14  if DataSpO2<16.1465 then node 22 else node 23
15  if ChangesFEV1<1.1895 then node 24 else node 25
16  if StdFEV1<0.0501056 then node 26 else node 27
17  if DataFEV1<0.694648 then node 28 else node 29
18  if DataIC<6.33849 then node 30 else node 31
19  class = Referral
20  class = Referral
21  if DataSystolic<2.73252 then node 32 else node 33
22  class = Referral
23  class = Home management
24  class = Referral
25  class = Home management
26  class = Home management
27  class = Referral
28  class = Referral
29  class = Home management
30  if MeanHR<43.3387 then node 34 else node 35
31  class = Referral
32  class = Home management
33  class = Referral
34  class = Referral
35  if StdWeight<0.298959 then node 36 else node 37
36  class = Referral
37  if MeanVT<0.313994 then node 38 else node 39
38  class = Referral
39  if StdIC<1.41457 then node 40 else node 41
40  if zScoreTemp<2.77685 then node 42 else node 43
41  class = Home management
42  class = Home management
43  if DataHR<59 then node 44 else node 45
44  class = Home management
45  class = Referral

Rules set 7

1  if MeanPEF<98.0073 then node 2 else node 3
2  if StdWeight<1.94697 then node 4 else node 5
3  if MeanBR<2.43713 then node 6 else node 7
4  if DataBR<2.45817 then node 8 else node 9
5  if MeanSpO2<18.1974 then node 10 else node 11
6  if MeanWeight<76.421 then node 12 else node 13
7  if StdSpO2<1.55537 then node 14 else node 15
8  if DataSpO2<13.977 then node 16 else node 17
9  if DataFEV1<1.10597 then node 18 else node 19
10 class = Referral
11 class = Home management
12 if MeanSystolic<2.90749 then node 20 else node 21
13 if ChangesVT<-5.91613 then node 22 else node 23
14 if DataSpO2<16.1465 then node 24 else node 25
15 if ChangesFEV1<1.1895 then node 26 else node 27
16 class = Home management
17 class = Referral
18 if DataHR<81.3333 then node 28 else node 29
19 class = Home management
20 if MeanHR<43.3387 then node 30 else node 31
21 class = Referral
22 class = Referral
23 if DataSystolic<2.73252 then node 32 else node 33
24 class = Referral
25 class = Home management
26 class = Referral
27 class = Home management
28 class = Referral
29 class = Home management
30 class = Referral
31 if DataIC<6.33849 then node 34 else node 35
32 class = Home management
33 class = Referral
34 if StdWeight<0.298959 then node 36 else node 37
35 class = Referral
36 class = Referral
37 if MeanVT<0.313994 then node 38 else node 39
38 class = Referral
39 if StdIC<1.41457 then node 40 else node 41
40 if zScoreTemp<2.77685 then node 42 else node 43
41 class = Home management
42 class = Home management
43 if DataHR<59 then node 44 else node 45
44 class = Home management
45 class = Referral

Rules set 8

1  if MeanPEF<98.0073 then node 2 else node 3
2  if StdFVC<0.120727 then node 4 else node 5
3  if MeanBR<2.43713 then node 6 else node 7
4  if zScoreWeight<1.01245 then node 8 else node 9
5  if StdFVC<0.156625 then node 10 else node 11
6  if MeanWeight<76.421 then node 12 else node 13
7  if StdSpO2<1.55537 then node 14 else node 15
8  class = Referral
9 if ChangesDiastolic<-6.90808 then node 16 else node 17
10 if ChangesFVC<-27.3141 then node 18 else node 19
11 if MeanSpO2<12.0842 then node 20 else node 21
12 if MeanHR<43.3387 then node 22 else node 23
13 if ChangesVT<-5.91613 then node 24 else node 25
14 if DataSpO2<16.1465 then node 26 else node 27
15 if ChangesFEV1<1.1895 then node 28 else node 29
16 class = Home management
17 class = Referral
18 class = Referral
19 class = Home management
20 class = Home management
21 if DataFEV6<9.48762 then node 30 else node 31
22 class = Referral
23 if DataTi2Tot<103.25 then node 32 else node 33
24 class = Referral
25 if DataSystolic<2.73252 then node 34 else node 35
26 class = Referral
27 class = Home management
28 class = Referral
29 class = Home management
30 class = Home management
31 if MeanHR<81.629 then node 36 else node 37
32 class = Referral
33 if StdWeight<0.298959 then node 38 else node 39
34 class = Home management
35 class = Referral
36 if zScoreVE<2.13726 then node 40 else node 41
37 class = Home management
38 class = Referral
39 if DataIC<6.33849 then node 42 else node 43
40 class = Referral
41 class = Home management
42 if MeanVT<0.313994 then node 44 else node 45
43 class = Referral
44 class = Referral
45 if StdIC<1.41457 then node 46 else node 47
46 if zScoreTemp<2.77685 then node 48 else node 49
47 class = Home management
48 class = Home management
49 if DataHR<59 then node 50 else node 51
50 class = Home management
51 class = Referral

