Identifying the Ideal Cardiopulmonary Exercise Test Variable to distinguish between the Cardiovascular and Respiratory Components of Functional Limitation and to detect Relevant Physiological Changes in Function

Dr Anthony Barron

MBBS BSc MRCP

A thesis submitted to Imperial College London

For the degree of

PhD

Cardiovascular physiology research group
International Centre for Circulatory Health
St Mary’s Hospital
Imperial College of Science, Technology and Medicine
59-61 North Wharf Road
London
W2 1LA

Submitted May 2013
Declaration of Authenticity

I hereby declare that the work presented within this thesis is my own work.

Dr Anthony Barron

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.
Abstract

Cardiopulmonary exercise testing (CPX) is established for the investigation of cardiac disease. In patients with heart failure multiple variables have shown prognostic benefit, although peak $\dot{V}_{O_2}$ remains the most widely used. It is accepted that peak $\dot{V}_{O_2}$ is affected by respiratory disease as well, and may be highly susceptible to influence from respiratory disease coexistent with cardiac disease. I propose an alternative variable will show significantly greater specificity for cardiovascular disease when compared to peak $\dot{V}_{O_2}$ (and other variables) and aim to identify this “ideal” variable through the investigation of patients undergoing isolated cardiac interventions.

Patients were recruited to the following groups: undergone/going cardiac resynchronisation therapy (CRT); heart failure; mitral valve surgery; COPD or mixed disease. Each patient underwent echocardiography, pulmonary function tests and blood sampling. They then performed (after an initial familiarisation CPX) a baseline CPX. In patients undergoing intervention (CRT and mitral surgery) they underwent another CPX 2-3 months after intervention, and a further CPX at 6 months in the mitral valve group.

I assessed the following characteristics to aid in finding an ideal variable: ability to discriminate between heart and lung disease; high reproducibility; relation to exercise capacity and disease severity; and appropriate changes with physiological interventions.

ROC curve analysis showed that breathing reserve, OUES and double product had the greatest areas under curve when differentiating COPD from heart failure. OUES also showed excellent test-retest reproducibility and was strongly correlated to disease severity. Following mitral surgery OUES fell less at 2 months than peak $\dot{V}_{O_2}$; OUES may therefore be influenced to a lesser degree by muscular maladaptation post-surgery.

The ideal cardiopulmonary exercise test variable for cardiac patients has yet to be described. OUES appears to show discriminating properties between heart and lung, is reproducible and is less influenced by peripheral changes when compared to peak $\dot{V}_{O_2}$. 
Dedication

I dedicate this thesis to my partner Christine for her selfless support and understanding during the PhD, and to my parents, Susan and Brian, without whom none of my career would have been possible.
Acknowledgements

There are so many people that have helped shape not only this thesis, but the way in which a culture of thinking and endeavour is nurtured at the International Centre for Circulatory Health that there wouldn’t be space to thank them all. The ICCH is a wonderfully unique research centre with incredible minds willing to share advice at any time. I can imagine nowhere in academia more welcoming. The following deserve special mention however.

First and foremost I would like to thank my supervisors:

Professor Darrel Francis, for identifying my potential a number of years ago, persevering with a vision to demystify much of the exercise physiology world, and continually striving to prove, not just believe. His approach and dedication to research has been constantly inspirational.

Dr Roland Wensel, there is no-one I know who has a greater understanding of exercise physiology. Our discussions relating to this fascinating field have consistently fuelled my enthusiasm and excitement for this thesis and beyond.

Professor Jamil Mayet, for allowing me the opportunity to undertake research here at ICCH, his constant encouragement both academically and clinically, and for the team he has assembled at ICCH, without which none of this would be possible.

Professor Alun Hughes, for his depthless knowledge of statistics and interest in my research. I am sure I caused him endless frustration covering the same statistical concepts time and again, but the knowledge and skills I take away will be priceless.

I would also like to offer sincere thanks to the physiologists Joe Carlton, Siana Jones, Suzanne Williams and especially Tom Gethin-Jones. Between them they have assisted me on more than 300 exercise tests, and Tom was ever present to ensure that CRT devices were appropriately set.

Dr Niti Dhutia, who undertook the thankless task of writing the incomprehensible (to me) Matlab code required within Chapter 4.

Dr Sarah Elkin, for helping me to recruit patients with COPD and teaching me so much about lung function and respiratory disease.
Mr Roberto Casula, Mr Prakash Punjabi and the surgical coordinating team, for allowing me to recruit their patients prior to mitral surgery.

Members of the cath lab and respiratory function lab staff, for helping with recruitment and for always being able to squeeze another lung function test in.

Dr Luke Howard and Kevin Murphy, for helping me take my first real steps in the world of exercise physiology.

The British Heart Foundation for supporting this research study and supplying excellent suggestions to help me improve the concepts and methods.

And finally to the patients and healthy controls who volunteered for this study. The thought of exercising until your maximum is not a pleasant one for many people, and to volunteer to do this, despite a significant chronic illness or before major surgery, in the name of science is truly humbling.
## Contents

Declaration of Authenticity ........................................................................................................ 2  
Abstract .................................................................................................................................. 3  
Dedication ............................................................................................................................... 4  
Acknowledgements .................................................................................................................. 5  
Contents .................................................................................................................................... 7  
List of Figures .......................................................................................................................... 15  
List of Tables ............................................................................................................................ 18  
Abbreviations ............................................................................................................................ 20  

### 1.0 Background

1.1 Heart failure and exercise intolerance .............................................................................. 23  
1.2 Skeletal muscle in CHF and the link between cardiac function and exercise capacity .................................................. 26  
1.3 The interaction of heart and lungs in CHF ....................................................................... 30  
1.4 Lung disease is common in CHF ..................................................................................... 31  
1.5 The respiratory patient with exercise intolerance or dyspnoea ................................... 32  
1.6 The muscle in COPD ....................................................................................................... 33  
1.7 Cardiopulmonary exercise testing, the requirement for oxygen and the role of the heart, lungs, vasculature and muscles in performing exercise ................................................. 34  
1.8 Common CPX Variables in the diagnosis and assessment of dyspnoea .................................................. 38  
   1.8.1 Peak $\dot{V}_O_2/\dot{V}_O_2$ Max ...................................................................................... 39  
   1.8.2 Anaerobic/ Lactate threshold .................................................................................... 40  
   1.8.3 Ventilatory equivalents ............................................................................................ 40  
   1.8.4 Breathing reserve .................................................................................................... 42  
   1.8.5 Oxygen efficiency .................................................................................................... 42  
   1.8.6 Oxygen Uptake: Work Rate relationship .................................................................. 43  
   1.8.7 An algorithm approach ............................................................................................ 45  
1.9 Aims ..................................................................................................................................... 47

### 2.0 Methods and Materials

2.1 Overview ............................................................................................................................. 50
2.2 Patient Identification and Recruitment ........................................ 50
  2.2.1 Mitral valve disease ............................................................... 50
  2.2.2 Heart failure ................................................................. 53
  2.2.3 COPD .......................................................... 53
  2.2.4 Inclusion and exclusion criteria ........................................... 54

2.3 Equipment and Testing ............................................................... 56
  2.3.1 Cardiopulmonary exercise testing .......................................... 56
    2.3.1.1 Overview of exercise testing and terminology .................. 56
    2.3.1.2 Calibration ......................................................... 57
    2.3.1.3 Ergometer and test set-up ...................................... 60
    2.3.1.4 Testing ............................................................... 65
    2.3.1.5 Data analysis and key variables .................................. 67
  2.3.2 Echocardiography ............................................................... 81
    2.3.2.1 Views ................................................................. 82
    2.3.2.2 Two-dimensional measurements .................................. 84
    2.3.2.3 Standard Doppler measurements ............................... 86
    2.3.2.4 Tissue Doppler measurements (TDI) .......................... 88
    2.3.2.5 Specific measurements for mitral regurgitation ............. 90
    2.3.2.6 Specific measurements for mitral stenosis .................... 92
  2.3.3 Pulmonary function tests .................................................... 93
    2.3.3.1 Calibration .......................................................... 93
    2.3.3.2 Spirometry ......................................................... 93
    2.3.3.3 Full lung subdivisions .......................................... 93
    2.3.3.4 Diffusion capacity ................................................ 94
  2.3.4 Venous blood sampling ...................................................... 96