**Rules set 9**

1 if MeanPEF<98.0073 then node 2 else node 3
2 if StdFVC<0.120727 then node 4 else node 5
3 if MeanBR<2.43713 then node 6 else node 7
4 if zScoreWeight<1.01245 then node 8 else node 9
5 if StdFVC<0.156625 then node 10 else node 11
6 if MeanTi2Tot<146.887 then node 12 else node 13
7 if StdSpO2<1.55537 then node 14 else node 15
8 class = Referral
if ChangesDiastolic<-6.90808 then node 16 else node 17
if MeanVT<0.406996 then node 18 else node 19
if MeanHR<68.0968 then node 20 else node 21
class = Home management
if DataSpO2<16.1465 then node 22 else node 23
if ChangesFEV1<1.1895 then node 24 else node 25
class = Home management
class = Referral
if StdFEV1<0.0501056 then node 26 else node 27
if DataFEV1<0.694648 then node 28 else node 29
if DataIC<6.33849 then node 30 else node 31
class = Referral
class = Home management
class = Referral
if MeanHR<43.3387 then node 32 else node 33
if StdWeight<0.298959 then node 34 else node 35
class = Referral
if ChangesVT<16.9891 then node 36 else node 37
if DataFVC<0.969576 then node 38 else node 39
if StdHR<14.8032 then node 40 else node 41
if DataIC<2.28922 then node 42 else node 43
class = Home management
class = Home management
class = Referral
if StdSpO2<1.49389 then node 44 else node 45
class = Referral
class = Home management

Rules set 10
1 if MeanPEF<98.0073 then node 2 else node 3
2 if StdTemp<0.085169 then node 4 else node 5
3 if MeanBR<2.43713 then node 6 else node 7
4 if ChangesSpO2<-9.35194 then node 8 else node 9
5 if StdSpO2<1.19377 then node 10 else node 11
6 if MeanWeight<76.421 then node 12 else node 13
7 if StdSpO2<1.55537 then node 14 else node 15
8 class = Home management
9 if DataFEV1<1.32509 then node 16 else node 17
10 class = Home management
11 if StdDiastolic<8.42177 then node 18 else node 19
12 if MeanSystolic<2.90749 then node 20 else node 21
13 if ChangesVT<-5.91613 then node 22 else node 23
14 if DataSpO2<16.1465 then node 24 else node 25
15 if ChangesFEV1<1.1895 then node 26 else node 27
16 class = Referral
17 class = Home management
18 if MeanFEV6<10.9587 then node 28 else node 29
19 if MeanHR<81.8548 then node 30 else node 31
20 if MeanHR<43.3387 then node 32 else node 33
21 class = Referral
22 class = Referral
23 if DataSystolic<2.73252 then node 34 else node 35
24 class = Referral
25 class = Home management
26 class = Referral
27 class = Home management
28 class = Home management
29 class = Referral
30 if zScoreSystolic<-3.06685 then node 36 else node 37
31 class = Home management
32 class = Referral
33 if DataIC<6.33849 then node 38 else node 39
34 class = Home management
35 class = Referral
36 class = Home management
37 class = Referral
38 if StdWeight<0.298959 then node 40 else node 41
39 class = Referral
40 class = Referral
41 if MeanVT<0.313994 then node 42 else node 43
42 class = Referral
43 if StdIC<1.41457 then node 44 else node 45
44 if zScoreTemp<2.77685 then node 46 else node 47
45 class = Home management
46 class = Home management
47 if DataHR<59 then node 48 else node 49
48 class = Home management
49 class = Referral