3.0 Retrospective Pilot analysis – the differential influence of spirometry on CPX variables
  3.1 Abstract .................................................................................... 98
  3.2 Introduction ............................................................................... 99
  3.3 Methods .................................................................................. 101
    3.3.1 Subjects ............................................................................ 101
    3.3.2 Cardiopulmonary exercise testing .................................... 101
3.3.3 Echocardiographic assessment of left ventricular function ................................................................. 101
3.3.4 Statistical analysis ............................................................................................................................ 102
3.4 Results ............................................................................................................................................. 103
  3.4.1 Patient characteristics .................................................................................................................... 103
  3.4.2 Relationship between exercise variables ......................................................................................... 105
  3.4.3 Effect of ventilatory measures on CPX variables .......................................................................... 105
  3.4.4 Impact of ventilatory measures on CPX variables in patients not limited by lungs .................. 107
  3.4.5 Effect of ejection fraction on CPX variables ................................................................................ 107
3.5 Discussion ....................................................................................................................................... 110
  3.5.1 The relation of CPX variables to one another .............................................................................. 110
  3.5.2 The impact of spirometric abnormalities on CPX ....................................................................... 111
  3.5.3 The impact of ejection fraction on CPX variables ....................................................................... 112
  3.5.4 Clinical implications of these results ......................................................................................... 112
  3.5.5 Implications for the thesis ........................................................................................................... 113
  3.5.6 Limitations .................................................................................................................................. 113
  3.5.7 Conclusions ............................................................................................................................... 114
4.0 Study of Health in Pomerania - Physiology of oxygen uptake kinetics
  4.1 Abstract ........................................................................................................................................... 117
  4.2 Introduction ..................................................................................................................................... 118
  4.3 Methods .......................................................................................................................................... 121
    4.3.1 Study population.......................................................................................................................... 121
    4.3.2 Pre-exercise diagnostics and exclusion criteria ........................................................................ 121
    4.3.3 Exercise testing ........................................................................................................................... 122
    4.3.4 CPX data analysis ....................................................................................................................... 123
    4.3.5 Test foreshortening .................................................................................................................... 123
    4.3.6 Statistical analysis ..................................................................................................................... 124
  4.4 Results ............................................................................................................................................ 125
    4.4.1 Baseline characteristics .............................................................................................................. 125
    4.4.2 Influence of demographics on CPX ......................................................................................... 125
    4.4.3 Influence of respiratory variables and smoking on CPX ...................................................... 134
    4.4.4 Influence of beta-blockade on CPX variables ......................................................................... 136
4.4.5 Reference distribution .............................................. 136
4.4.6 Comparison with previous predictive equations for OUES ...... 139
4.4.7 Impact of foreshortened exercise on OUES ...................... 139
4.4.8 Impact of foreshortened exercise on the $V_{O_2} - WR$ slope ...... 143

4.5 Discussion ...................................................................... 145
4.5.1 Study of Health in Pomerania ........................................ 145
4.5.2 Predictive equations ..................................................... 146
4.5.3 Oxygen uptake efficiency slope .................................... 147
4.5.4 Oxygen uptake to work rate relationship ......................... 148
4.5.5 Heart rate to oxygen uptake relationship ......................... 150
4.5.6 Limitations ............................................................... 151
4.5.7 Conclusions ............................................................. 152

5.0 Reproducibility/Repeatability

5.1 Abstract ......................................................................... 154
5.2 Introduction ..................................................................... 155
5.3 Methods .......................................................................... 160
5.3.1 Study 1 – Healthy controls ........................................... 160
5.3.2 Study 2 – The Observational cohort ............................... 161
5.3.3 Statistical analysis ....................................................... 161
5.4 Results ............................................................................ 164
5.4.1 Study 1 - Participant and test characteristics ...................... 164
5.4.2 Study 1 – Identification of intensity onset ......................... 164
5.4.3 Study 1 – Repeatability of 2 identical CPX tests ................. 167
5.4.4 Study 1 – Repeatability of CPX tests using different ramp protocols ......................................................... 170
5.4.5 Study 2 – Patient and test characteristics ......................... 172
5.4.6 Study 2 – Difference between test 1 and 2 ....................... 175
5.4.7 Study 2 – Test-retest reliability of various CPX variables ...... 178
5.4.8 Study 2 - Determinants of test-retest reliability .................. 179
5.4.9 Changes in slopes throughout testing .............................. 181
5.5 Discussion ...................................................................... 184
5.5.1 Test-retest reliability and the measures of reproducibility ............................................................. 184
5.5.2 The difference between 2 tests –
A familiarisation effect? ...................................................... 186

5.5.3 The effect of potential determinants on test-retest reliability ................................................................. 186

5.5.4 The effect of a change in ramp protocol on test-retest reliability ................................................................. 187

5.5.5 Limitations ................................................................. 189

5.5.6 Conclusions ................................................................. 190

6.0 The Observational Study

6.1 Abstract ................................................................................................................................. 192

6.2 Introduction .......................................................................................................................... 193

6.3 Methods ............................................................................................................................... 195
  6.3.1 Patient recruitment ........................................................................................................... 195
  6.3.2 CPX analysis ................................................................................................................... 195
  6.3.3 Hypotheses ..................................................................................................................... 196
  6.3.4 Statistical analysis .......................................................................................................... 196

6.4 Results .................................................................................................................................. 198
  6.4.1 Patient recruitment and characteristics ........................................................................... 198
  6.4.2 CPX results and influencing factors on exercise capacity ........... 201
  6.4.3 The influence of the lungs on CPX variables .............................................................. 201
  6.4.4 Differences in variables between groups ................................................................. 202
  6.4.5 ROC curve analysis ...................................................................................................... 206

6.5 Discussion ............................................................................................................................ 211
  6.5.1 Oxygen uptake efficiency slope ..................................................................................... 212
  6.5.2 Breathing reserve ........................................................................................................... 213
  6.5.3 Anaerobic threshold ........................................................................................................ 214
  6.5.4 Other variables ............................................................................................................... 215
  6.5.5 Clinical relevance .......................................................................................................... 217
  6.5.6 Limitations ..................................................................................................................... 218
  6.5.7 Conclusions ................................................................................................................... 218

7.0 Correlation of CPX variables to disease severity

7.1 Abstract ................................................................................................................................. 220

7.2 Introduction ........................................................................................................................ 221
7.3 Methods ................................................................. 222
  7.3.1 Patient recruitment ........................................... 222
  7.3.2 CPX analysis .................................................... 222
  7.3.3 Echocardiographic analysis ................................. 223
  7.3.4 Hypotheses ..................................................... 223
  7.3.5 Statistical analysis ........................................... 223
7.4 Results ................................................................. 225
  7.4.1 Patient characteristics ....................................... 225
  7.4.2 Relation to symptom class .................................. 225
  7.4.3 Correlation with markers of cardiac dysfunction .... 226
  7.4.4 Correlation with peak $\dot{V}_{O2}$ .......................... 230
7.5 Discussion ............................................................. 234
  7.5.1 NYHA classification ......................................... 234
  7.5.2 Objective measures of respiratory and cardiac dysfunction .... 235
  7.5.3 Correlation between peak $\dot{V}_{O2}$ and other cardiopulmonary exercise variables .......................................................... 236
  7.5.4 Variables relating to disease severity ...................... 237
  7.5.5 Limitations ....................................................... 238
  7.5.6 Conclusions ...................................................... 238

8.0 The Interventional Study – The Mitral Valve
8.1 Abstract .............................................................. 240
8.2 Introduction .......................................................... 241
8.3 Methods ............................................................... 243
  8.3.1 Patient recruitment ........................................... 243
  8.3.2 Patient testing preoperatively ............................... 243
  8.3.3 Patient testing postoperatively .............................. 244
  8.3.4 Echocardiography ............................................. 244
  8.3.5 Hypotheses ..................................................... 245
  8.3.6 Statistical analysis ........................................... 245
8.4 Results ................................................................. 247
  8.4.1 Patient recruitment and characteristics .................. 247
  8.4.2 Impact of surgery on respiratory function ............... 249
  8.4.3 Other changes postoperatively and echocardiographic findings .......................................................... 249
8.4.4 Change in CPX variables at 2 months ........................................ 251
8.4.5 Change in CPX variables at 6 months ........................................ 254
8.4.6 Influencing factors on postoperative CPX variables ................. 254
8.4.7 Changes in patients with reduced exercise capacity ............... 256