Rules set 11

1 if MeanPEF<97.9755 then node 2 else node 3
2 if StdFVC<0.120727 then node 4 else node 5
3 if MeanDiastolic<67.355 then node 6 else node 7
4 if ChangesSpO2<-9.35194 then node 8 else node 9
5 if StdFVC<0.156625 then node 10 else node 11
6 class = Home management
7 if MeanBR<2.43713 then node 12 else node 13
8 class = Home management
9 class = Referral
10 class = Home management
11 if MeanSpO2<12.0842 then node 14 else node 15
12 if MeanHR<43.3387 then node 16 else node 17
13 if StdSpO2<1.55537 then node 18 else node 19
14 class = Home management
15 if DataFEV6<9.48762 then node 20 else node 21
16 class = Referral
if ChangesFEV1<-22.6247 then node 22 else node 23
if DataSpO2<16.1465 then node 24 else node 25
if ChangesFEV1<1.1895 then node 26 else node 27
      class = Home management
if MeanHR<81.629 then node 28 else node 29
      class = Referral
if DataIC<6.33849 then node 30 else node 31
      class = Referral
class = Home management
if MeanHR<81.629 then node 28 else node 29
      class = Referral
if DataIC<6.33849 then node 30 else node 31
      class = Referral
class = Home management
if MeanFEV1<0.865317 then node 32 else node 33
      class = Home management
if StdWeight<0.298959 then node 34 else node 35
      class = Referral
if MeanVT<0.406996 then node 16 else node 17
      class = Home management
if MeanBR<2.44317 then node 12 else node 13
      class = Referral
if ChangesDiastolic<-6.90808 then node 14 else node 15
      class = Home management
if MeanVT<0.406996 then node 16 else node 17
if MeanWeight<76.421 then node 18 else node 19
if ChangeFEV1<0.0541449 then node 20 else node 21
      class = Home management
      class = Referral
if StdFEV1<0.0501056 then node 22 else node 23
if DataFEV1<0.694648 then node 24 else node 25
if MeanHR<43.3387 then node 26 else node 27

Rules set 12

1  if MeanPEF<98.0073 then node 2 else node 3
2  if StdFVC<0.120727 then node 4 else node 5
3  if ChangesBR<-17.5568 then node 6 else node 7
4  if zScoreWeight<1.01245 then node 8 else node 9
5  if StdFVC<0.156625 then node 10 else node 11
      class = Referral
7  if MeanBR<2.44317 then node 12 else node 13
      class = Referral
9  if ChangesDiastolic<-6.90808 then node 14 else node 15
      class = Home management
11  if MeanVT<0.406996 then node 16 else node 17
12  if MeanWeight<76.421 then node 18 else node 19
13  if StdFEV1<0.0541449 then node 20 else node 21
      class = Home management
      class = Referral
16  if StdFEV1<0.0501056 then node 22 else node 23
17  if DataFEV1<0.694648 then node 24 else node 25
18  if MeanHR<43.3387 then node 26 else node 27
if ChangesVT<-5.91613 then node 28 else node 29
class = Referral
if DataFEV1<1.19971 then node 30 else node 31
class = Home management
class = Referral
class = Home management
class = Referral
if MeanFEV1<1.20037 then node 32 else node 33
class = Referral
if DataSystolic<2.73252 then node 34 else node 35
class = Home management
class = Referral
if DataIC<6.33849 then node 36 else node 37
class = Referral
class = Home management
class = Referral
if StdWeight<0.298959 then node 38 else node 39
class = Referral
class = Referral
if MeanVT<0.313994 then node 40 else node 41
class = Referral
if StdIC<1.41457 then node 42 else node 43
class = Home management
class = Home management
if DataHR<59 then node 46 else node 47
class = Home management
class = Referral
Appendix D: Chapter seven 14 sets of rules obtained from the 14 trained CARTs using data from Database II.

Rules set 1

1  if MeanDiastolic<71.0161 then node 2 else node 3
2  if zScoreWeight<-1.53059 then node 4 else node 5
3  if SConsQ<0.5 then node 6 else node 7
4   class = High risk
5  if ChangesWeight<-0.39354 then node 8 else node 9
6   class = Low risk
7  if SConsQ<1.5 then node 10 else node 11
8  if ChangesSpO2<-3.83943 then node 12 else node 13
9  if ChangesSystolic<-4.82863 then node 14 else node 15
10 if StdDiastolic<17.0654 then node 16 else node 17
11 if ChangesWeight<0.0239398 then node 18 else node 19
12   class = High risk
13   class = Low risk
14   class = High risk
15  if zScoreWeight<1.08049 then node 20 else node 21
16   class = High risk
17   class = Low risk
18   class = High risk
19  if ChangesWeight<0.960069 then node 22 else node 23
20   class = Low risk
21  if MeanWeight<57.7266 then node 24 else node 25
22   class = Low risk
23  if SColorQ<2.5 then node 26 else node 27
24   class = High risk
25  if DataWeight<62.45 then node 28 else node 29
26   class = High risk
27  if StdTemp<0.431263 then node 30 else node 31
28   class = Low risk
29  if DataWeight<86.075 then node 32 else node 33
30   class = High risk
31   class = Low risk
32   class = High risk
33  if DataSystolic<139.5 then node 34 else node 35
34   class = Low risk
35   class = High risk

Rules set 2

1  if MeanDiastolic<71.0161 then node 2 else node 3
2  if DataTemp<36.95 then node 4 else node 5
3  if zScoreWeight<0.105577 then node 6 else node 7
4  if ChangesSystolic<16.7219 then node 8 else node 9
5   class = High risk
6   class = High risk
7  if ChangesWeight<0.42063 then node 10 else node 11
8  if DataTemp<35.425 then node 12 else node 13
9   class = High risk
10  class = Low risk
11  if DataWeight<56 then node 14 else node 15
12  class = High risk
13  if StdSystolic<7.19256 then node 16 else node 17
14  class = Low risk
15  if StdWeight<0.420408 then node 18 else node 19
16  class = High risk
17  if StdWeight<1.59289 then node 20 else node 21
18  class = Low risk
19  if SColorQ<2.75 then node 22 else node 23
20  if zScoreSpO2<-1.60028 then node 24 else node 25
21  if StdWeight<1.81998 then node 26 else node 27
22  class = High risk
23  if MeanSystolic<135.081 then node 28 else node 29
24  class = High risk
25  class = Low risk
26  class = High risk
27  if StdSystolic<7.90924 then node 30 else node 31
28  class = High risk
29  class = Low risk
30  class = High risk
31  if DataSystolic<139.5 then node 32 else node 33
32  class = Low risk
33  class = High risk

Rules set 3

1  if zScoreWeight<-1.53059 then node 2 else node 3
2  class = High risk
3  if MeanDiastolic<71.0161 then node 4 else node 5
4  if ChangesWeight<-0.39354 then node 6 else node 7
5  if MeanSystolic<135.081 then node 8 else node 9
6  if ChangesSpO2<-3.83943 then node 10 else node 11
7  if ChangesSystolic<-4.82863 then node 12 else node 13
8  class = High risk
9  if DataWeight<55.425 then node 14 else node 15
10  class = High risk
11  class = Low risk
12  class = High risk
13  if zScoreWeight<1.08049 then node 16 else node 17
14  class = High risk
15  if DataWeight<96.725 then node 18 else node 19
16  class = Low risk
17  if MeanWeight<57.7266 then node 20 else node 21
18  class = Low risk
19  if DataWeight<97.725 then node 22 else node 23
20  class = High risk
21  if DataWeight<62.45 then node 24 else node 25
22  class = High risk
23  if SAmountQ<1 then node 26 else node 27
24  class = Low risk
25  if DataWeight<86.075 then node 28 else node 29
26  class = Low risk
27  class = High risk

207
class = High risk
if ChangesSystolic<2.53829 then node 30 else node 31
class = Low risk
class = High risk

Rules set 4

1 if MeanDiastolic<70.4574 then node 2 else node 3
2 if zScoreHR<0.90898 then node 4 else node 5
3 if zScoreWeight<0.105577 then node 6 else node 7
4 if ChangesSystolic<-9.12919 then node 8 else node 9
5 if StdWeight<0.83806 then node 10 else node 11
class = High risk
6 if ChangesWeight<1.45621 then node 12 else node 13
class = High risk
7 if zScoreWeight<-1.57911 then node 14 else node 15
class = Low risk
8 if SAmountQ<0.5 then node 16 else node 17
class = Low risk
9 if CoughQ<1.5 then node 18 else node 19
class = High risk
10 if DataDiastolic<57 then node 20 else node 21
class = Low risk
11 if CoughQ<1.5 then node 22 else node 23
class = Low risk