8.5 Discussion ................................................................................. 257
8.5.1 Reductions in exercise capacity pre- and postoperatively ...... 257
8.5.2 Changes in CPX variables ..................................................... 258
8.5.3 $\dot{V}_E/\dot{V}_{CO_2}$ relationship ................................................. 260
8.5.4 Effect of surgery on lung function ............................................ 260
8.5.5 Limitations ............................................................................. 261
8.5.6 Conclusions ........................................................................... 262

9.0 The Interventional Study – Cardiac Resynchronisation Therapy

9.1 Abstract .................................................................................... 264
9.2 Introduction .............................................................................. 265
9.3 Methods ................................................................................... 267
  9.3.1 Patient recruitment ............................................................... 267
  9.3.2 Patient testing ..................................................................... 267
  9.3.3 Hypotheses .......................................................................... 269
  9.3.4 Statistical analysis ................................................................. 269
9.4 Results ...................................................................................... 270
  9.4.1 Patient recruitment and characteristics ................................. 270
  9.4.2 Changes between pre- and post-implantation ...................... 272
  9.4.3 Comparing the 2 groups ....................................................... 272
  9.4.4 CPX results – The pre-post study ...................................... 275
  9.4.5 CPX results – The on-off study .......................................... 275
9.5 Discussion ................................................................................ 279
  9.5.1 Changes in ventilation .......................................................... 279
  9.5.2 Lack of improvement in exercise capacity .......................... 280
  9.5.3 Differences between variables ............................................ 282
  9.5.4 Limitations .......................................................................... 282
  9.5.5 Conclusions ......................................................................... 283
10.0 Synthesis

10.1 Overview .......................................................... 285
10.2 Respiratory exchange ratio and breathing reserve ......................... 286
10.3 Oxygen uptake efficiency slope ................................................. 287
10.4 Oxygen uptake at the anaerobic threshold ........................................ 289
10.5 Oxygen uptake to work rate slope .............................................. 290
10.6 $O_2$ Pulse and the heart rate to oxygen uptake slope .......................... 290
10.7 Ventilatory equivalents ................................................................... 291
10.8 Double product and circulatory power ............................................. 293
10.9 Limitations .................................................................................... 293
10.10 Conclusions ................................................................................ 295

11.0 Bibliography .............................................................................. 296

12.0 Appendix 1 – Data analysis within Excel 2007/2010
    12.1 Identification of AT and slopes ................................................... 309
    12.2 Measurement of OUES and foreshortened OUES ......................... 310
    12.3 Measurement of $\dot{V}_E/\dot{V}_{CO_2}$ slope up to the VCP and post-VCP ...... 311

13.0 Appendix 2
    13.1 Participant Information Leaflet – Valve disease .......................... 312
    13.2 Participant Information Leaflet - COPD ....................................... 314
    13.3 Consent form .............................................................................. 316
    13.4 Research Ethics Committee acceptance ....................................... 318
    13.5 Permission for use of figure ......................................................... 320

14.0 Publications .................................................................................. 321

15.0 Abstracts ...................................................................................... 322
List of Figures

Figure 1.1: A conceptualisation of the interaction between the heart, lungs and muscles in CHF 25
Figure 1.2: The intracellular production of energy, CO₂ and water from carbohydrate and O₂ 35
Figure 1.3: Minute ventilation and carbon dioxide elimination rise together during incremental exercise 41
Figure 1.4: Graphical representation of how OUES is calculated 44
Figure 1.5: Flow Chart for the differential diagnosis of exertional dyspnoea 46

Figure 2.1: European Society of Cardiology guidance 2012 for treatment of organic mitral regurgitation 52
Figure 2.2: Calibration of the Gas Analyzer Results Table 58
Figure 2.3: Calibration of turbine flow 59
Figure 2.4: Ergoselect bicycle ergometer 61
Figure 2.5: Software visualisation of spirometry 63
Figure 2.6: Positioning of ECG electrodes on the subject’s chest 64
Figure 2.7: How peak data was manually averaged in Excel 70
Figure 2.8: The identification of the \( \dot{V}_{O_2} \) at the anaerobic threshold 72
Figure 2.9: Identification of the anaerobic threshold using 3 supplementary plots 74
Figure 2.10: Identifying the Ventilatory compensation point (VCP) 75
Figure 2.11: How submaximal OUES limits were constructed 79
Figure 2.12: Measurement of ventricular volumes using the Simpson’s method 85
Figure 2.13: Transmitral filling pattern 87
Figure 2.14: Tissue Doppler imaging of the septal wall at the mitral annulus 89
Figure 2.15: The calculation of mitral regurgitation severity via the PISA method 91
Figure 2.16: The subdivisions of the lung 95

Figure 3.1: Correlation between the 4 CPX variables; peak \( \dot{V}_{O_2} \) (L/min), OUES (L/min/10-fold increase in ventilation), \( \dot{V}_E/\dot{V}_{CO_2} \) slope, and \( \dot{V}_{O_2} \) at AT (L/min) 106
Figure 3.2: Box and whisker plot of the impact of FEV₁ and ejection fraction on the four CPX variables 108
Figure 3.3: Box and whisker plot of the impact of spirometric abnormalities on the four CPX variables when patients are divided into normal, obstructive, restrictive or mixed pattern 109
Figure 4.1: The distribution of peak $\dot{V}_{O_2}$ and OUES with age fit a polynomial distribution

Figure 4.2: Reference ranges for OUES for males and females with relation to age and subdivided according to body mass index (BMI) groups

Figure 4.3: Reference ranges for OUES/kg for males and females with relation to age and subdivided according to body mass index (BMI) groups

Figure 4.4: Reference ranges for $\dot{V}_{O_2} - WR$ slope for males and females with relation to age and subdivided according to body mass index (BMI) groups

Figure 4.5: Reference ranges for $HR - \dot{V}_{O_2}$ slope for males and females with relation to age and subdivided according to body mass index (BMI) groups

Figure 4.6: Comparison of my new reference equations with 2 pre-existing reference ranges

Figure 4.7: The effect of data coping on OUES and peak $\dot{V}_{O_2}$

Figure 4.8A: Multiple mean $\dot{V}_{O_2} - WR$ slopes were calculated from all the participants within four groups. B: $\dot{V}_{O_2}$ was then plotted against work rate for all participants within these same four groups

Figure 5.1: Distribution of work rate in an incremental test where participant perception of increasing intensity is first noticed by the participant

Figure 5.2: Median and interquartile range for peak $\dot{V}_{O_2}$ (mL/min/kg) by test order

Figure 5.3: Proportion of each ramp protocol chosen for test 2 based on the results of test 1

Figure 5.4: Bland-Altman plots of 9 important CPX variables

Figure 5.5: Sparkline plots for the change in OUES, $\dot{V}_{E}/\dot{V}_{CO_2}$, and $\dot{V}_{O_2} - WR$ slope in 24 participants over time

Figure 6.1: Flow Chart for the differential diagnosis of exertional dyspnoea

Figure 6.2: Dotplots for 9 CPX variables in patients with COPD, CHF, mitral valve disease (MV) and mixed disease

Figure 6.3: Receiver Operator Characteristic (ROC) curves to differentiate between COPD and CHF

Figure 6.4: Diagnostic capabilities of OUES and BR cut-offs

Figure 7.1: Change in 4 CPX variables with symptom class

Figure 7.2: Correlation between tissue Doppler imaging, OUES and peak $\dot{V}_{O_2}$

Figure 7.3: Correlation between log BNP and both OUES and peak $\dot{V}_{O_2}$
Figure 7.4: Scatter plot of percent of predicted OUES against percent of predicted peak $\dot{V}_{O_2}$

Figure 8.1: Change in 12 variables between a preoperative CPX and a CPX performed, on average, 2 months postoperatively

Figure 8.2: Change in 12 variables between a preoperative CPX and a CPX performed, on average, 6 months postoperatively

Figure 9.1: Sparkline plots of the change of $\dot{V}_{O_2}$ over time for the 10 patients post-implantation

Figure 9.2: Change in 13 variables between a pre-implant CPX and a CPX performed on average 2 months post-implant

Figure 9.3: Change in 13 variables between a CPX performed with CRT enabled and another with CRT disabled
List of Tables

Table 3.1: Baseline characteristics.  
Table 4.1: Baseline characteristics of the 1202 participants undergoing CPX testing with measurement of oxygen kinetic variables  
Table 4.2: The influence of age, height, weight and BMI on the four variables of oxygen kinetics in both males and females  
Table 4.3: The influence of four respiratory variables on the four variables of oxygen kinetics in both males and females  
Table 4.4: Multivariate regression analysis adjusted for age, height and weight in males and females for the effect of taking beta-blockers on CPX variables  
Table 4.5: Predictive equations of reference ranges for 5th, 50th and 95th percentiles  
Table 4.6: Correlation coefficient values for comparison of each pair of foreshortened OUES readings, calculated for all data present for the range stated  
Table 5.1: Baseline characteristics for eight healthy adults undertaking four CPX tests  
Table 5.2: The mean value, mean difference, standard deviation, standard deviation of the difference and the bootstrapped intraclass correlation coefficient with confidence interval when comparing the two standard tests undertaken by 8 healthy volunteers  
Table 5.3: The mean difference, standard deviation of the difference: mean ratio and the bootstrapped intraclass correlation coefficient with confidence interval when comparing the first standard test to either the steep or shallow tests undertaken by 8 healthy volunteers  
Table 5.4: Patient characteristics overall and for each of the 3 disease categories, chronic obstructive pulmonary disease (COPD), chronic heart failure (CHF) and mitral valve disease (MV)  
Table 5.5: Test-Retest reliability of the full cohort  
Table 5.6: Test-Retest reliability measurements between disease categories for selected variables  
Table 6.1: Reference ranges for peak $V_{O2}$ (mL/min), $O_2$ pulse (mL/beat) and OUES (mL/min/10-fold increase in $V_{E}$) for males (M) and females (F) based on the SHIP data
Table 6.2: Patient characteristics

Table 6.3: CPX results for a number of principal variables for the whole cohort and then divided into disease categories

Table 6.4: Area under curve (AUC) for a number of CPX variables following 3 ROC curve analyses

Table 7.1: Association (by ANOVA) between symptom class (I-IV) and CPX variables in the full group, patients with COPD, and patients with mitral valve disease and CHF

Table 7.2: Correlation between peak $\dot{V}_{O_2}$ and other CPX variables

Table 8.1: Patient baseline and operative characteristics

Table 8.2: Comparison of lung function variables before and after mitral surgery

Table 8.3: CPX variables preoperatively and at an average of 2 and 6 months postoperatively in patients undergoing mitral valve surgery for severe mitral regurgitation (n=20) or stenosis (n=1)

Table 9.1: Patient baseline characteristics

Table 9.2: CPX results for the pre- and post-implant patients, and the CRT on-off patients
Frequently used Abbreviations

ANOVA = Oneway analysis of variance
AT = Anaerobic threshold
ATP/ADP = Adenosine tri-/diphosphate
AUC = Area under curve
BMI = Body mass index
BNP = B-type natriuretic peptide
BR = Breathing reserve
BSA = Body surface area
CHF = Chronic heart failure
COPD = Chronic obstructive pulmonary disease
CoV = Coefficient of variation
CPX = Cardiopulmonary exercise test
CRT = Cardiac resynchronisation therapy
DCM = Dilated cardiomyopathy
DLCO = Diffusion capacity of the lung for carbon monoxide
DP = Double product
EF = Ejection fraction
eGFR = Estimated glomerular filtration rate
FEV1 = Forced expiratory volume in 1 second
FS = Fractional shortening
FVC = Forced vital capacity
GOLD = Global initiative for chronic obstructive lung disease
ICC = Intraclass correlation coefficient
ICHNT = Imperial College Healthcare NHS Trust
IHD = Ischaemic heart disease
IQR = Interquartile range
KCO = Transfer coefficient for carbon monoxide
LAV = Left atrial volume
LOA = Limits of agreement
LV = Left ventricular
LVAD = Left ventricular assist device
MDRD = Modification of Diet in Renal Disease
MV/MR/MS = Mitral valve/regurgitation/ stenosis
MVV = Maximum voluntary ventilation
NYHA = New York Heart Association
OUES/OUEP = Oxygen uptake efficiency slope/plateau
PaCO2/ PaO2 = Arterial partial pressure of carbon dioxide/oxygen
PETCO2/ PETO2 = End-tidal partial pressure of carbon dioxide/oxygen
PISA = Proximal isovelocity surface area
RER/RQ = Respiratory exchange ratio/ Respiratory quotient
R_f = Respiratory frequency
ROC = Receiver operator characteristic
RV = Right ventricular
SDD = Standard deviation of the difference
SHIP = Study of health in Pomerania
TAPSE = Tricuspid annular plane systolic excursion
TDI = Tissue Doppler imaging/echocardiography
TLC = Total lung capacity
V_A = Alveolar volume
VCP = Ventilatory compensation point
\( \dot{V}_E \) = Minute ventilation
\( \dot{V}_{O_2}/\dot{V}_{CO_2} \) = Oxygen uptake/ carbon dioxide elimination
V_T = Tidal volume
VTI = Velocity time integral
WHO = World Health Organisation
WR = Work rate
1.0 Background
1.1 Heart Failure and Exercise Intolerance

Chronic heart failure (CHF) is a common condition typically caused by a weakness or stiffness of the heart muscle. It is a difficult disease entity to fully define; despite numerous attempts, all with their own merits, none have been able to encapsulate all points. The current definition used within the UK’s National Heart Failure Audit 2010-11 states:

“Heart failure is a complex syndrome that can result from any structural or functional cardiac disorder that impairs the ability of the heart to function as a pump to support the circulation in the normal physiological range. The syndrome of heart failure is characterised by symptoms such as breathlessness and fatigue, and signs such as fluid retention.”

CHF is correctly felt to be a syndrome, rather than a disease per se and is characterised by exercise intolerance and exertional dyspnoea. Aetiology of heart failure is varied; this heterogeneity includes ischaemic heart disease, cardiomyopathy, valvular disease, and arrhythmias. Heart failure is amongst the commonest chronic diseases in Westernised countries and within the UK affects more than 1% of the population (UK National Heart Failure Audit 2011). Heart disease overall accounts for more deaths than any other (UK Office of National Statistics, WHO website) with heart failure accounting for a large proportion of these deaths.

There are a number of pathophysiological abnormalities in heart failure. It is characterised by low cardiac output, or an inability to raise cardiac output appropriately during times of increased demand; an increase in intra-cardiac pressures, usually in an attempt to maintain cardiac output, at the expense of increased proximal pressures in vascular beds such as the lungs or peripheries leading to symptoms of oedema; and increased neurohormonal activation with systemic consequences that are as yet not fully understood but include increased sympathetic activation, activation of the renin-angiotension-aldosterone system, abnormal endothelin activity and a chronic inflammatory response.

CHF is not only common but leads to significant morbidity and mortality. Hospitalisations are frequent, long (median 9 days in the UK National Heart Failure Audit 2010-11) and associated with a high mortality rate (33% during or soon after admission). Exercise limitation is a frequent principal presenting complaint of patients admitted to hospital with CHF in the UK, but in the ambulant outpatient population with heart failure is likely to be even more common as the main symptom. It has a significant impact on quality of life.
The degree of exercise intolerance has been shown to be a key determinant of prognosis in an ambulant population with CHF in numerous studies using multiple measures of exercise capacity such as the New York Heart Association (NYHA) functional classification, 6 minute walk test and cardiopulmonary exercise testing. NYHA classification was established in 1928 with a significant re-clarification in 1964. It suffers from poor reproducibility (Goldman et al 1981, Raphael et al 2007) and is generally within studies a non-predictor or, at best, a weak univariate predictor of prognosis in CHF (Willens et al 1987, Van den Broek et al 1992, Cowie et al 2000, Hobbs et al 2007). Since the introduction of objective measures of exercise capacity it rarely maintains significance in most multivariate analyses (Likoff et al 1987, Cohn et al 1993). The 6 minute walk test (6MWT) is a simple clinical test, potentially less subjective than NYHA classification, and can be done with minimal equipment. However it is at best a weak indicator of prognosis (Roul et al 1998) with only the most limited patients (6 minute walk distance <300m) showing a correlation to prognosis. It does not appear to provide complementary prognostic data beyond formal exercise testing and NYHA classification (Opasich et al 2001, Sharma et al 2001). In contrast to NYHA classification and the 6WMT, formal exercise testing with gas analysis, known as cardiopulmonary exercise testing (CPX) has shown in dozens of studies strong univariate and multivariate power in determining the prognosis of patients with CHF. The extensive literature on CPX in CHF prognosis will be discussed later alongside description of commonly measured CPX variables. However, of clinical relevance to heart failure physicians, is that objective variables of exercise capacity, as measured by CPX, are key determinants for the identification of appropriate patients for heart transplantation and often supersede measures of cardiac morphological impairment.