Rules set 5

1 if CoughQ<1.5 then node 2 else node 3
2 if DataWeight<54.25 then node 4 else node 5
3 if MeanDiastolic<72.1613 then node 6 else node 7
class = High risk
4 if ChangesSpO2<0.35432 then node 8 else node 9
5 if zScoreSpO2<-1.35071 then node 10 else node 11
6 if StdSpO2<3.41165 then node 12 else node 13
class = Low risk
7 if zScoreWeight<-1.53059 then node 14 else node 15
class = High risk
8 if StdWeight<0.167589 then node 16 else node 17
class = Low risk
9 if DataWeight<93.85 then node 18 else node 19
class = High risk
10 if zScoreWeight<0.893519 then node 20 else node 21
class = High risk
11 if StdHR<14.965 then node 22 else node 23
class = High risk
12 if CoughQ<1.5 then node 24 else node 25
class = Low risk
20  class = Low risk
21  if StdDiastolic<5.29131 then node 24 else node 25
22  if DataDiastolic<57 then node 26 else node 27
23  class = Low risk
24  class = Low risk
25  class = High risk
26  class = Low risk
27  class = High risk

Rules set 6

1  if MeanDiastolic<71.0161 then node 2 else node 3
2  if StdDiastolic<7.82856 then node 4 else node 5
3  if zScoreWeight<0.105577 then node 6 else node 7
4  if ChangesTemp<0.74312 then node 8 else node 9
5  if ChangesSpO2<1.27003 then node 10 else node 11
6  class = High risk
7  if ChangesWeight<0.42063 then node 12 else node 13
8  if MeanWeight<57.0506 then node 14 else node 15
9  class = High risk
10  if ChangesWeight<-0.410409 then node 16 else node 17
11  class = High risk
12  class = Low risk
13  if DataWeight<56 then node 18 else node 19
14  class = High risk
15  class = Low risk
16  class = Low risk
17  if DataWeight<64.375 then node 20 else node 21
18  class = Low risk
19  if StdWeight<0.420408 then node 22 else node 23
20  class = Low risk
21  class = High risk
22  class = Low risk
23  if SColorQ<2.75 then node 24 else node 25
24  class = High risk
25  if MeanSystolic<135.081 then node 26 else node 27
26  class = High risk
27  class = Low risk

Rules set 7

1  if MeanDiastolic<71.0161 then node 2 else node 3
2  if zScoreWeight<-1.53059 then node 4 else node 5
3  if StdDiastolic<17.0654 then node 6 else node 7
4  class = High risk
5  if ChangesWeight<-0.39354 then node 8 else node 9
6  if SColorQ<3.5 then node 10 else node 11
7  class = Low risk
8  if ChangesSpO2<-3.83943 then node 12 else node 13
9  if ChangesSystolic<-4.82863 then node 14 else node 15
10  if DataDiastolic<-56.5 then node 16 else node 17
11  class = High risk

209
Rules set 8

1 if MeanDiastolic<71.0161 then node 2 else node 3
2 if zScoreWeight<-1.53059 then node 4 else node 5
3 if MeanSystolic<135.081 then node 6 else node 7
4 class = High risk
5 if ChangesWeight<-0.39354 then node 8 else node 9
6 class = High risk
7 if ChangesWeight<0.0346369 then node 10 else node 11
8 if ChangesSpO2<-3.83943 then node 12 else node 13
9 if ChangesSystolic<-4.82863 then node 14 else node 15
class = High risk
11 if zScoreSpO2<-0.577707 then node 16 else node 17
class = High risk
13 class = Low risk
14 class = Low risk
15 if zScoreWeight<1.08049 then node 18 else node 19
class = High risk
19 if MeanWeight<57.7266 then node 20 else node 21
class = High risk
21 if DataWeight<62.45 then node 22 else node 23
class = Low risk
23 if DataWeight<86.075 then node 24 else node 25
class = High risk
25 if DataSystolic<139.5 then node 26 else node 27
class = Low risk
27 class = High risk