However we define and measure it, exercise intolerance in CHF is common with undeniable importance to prognosis but the physiological abnormalities behind exercise intolerance in CHF remain poorly understood. It is not as simple as low cardiac output during exercise; the poor correlation of the haemodynamic disorder to exercise capacity has long been established (Wilson et al 1983, Szlachcic et al 1985, Cohn et al 1993, Wilson et al 1995). First and foremost, CHF must be seen as a multisystem disease, with common interactions between the heart and the lungs, skeletal muscle, peripheral vasculature, neurohormonal systems and the kidneys. The skeletal muscles and lungs deserve special mention in their interaction with the heart in patients with CHF. These 3 organ systems, which are stressed and measured during CPX, interact together, with typical abnormalities in all, rather than just the heart, in patients with CHF (Figure 1.1).
Figure 1.1: A conceptualisation of the interaction between the heart, lungs and muscles in CHF. The interaction of all 3 leads to reduced exercise capacity.
1.2 Skeletal muscle in CHF and the link between cardiac function and exercise capacity

There is substantial evidence that the degree of cardiac abnormality does not correlate strongly with exercise capacity. In 1981 Franciosa et al showed that, in patients with CHF, non-invasive measurements of cardiac function at rest did not correlate to total treadmill time. Following 5 weeks of randomisation to placebo or vasodilator therapy, patients increased their treadmill time by, on average, 1.5 minutes, without a significant change in measures of left ventricular (LV) function, and no correlation between the improvement in exercise capacity and degree of change of LV variables.

We find greater agreement between abnormalities of invasive haemodynamics and exercise capacity, but a large degree of unexplained heterogeneity still remains. In a study of 27 patients with CHF Szlachcic et al showed that of all resting invasive measurements, only capillary wedge pressure correlated to exercise capacity (as defined by peak oxygen consumption which will be discussed extensively later) (Szlachcic et al 1985). Of variables on exercise, only cardiac index (cardiac output indexed to body surface area (BSA)) and peak heart rate correlated to exercise capacity, and heart rate was the only variable where the change from rest to exercise correlated to exercise capacity. Values obtained during exercise, rest and the degree of change from rest to exercise for stroke volume (indexed for BSA), systemic and pulmonary vascular resistance, left ventricular stroke work (indexed for BSA) and left and right ejection fractions did not correlate to exercise capacity. Wilson et al showed that in patients with reduced exercise capacity awaiting heart transplant, there was no difference in exercise capacity when patients were grouped according to the degree of cardiac output response and/or elevated wedge pressure (Wilson et al 1995). In another study the initiation of vasodilators led to improvements in exercise capacity, and improvements in invasive haemodynamics, but these two improvements were not correlated (Massie et al 1981). Separately the use of dobutamine was able to substantially increase cardiac index in one study, but improvements in oxygen consumption were modest without improvements in treadmill time (Maskin et al 1983). Cardiac index at peak exercise, measured non-invasively, correlates to oxygen consumption (Shelton et al 2010, Hummel et al 2012) but whilst being lower in patients with heart failure than in healthy adults, rises during equivalent exercise by the same amount (Shelton et al 2010). This suggests that cardiac output does not limit exercise, at least at submaximal levels. In agreement with this is a study by Esposito et al, which showed that peak oxygen consumption during maximal cycle exercise was 33% lower in patients with CHF compared to healthy controls. During single leg knee extension (when the required cardiac output will be much lower than required for cycle exercise, freeing the heart from its role as limiter) oxygen consumption
consumption was also lower than in healthy controls (Esposito et al 2010). If cardiac output was the only limiting factor then we would also expect that at peak exercise oxygen consumption could not increase further, but Jondeau et al showed that the addition of arm crank exercise to cycling increased this significantly in patients with severe CHF, but not healthy controls, suggesting that prior to the addition of arm crank exercise the leg muscles had reached a maximum, beyond which they cannot extract more oxygen even if more is made available, but the addition of a further muscle group (the arms) did allow further cardiac output (Jondeau et al 1992).

Positive interventions to the heart which acutely improve haemodynamics therefore do not appear to rapidly improve exercise capacity. However the reverse is not true. When patients with left ventricular assist device (LVAD) implantation for severe left ventricular dysfunction have their LVAD support reduced (by decreasing the pump speed) there is an immediate reduction in cardiac output and exercise capacity (Jakovljevic et al 2010, Noor et al 2012).

If cardiac output and other measures of cardiac function are only loosely related to oxygen consumption as a marker of exercise capacity in CHF, then maybe alternatively the haemodynamic abnormalities are strongly related to the degree of breathlessness, and this, rather than oxygen, limits exercise capacity. The cause of the increased sensation of breathlessness in CHF was, for a long time, believed to be increased ventilation to perfusion mismatch, so that, even if for any given activity oxygen consumption was the same as for healthy adults, ventilation to both consume this oxygen and eliminate carbon dioxide was higher. Perhaps one of the first suggestions that organs beyond the heart were associated in the abnormal ventilatory response to exercise was observed in a study by Fink et al which looked at the relationship between minute ventilation ($V_e$) and carbon dioxide ($V_{CO2}$) production. Wedge pressure correlated weakly with the $V_e/V_{CO2}$ relationship at rest, and not at all at peak exercise. Wedge pressure was then deliberately reduced through the use of prazosin or dobutamine; in neither group did the $V_e/V_{CO2}$ relationship significantly improve (Fink et al 1986). Substantial work over the past two decades has gone on to elucidate more about the cause of breathlessness in heart failure. In a study by Wensel et al an increase in dead space, which will be affected by abnormal ventilation: perfusion mismatch, correlated with the relationship between carbon dioxide (CO$_2$) and ventilation, but in addition an independent mechanism, the depression of arterial CO$_2$, correlated more closely (Wensel et al 2004). It was therefore proposed that there is a strong drive to ventilation outside of the lungs. An explanation was that lactate may be this stimulus, however blood lactate levels did not correlate with the arterial CO$_2$ levels (paradoxically
the patients with the lowest CO₂ values also had the lowest hydrogen ion concentration. Interestingly it was shown, in the patients with the most abnormal ventilation to CO₂ relationship that a pronounced rise in systemic lactate levels occurred during recovery. Again the relationship with arterial CO₂ behaved in a paradoxical manner; the increased hydrogen ions should have stimulated a respiratory compensation and a reduction in arterial CO₂, but the opposite occurred (Wensel et al 2005).

All this evidence therefore suggests a situation where cardiac haemodynamics are abnormal at rest and on exercise, but the degree of abnormality does not appear to relate closely to the severity of exercise limitation, although perhaps cardiac index and wedge pressure are more tightly related to exercise capacity than other variables. Cardiac output is reduced but can improve with intervention; however this does not generally restore exercise capacity, at least not immediately. Finally ventilation: perfusion mismatch arising from abnormal pulmonary vasoregulation and blood flow distribution only explains some of the increased ventilation and symptoms of breathlessness found in patients with CHF. Despite changes within the lung, arterial blood leaving the heart is generally fully oxygenated, and so it is to the muscle that we look to explain the reductions in exercise capacity.

The skeletal muscle is commonly altered in CHF (Clark et al 1996) and the cause of this “myopathy of heart failure” is multifactorial; including deconditioning through underuse, the toxic accumulation of neurohormonal substances, the systemic inflammatory response, and possibly iatrogenic causes such as statins and the well-meaning, but ultimately harmful, medical discouragement of exercise.

Total muscle mass is typically reduced (Mancini et al 1992) and there are characteristic structural alterations such as a switch in muscle fibre type. In a study by Mancini et al, type IIb fibres were found as a greater percentage of total muscle fibres in patients with CHF. Type IIb fibres are glycolytic and much less fatigue resistant than the type I oxidative fibres. Oxygen consumption inversely correlated with the percentage of type IIb fibres (r= -0.81), and positively correlated with the percentage of type I fibres (r= 0.68) (Mancini et al 1989). Other findings include fibre atrophy (Lipkin et al 1988), decreased mitochondrial density (Drexler et al 1992, Esposito et al 2010) and abnormal enzyme activity (Mancini et al 1989, Sullivan et al 1990, Ralston et al 1991, Drexler et al 1992, Okita et al 1998). There is conflict regarding the changes within the capillary density of the muscle, with Esposito et al and Mancini et al finding a trend to increased capillary density. Sullivan et al showed that capillaries per fibre were reduced although density based on cross-sectional area was normal. Overall it appears likely that the area of muscle supplied by each capillary is largely unchanged, but other
factors, such as sympathetic activation, probably restrict flow. The consequences of these changes are a reduction in muscular peak power (Lipkin et al 1988) and early fatigability (Wilson et al 1996). The muscles of patients with CHF are unique in that they display signs of oxygen supply and demand limitation. Athletes typically display signs of oxygen supply limitation; that is the introduction of hyperoxia improves oxygen consumption at the muscle (Knight et al 1993), whereas sedentary individuals are oxygen demand limited; the addition of oxygen does not increase oxygen consumption further (Cardus et al 1998). Patients with CHF show improvements in oxygen consumption at the muscle with oxygen therapy, suggesting a degree of supply limitation, but even with 100% oxygen, they still do not match healthy adults in normoxia (Esposito et al 2010). This has been described as a reduction in both the convective (delivery of oxygen within the larger vessels) and diffusive (ability of the muscles to extract the delivered oxygen) capabilities of the muscle.

The muscle may also give us the answer to the “non-respiratory” portion of the high ventilatory levels seen in CHF, which appears to be caused by an autonomic ventilatory response activated by signals from exercising muscles – the ergoreflex. This has been shown to increase ventilatory drive in patients compared with healthy controls during similar levels of exercise (Piepoli et al 1996, Scott et al 2002). The addition of sodium bicarbonate (which will reduce acidosis) but not a glycogen free diet (which will reduce lactate) can abrogate this reflex, suggesting that it is principally driven by high intramuscular hydrogen ion concentration (Scott et al 2002). This may therefore allow us to reconcile some of the paradoxical findings by Wensel et al. In patients with CHF, high intramuscular concentrations of hydrogen ions drives ventilation through the ergoreflex, this lowers arterial CO₂. Possibly through reduced tissue blood flow, venous efflux is reduced during exercise, so that the increased hydrogen ion concentration is not measured in the blood. On recovery these ions spill over into the blood, reducing the ergoreflex but giving the impression of a systemic acidosis.

As further support in the argument for the central role of the muscle in CHF, exercise rehabilitation, and muscle strengthening have been shown to increase indices of muscle function and, in tandem, exercise capacity and prognosis. Critically prognosis is improved, as shown by a meta-analysis of 9 studies utilising exercise rehabilitation programmes (Piepoli et al 2004). However this doesn’t give direct evidence that the muscle is central to CHF, as whole body exercise could be improving the cardiac haemodynamics of the patients and therefore prognosis. Following an 8 week training program consisting only of knee extension exercises (deliberately designed to avoid stressing, and thereby strengthening the cardiovascular system), Esposito et al showed that peak oxygen consumption could increase to levels seen in matched sedentary controls, despite significantly lower values pre-programme (Esposito et al 2011). Cardiac output was not significantly altered,
but the convective and diffusive components of oxygen delivery to the muscles both increased. It has also been postulated that the main benefits of exercise training would be to improve the efficiency and comfort of submaximal activity rather than measures of maximal exercise capacity; submaximal exercise is after all more representative of day-to-day activity. Following 6 months of a cycle exercise programme patients significantly improved their maximal oxygen consumption by approximately 10% (unlike controls with CHF who did not undertake regular exercise and showed no improvements in maximal oxygen consumption). However more impressive were the improvements in variables based on the oxygen kinetics of submaximal exercise during constant work rate testing, with changes of nearer 20% seen (Mezzani et al 2012); these improvements are believed to reflect improved symptomatology.

In conclusion skeletal muscular abnormalities in CHF probably explain the heterogeneity in exercise capacity previously unexplained by markers of cardiac function. Structural, metabolic and neural changes lead to reduced convective and diffusive components of oxygen delivery to the muscle, an increased drive to ventilation (and therefore breathlessness) and exercise training, even when only directed at small muscle groups, improves exercise capacity without necessarily improving cardiac function.

1.3 The interaction of heart and lungs in CHF

CHF commonly affects the lungs, not just through an increase in interstitial fluid, but with easily identifiable, objective abnormalities such as restrictive defects, airway obstruction, diffusion abnormalities and inspiratory muscle weakness (Hosenpud et al 1990, Naum et al 1992, Dimopoulou et al 1998, Hughes et al 1999, Daganou et al 1999).

Left atrial pressure in heart failure is almost always elevated, if not at rest, on exercise, causing pulmonary vascular congestion. Initially, to limit the onset of interstitial oedema, lymphatic drainage is able to compensate, however once lymphatic clearance is maximal (this often occurs early due to systemic venous pressure elevation) interstitial oedema intervenes. With chronic elevations in pulmonary pressures, protective mechanisms occur. Within the pulmonary vasculature these lead to medial and intimal thickening, which affects perfusion of the alveoli, and in the capillary endothelium structural changes reduce microvascular permeability; these limit the transfer of oxygen and carbon dioxide, the key role of the lungs (Chua et al 1995).
On spirometry the commonest abnormality seen is a reduction in vital capacity – a restrictive pattern. Less commonly an obstructive pattern is noted. The latter phenomenon is more commonly seen during an acute decompensation of heart failure, rather than the stable chronic state.

1.4 Lung disease is common in CHF

Heart disease and respiratory disease, exemplified by CHF and chronic obstructive pulmonary disease (COPD) respectively, are very common within the UK. Furthermore respiratory disease often coexists with cardiovascular disease (Rutten et al 2005, Mascarenhas et al 2008, Iversen et al 2008, Hawkins et al 2010). In a recent study the prevalence of COPD in an out-patient cohort with CHF was 39.2% (Mascarenhas et al 2008). Similarly approximately 30% of patients with stable COPD have unrecognised heart failure (Rutten et al 2005). Within the UK this prevalence of coexistence appears to be rising (19.8% of heart failure patients had coexistent COPD in 1999, rising to 23.8% in 2004 (Hawkins et al 2010)). The reason for the high coexistence rates may relate to certain undesirable risk factors, principally smoking.

When both conditions coexist, management of either condition may be compromised by the presence of the other. A common problem encountered is the reluctance of physicians to initiate beta-blockers – drugs proven to significantly reduce mortality in CHF (CIBIS investigators and Committees 1994, MERIT-HF Study Group 1999, Packer et al 2002) – in the presence of COPD; this despite plenty of evidence to show safety in this patient group (Gottlieb et al 1998, Chen et al 2001, Salpeter et al 2002, Kotlyar et al 2002, Sirak et al 2004). In one study of CHF patients only 22% with a diagnosis of COPD received beta-blockers, and mortality, over an average of just under 2 years, was 17% compared with 31% for the remainder of patients not treated with beta-blockers (Staszewsky et al 2007). In a large retrospective analysis of 200,000 patients following a myocardial infarction, only 22.1% (compared with 34.3% of the total cohort) of patients with COPD received beta-blocker. Mortality was 16.8% at 2 years in those prescribed beta-blockers compared with 27.8% if they were not (Gottlieb et al 1998). However, given that beta-blockers may not benefit survival in patients with severe COPD (Chen et al 2001), these results may simply reflect a more severely affected spectrum of patients in the non-beta blocked groups. Randomised trials will be required to answer the question definitively.
1.5 The respiratory patient with exercise intolerance or dyspnoea

It is estimated that two thirds of patients presenting with exertional limitation or chronic dyspnoea have a cardiac or pulmonary cause for the symptoms and as described above these two groups of diseases commonly coexist.