Rules set 9

1 if MeanDiastolic<71.0161 then node 2 else node 3
2 if zScoreWeight<-1.53059 then node 4 else node 5
3 if zScoreWeight<0.105577 then node 6 else node 7
class = High risk
4 if ChangesWeight<1.03792 then node 8 else node 9
class = High risk
5 if ChangesWeight<0.42063 then node 10 else node 11
8 if zScoreSpO2<-1.26716 then node 12 else node 13
9 if ChangesWeight<1.64813 then node 14 else node 15
class = Low risk
11 if DataWeight<56 then node 16 else node 17
14 class = Low risk
class = High risk
if ChangesSpO2<4.50942 then node 18 else node 19
class = High risk
if ChangesWeight<1.90386 then node 20 else node 21
class = Low risk
if StdWeight<0.420408 then node 22 else node 23
class = Low risk
class = High risk
class = Low risk
if DataWeight<62.525 then node 24 else node 25
class = Low risk
if SColorQ<2.75 then node 26 else node 27
class = Low risk
if CoughQ<1.5 then node 28 else node 29
class = High risk
if MeanSystolic<135.081 then node 30 else node 31
class = Low risk
if StdSpO2<1.17587 then node 32 else node 33
class = Low risk
class = High risk
class = Low risk
class = High risk
if CoughQ<1.5 then node 2 else node 3
if ChangesSpO2<0.35432 then node 4 else node 5
class = Low risk
if MeanDiastolic<70.4574 then node 6 else node 7
class = High risk
if zScoreSpO2<-1.35071 then node 8 else node 9
class = Low risk
if StdDiastolic<17.0654 then node 10 else node 11
class = High risk
if SColorQ<3.5 then node 14 else node 15
class = Low risk
if BreathQ<2.75 then node 16 else node 17
class = High risk
if DataWeight<39.325 then node 18 else node 19
class = High risk
if DataDiastolic<56.5 then node 20 else node 21
class = High risk
if SColorQ<2.5 then node 22 else node 23
class = High risk
if MeanWeight<98.5597 then node 24 else node 25
class = High risk
if SColorQ<0.5 then node 26 else node 27
class = Low risk
if DataHR<110.183 then node 28 else node 29
class = Low risk

Rules set 10

1 if CoughQ<1.5 then node 2 else node 3
2 if ChangesSpO2<0.35432 then node 4 else node 5
3 if MeanDiastolic<70.4574 then node 6 else node 7
4 class = Low risk
5 class = High risk
6 if zScoreSpO2<-1.35071 then node 8 else node 9
7 if StdDiastolic<17.0654 then node 10 else node 11
8 class = High risk
9 if MeanHR<85.6686 then node 12 else node 13
10 if SColorQ<3.5 then node 14 else node 15
11 class = Low risk
12 if BreathQ<2.75 then node 16 else node 17
13 if DataWeight<39.325 then node 18 else node 19
14 if DataDiastolic<56.5 then node 20 else node 21
15 class = High risk
16 if SColorQ<2.5 then node 22 else node 23
17 if MeanWeight<98.5597 then node 24 else node 25
18 class = High risk
19 if SColorQ<0.5 then node 26 else node 27
20 class = Low risk
21 class = High risk
22 class = Low risk
23 class = High risk
24 class = High risk
25 class = Low risk
26 class = High risk
27 if DataHR<110.183 then node 28 else node 29
28 class = Low risk
class = High risk

Rules set 11

1 if zScoreWeight<-1.53059 then node 2 else node 3
2 class = High risk
3 if MeanDiastolic<73.873 then node 4 else node 5
4 if zScoreWeight<0.893519 then node 6 else node 7
5 if StdDiastolic<17.0654 then node 8 else node 9
6 if zScoreSpO2<-1.26716 then node 10 else node 11
7 if MeanWeight<59.2581 then node 12 else node 13
8 if StdWeight<0.461419 then node 14 else node 15
9 class = Low risk
10 class = High risk
11 if ChangesSpO2<7.52917 then node 16 else node 17
12 class = Low risk
13 if StdWeight<1.26121 then node 18 else node 19
14 class = Low risk
15 if MeanSpO2<97.0442 then node 20 else node 21
16 class = Low risk
17 class = High risk
18 class = High risk
19 if DataWeight<86.075 then node 22 else node 23
20 if DataDiastolic<57 then node 24 else node 25
21 class = Low risk
22 class = High risk
23 if CoughQ<1.5 then node 26 else node 27
24 class = Low risk
25 class = High risk
26 if ChangesSpO2<0.35432 then node 28 else node 29
27 if StdHR<6.6543 then node 30 else node 31
28 class = Low risk
29 class = High risk
30 class = Low risk
31 class = High risk

Rules set 12

1 if MeanDiastolic<71.0161 then node 2 else node 3
2 if ChangesSystolic<-3.31788 then node 4 else node 5
3 if ChangesWeight<0.0346369 then node 6 else node 7
4 if ChangesHR<8.89536 then node 8 else node 9
5 if zScoreWeight<-1.53059 then node 10 else node 11
6 class = High risk
7 if ChangesWeight<0.42063 then node 12 else node 13
8 class = High risk
9 class = Low risk
10 class = High risk
11 if ChangesWeight<1.85263 then node 14 else node 15
12 class = Low risk
13 if DataWeight<56 then node 16 else node 17
14 if StdHR<6.08989 then node 18 else node 19
15 if StdWeight<0.690505 then node 20 else node 21
16 class = Low risk
17 if StdWeight<0.420408 then node 22 else node 23
18 class = High risk
19 class = Low risk
20 class = Low risk
21 if SColorQ<0.5 then node 24 else node 25
22 class = Low risk
23 if SColorQ<2.75 then node 26 else node 27
24 class = Low risk
25 if ChangesTemp<-0.973033 then node 28 else node 29
26 class = High risk
27 if MeanSystolic<135.081 then node 30 else node 31
28 class = High risk
29 class = Low risk
30 class = High risk
31 class = Low risk

Rules set 13

1 if MeanDiastolic<71.0161 then node 2 else node 3
2 if ChangesSystolic<-5.84241 then node 4 else node 5
3 if zScoreWeight<0.105577 then node 6 else node 7
4 if zScoreHR<1.39742 then node 8 else node 9
5 if zScoreHR<0.891187 then node 10 else node 11
6 class = High risk
7 if ChangesWeight<0.42063 then node 12 else node 13
8 class = High risk
9 class = Low risk
10 if zScoreWeight<-1.53059 then node 14 else node 15
11 if MeanWeight<40.8906 then node 16 else node 17
12 class = Low risk
13 if DataWeight<56 then node 18 else node 19
14 class = High risk
15 class = Low risk
16 class = Low risk
17 if StdWeight<0.167589 then node 20 else node 21
18 class = Low risk
19 if StdWeight<0.420408 then node 22 else node 23
20 class = Low risk
21 if SAmountQ<0.5 then node 24 else node 25
22 class = Low risk
23 if SColorQ<2.75 then node 26 else node 27
24 class = High risk
25 if DataHR<107.841 then node 28 else node 29
26 class = High risk
27 if MeanSystolic<135.081 then node 30 else node 31
28 class = High risk
29 class = Low risk
30 class = High risk
31 class = Low risk

213
Rules set 14

1 if MeanDiastolic<71.0161 then node 2 else node 3
2 if zScoreWeight<-1.53059 then node 4 else node 5
3 if zScoreWeight<0.105577 then node 6 else node 7
4 class = High risk
5 if zScoreWeight<0.792941 then node 8 else node 9
6 class = High risk
7 if ChangesWeight<0.42063 then node 10 else node 11
8 if zScoreSpO2<-1.26716 then node 12 else node 13
9 if DataWeight<62.525 then node 14 else node 15
10 class = Low risk
11 if DataWeight<56 then node 16 else node 17
12 class = High risk
13 if ChangesSpO2<4.50942 then node 18 else node 19
14 class = Low risk
15 if StdWeight<0.167589 then node 20 else node 21
16 class = Low risk
17 if StdWeight<0.420408 then node 22 else node 23
18 class = Low risk
19 class = High risk
20 class = Low risk
21 if StdWeight<1.26121 then node 24 else node 25
22 class = Low risk
23 if SColorQ<2.75 then node 26 else node 27
24 class = High risk
25 if DataWeight<86.075 then node 28 else node 29
26 class = High risk
27 if MeanSystolic<135.081 then node 30 else node 31
28 class = High risk
29 if ChangesSpO2<1.44617 then node 32 else node 33
30 class = High risk
31 class = Low risk
32 if DataHR<114.996 then node 34 else node 35
33 class = High risk
34 class = Low risk
35 class = High risk