Pulmonary disease commonly presents with dyspnoea as the predominant symptom. Chronic Obstructive Pulmonary Disease (COPD), asthma, interstitial lung disease, chest wall disease (including muscular abnormalities of the chest wall muscles or diaphragm) are amongst the most common diseases. With the exception of asthma, these diseases are all generally progressive, leading to deterioration in symptoms including dyspnoea. There is a strong environmental cause linked in many patients.

COPD is one of the most common chronic diseases in Westernised cultures with a proposed 6.9% prevalence of mild COPD in adults in the USA aged between 25-75 years (Mannino et al 2000). It is characterised by airflow limitation, which unlike asthma, is not fully reversible. This is usually progressive with an abnormal inflammatory response of the lungs. It can be viewed as a combination of chronic bronchitis, a clinical diagnosis defined by the presence of a chronic productive cough for 3 months in subsequent years; and emphysema, a pathological diagnosis with airspace dilatation of the distal bronchial tree (American Thoracic Society/European Respiratory Society Standards for the Diagnosis and Management of Patients with COPD 2004). Currently diagnosis is made using a combination of typical symptoms (chronic dyspnoea, cough, sputum production), appropriate environmental exposures (principally cigarette smoking) and appropriate spirometric findings (forced expiration in 1 second (FEV$_1$)/Forced Vital Capacity (FVC) ratio of <0.7) (GOLD 2011) or more rarely in the presence of an inherited disease predisposing to COPD (e.g. α-1 antitrypsin deficiency). Use of spirometry alone would lead to an over-diagnosis - a normal FEV$_1$, with supranormal FVC resulting in a low ratio, is not uncommon. In these cases a lack of symptoms and exposure risks limits the chance of a misdiagnosis. Whilst COPD is a disease of the respiratory parenchyma and airways it often has effects beyond the lungs. There are consequences for both sides of the heart (Barr et al 2010), and importantly the musculature (Donaldson et al 2012) which we will discuss in detail. There are also multi-organ consequences of the treatment, for example osteoporosis and myopathy with recurrent steroid use. Hence COPD, whilst identified as the most common, chronic progressive respiratory disease, should not be thought of as purely a disease of the lungs.
1.6 The muscle in COPD

Similarly to heart failure, COPD is a multiorgan disease with consequences to the musculature as well as the lungs. In CHF it is generally the peripheral muscles such as the calves and quadriceps that have been investigated. In COPD patients the ventilatory muscles such as the diaphragm also deserve special mention.

COPD behaves similarly to CHF, in that there is a reduction in strength of the peripheral muscles (Man et al 2003, Man et al 2005), with early fatigability (Mador et al 2003, Allaire et al 2004). There is a shift in fibre type, away from oxidative type I fibres, towards the glycolytic type II fibres (Jobin et al 1998). Capillary density reduces (Jobin et al 1998) and there are reductions in the enzyme activity involved in aerobic activity within the muscle (Maltais et al 1996). Conversely the diaphragmatic muscle is able to increase its force generating capabilities (Similowski et al 1991), with a shift towards type 1 fibres (Levine et al 1997). This likely represents the continued training exposure of the high ventilatory demand of COPD, and gives evidence that muscular training, as it does for CHF, can abrogate some of the abnormalities found. Again, in a similar fashion to CHF, the degree of peripheral muscular dysfunction correlates relatively poorly to resting determination (spirometry) of COPD severity (Gosker et al 2003, Seymour et al 2010).

Therefore, for physicians treating patients presenting with exertional limitation, a difficulty with diagnosis, follow-up and management can exist; COPD and CHF are common, frequently overlap, and can lead to consequences within the other organ system. Diagnosis alone may be challenging, but in patients with coexistence of these two conditions it is very difficult to elucidate the relative contribution of each to the symptoms of dyspnoea and exercise intolerance. Importantly both conditions can lead to similar changes in skeletal muscle which will exacerbate symptoms of exercise limitation.

Cardiopulmonary exercise testing has the potential to elucidate the degree of exercise limitation and identify principal limiting physiology in a patient, but it is unclear at present if a variable exists for either organ system that can quantitate the degree of specific organ dysfunction. It is the aim of this thesis to discover if such a variable on CPX exists. I will first discuss how CPX works, followed by the aims of the thesis and potential candidates for “ideal variables”.

33
1.7 Cardiopulmonary exercise testing, the requirement for oxygen and the role of the heart, lungs, vasculature and muscles in performing exercise

Gas exchange is the principal role of the heart and lungs. Cardiopulmonary exercise testing is a clinical test involving exercising an individual whilst making real-time measurements of their ventilation, and the concentrations of oxygen and carbon dioxide within the ventilation allowing us to calculate gas exchange in incredible detail. In order to understand CPX, an understanding of the physiological processes of gas exchange within the body, and how they go wrong in disease, is necessary.

Oxygen (O\textsubscript{2}) is required by all metabolic tissues, without this life could not exist. The consequence of this oxidative metabolism is the production of carbon dioxide (CO\textsubscript{2}) as a waste product. Although it is a continuous cycle, we will view it as starting with an inspiration; negative pressure is generated within the chest cavity and air travels from the mouth down the airways and into the alveoli of the lungs. Assuming normal atmospheric concentrations of O\textsubscript{2} and no significant left to right shunting within the heart, there will then exist a concentration gradient across the alveolar surface to the pulmonary capillaries, across which O\textsubscript{2} will diffuse, to bind to the metalloprotein haemoglobin, raising blood oxygen levels. Blood passes to the left side of the heart where rhythmic contractions eject a volume of blood approximately once a second. This blood travels down arteries, then arterioles to target organs. Here the blood vessels, now called capillaries, get smaller and densely infiltrate the tissue to allow easy diffusion of oxygen, again down a concentration gradient, into metabolising cells. Along with glucose (or alternative energy sources such as protein or fat), this oxygen is involved in the production of finite packets of chemical energy known as adenosine triphosphate (ATP) from adenosine diphosphate (ADP) and inorganic phosphates (P\textsubscript{i}), which allow the tissue to perform its actions. In the process of the conversion of an energy source and O\textsubscript{2} to ATP, termed cellular respiration, carbon dioxide and water are produced as waste products. When glucose is the principal energy source the conversion is as below:

\[
C_{6}H_{12}O_{6} + 6O_{2} \rightarrow 6CO_{2} + 6H_{2}O
\]

There are 3 principal parts to cellular respiration, through which ATP is produced (Figure 1.2); firstly glycolysis, which is not oxygen dependent, takes the raw glucose and produces intermediate metabolites, one of which, pyruvate, then passes into the second part, the Kreb’s cycle, with the production of some ATP, carbon dioxide and further intermediate metabolites. Finally many of these intermediates enter the electron transport
chain, which in the presence of oxygen produces the majority of the 36 ATP units. This process is termed *aerobic respiration*.

**Figure 1.2:** The intracellular production of energy, CO$_2$ and water from carbohydrate and O$_2$. Within the cytosol glucose is converted to pyruvate with the production of the intermediate NADH+H$^+$. Pyruvate crosses the mitochondrial membrane where it enters the Kreb’s cycle, producing further NADH+H$^+$ and CO$_2$. The NADH+H$^+$ is then oxidised to produce water and ATP. The original NAD is regenerated.
Parts 2 and 3 are dependent on each other and oxygen; but when oxygen supply cannot keep up with demand the intermediates (NADH + H\(^+\)) quickly build up limiting further energy production. In contrast glycolysis produces metabolites which, in the absence of sufficient oxygen, can build up within the cytosol of the cell without stopping further glycolysis. This allows a small production of energy (3 ATP units compared to 36 in the presence of oxygen) and so is only used by the tissue as a short-term solution, firstly because it is highly inefficient (only 8.1% of each glucose molecule’s potential energy is utilised) and secondly because this intermediate is necessarily converted to lactate, an acidic by-product, which impairs cellular activity as it builds up; this process is called *anaerobic respiration*. The lactate allows the regeneration of NAD, a necessary intermediate in glycolysis.

Fortunately when the conditions are improved and adequate oxygen supply resumes (typically following the cessation of exercise) this formation of lactate can be reversed, with these metabolites entering the Kreb’s cycle and allowing the generation of the full quota of potential energy. The production of lactate forms a central part of the analysis of gas exchange.

Following cellular respiration carbon dioxide is produced as a waste product. This is found in higher concentrations in the metabolic tissue than the blood vessels bringing blood to the tissue so CO\(_2\) diffuses in the opposite direction to O\(_2\), into the blood. This blood then returns to the heart, and thereon to the lungs. In contrast to the transfer of oxygen, CO\(_2\) is in greater concentration in the blood of the capillary bed of the lungs than the alveolar air, so moves out into the air, in the opposite direction to the oxygen transfer. By generating positive intra-thoracic pressures, expiration occurs, expelling this air (now lower in O\(_2\) and filled with CO\(_2\)) into the atmosphere, allowing the cycle to start again.

Exercise intolerance can occur with a defect in any, or many, of these processes described above. Abnormalities within the lungs will lead to an inability to ventilate sufficiently at an earlier point during activity, typically resulting in a build up of CO\(_2\). Oxygen delivery to the muscles can remain relatively normal and cessation of exercise is typically a consequence of a mechanical inability to ventilate further at higher rates/volumes, with a perception of respiratory difficulty even if physiological reserve exists. However as described above, the lungs are not in isolation in respiratory disease, there are often associated problems with the heart and circulation, limiting the ability to distribute blood around the body; within the muscles, so that capillary density may be reduced, leading to a lower relative supply of O\(_2\); and potentially changes within the blood itself.
In heart failure a primary abnormality is reduced oxygen delivery to the muscles. At rest oxygen consumption may appear normal, but a reduced cardiac output requires compensation with increased oxygen extraction (Shelton et al 2010). During exercise the cardiac output increasingly fails to increase appropriately to deliver blood to exercising muscles, requiring even greater oxygen extraction from the muscle. Oxygen extraction cannot go above a certain threshold, approximately 80%, so earlier than expected the Kreb’s cycle and electron transport chain are starved of oxygen. Glycolysis however can continue to produce some energy for exercise. One of the intermediates, NAD, is required to help produce ATP, but builds up in its reduced form, NADH + H⁺ in the absence of adequate oxygen. This NADH + H⁺ can be oxidised back to NAD by the reduction of pyruvate to lactate, however the build-up of the acidic lactate, can have unwanted consequences. This generally impairs the ability of the muscle to perform work and is almost always perceived as painful or discomforting, leading to early exercise cessation, although it also has beneficial properties such as shifting the haemoglobin dissociation curve to the right (the Bohr effect) facilitating the unloading of further oxygen stores within the blood.

Isolated problems within the muscle will also lead to symptoms of exercise intolerance. This may be anatomical, with a reduction in the motor units of the muscle (which can even be seen after a long period of relative sedentary behaviour) or a mismatch of perfusing capillaries to muscle; or metabolic, usually with an inherited problem with glycolysis, the Kreb’s cycle, the electron transport chain, or another metabolic pathway required to supplement and complement them. Mitochondrial myopathies are probably much more common than either believed or diagnosed (Flaherty et al 2001). There may be a problem with the arterial or venous vessels used to supply blood to, or remove blood from, the muscles respectively, such as peripheral vascular disease. Blood may also have limitations to its oxygen transport capacity, as seen with anaemia and certain haemoglobinopathies.

CPX allows a way, through measuring the uptake of O₂ and excretion of CO₂ at the mouth, of understanding how all these processes perform, and how they can go wrong in disease.

Measurements of oxygen uptake (\(\dot{V}_{\text{O}_2}\)) are central to CPX and this uptake measured at the lungs is believed to equal the consumption of oxygen at the tissues (\(\dot{Q}_{\text{O}_2}\)), and therefore the term oxygen consumption, whilst scientifically incorrect, is often used; they will be used interchangeably within this manuscript. In an equal (albeit opposite) way \(\dot{V}_{\text{CO}_2}\), the elimination of carbon dioxide by the lungs, is calculated and reflects the production of CO₂ at the muscles. However, although oxygen uptake/consumption is increased in only one manner - the aerobic conversion of fuel to ATP for energy - carbon dioxide elimination/ production occurs in 3
main ways; as a end-product of the conversion of fuel and oxygen for energy, to buffer the build-up of hydrogen ions from lactate production, and as respiratory compensation for the metabolic acidosis caused when hydrogen ion (from the lactate) build-up overwhelms local bicarbonate buffering capacity.

As exercise increases in intensity, oxygen uptake and carbon dioxide elimination increase, the latter at a progressively faster rate than the former. When exercise stops, although no new work is being done, there is an oxygen deficit to be replaced, so \( \dot{V}_{O_2} \) drops towards baseline over a few minutes. Lactate in the muscle is re-integrated into the Kreb’s cycle, and from there the electron transport chain, to form energy aerobically, and replenish cellular stores for future exercise. \( \dot{V}_{CO_2} \) takes much longer to return to normal because of the acidosis.

### 1.8 Common CPX Variables in the diagnosis and assessment of dyspnoea

Patients commonly complain of exercise intolerance, exercise fatigue or breathlessness on exercise. History and examination alone commonly fails to find a diagnosis (Pratter et al 1989, DePaso et al 1991). This is principally because the descriptions of the symptoms are themselves of limited diagnostic value as they vary minimally between diseases. Common clinical tests to identify the aetiology include pulmonary function, chest radiography, blood chemistry and haematology, electrocardiography (ECG) and echocardiography. But as described earlier, the use of spirometry to identify respiratory disease may be flawed by the common prevalence of abnormalities on spirometry in patients with cardiac disease. Likewise echocardiographic abnormalities can occur in patients with respiratory disease. A possible distinguisher between cardiac and respiratory disease is B-type natriuretic peptide (BNP), secreted by cardiomyocytes, which rises in heart failure. However patients with stable, compensated heart failure may normalise their BNP and in patients with COPD, elevations in BNP have been shown (Ando et al 1996).

Given that the symptoms are typically present on exercise and not at rest, it makes sense that to differentiate the cause of exercise limitation, exercise should be utilised. CPX generates a number of variables that can be useful both in identifying the presence of true exercise intolerance and the cause of exercise limitation. Once a diagnosis is made CPX is invaluable in determining the prognosis of common cardiac and respiratory diseases. This is important because both CHF and COPD have significant prognostic implications. There are 2 well quoted statistics for patients with CHF; 1 year mortality is 50% following an admission, and secondly it carries a worse prognosis than many cancers. Whilst these facts were identified prior to current therapeutic options, CHF still causes significant mortality. In the UK in 2008/9, following a hospital admission with a diagnosis of
heart failure, in-hospital mortality was 12%, and a further 22% of patients died during follow-up (average 158 days) (Cleland et al 2011), giving a predicted total 1 year mortality >40%. These are similar values to those seen at the end of the last century where 1 year mortality in UK patients diagnosed with heart failure (either as an admission or through out-patient services) was 38% (Cowie et al 2000). When population screening is used to identify patients with heart failure (rather than wait for presentation to health services) prognosis is better but still significantly reduced compared with healthy adults; a 5 year mortality rate of >40% was found in one UK study (Hobbs et al 2007) and 41% in a Dutch study (Mosterd et al 2001).

COPD has an overall better prognosis than CHF. Following admission for COPD, 1 year mortality was 22% (Almagro et al 2002). Another study showed a median survival of 3.1 years (Incalzi et al 1997). Despite these differences in prognosis CPX has excellent utility for determining prognosis in both (Mancini et al 1991, Cohn et al 1993, Hiraga et al 2003, Oga et al 2003, Cote et al 2007).

1.8.1 Peak $\dot{V}_{O_2}/N_{O_2, Max}$

Oxygen uptake (practically synonymous with oxygen consumption) is the principal measurement of CPX. It increases throughout incremental exercise, and around the termination of exercise, when exercise is at its most intense, $\dot{V}_{O_2}$ will be at its highest. This highest value is termed the peak $\dot{V}_{O_2}$. Peak $\dot{V}_{O_2}$ is the most commonly described CPX variable.

$\dot{V}_{O_2}$ max is often used synonymously with peak $\dot{V}_{O_2}$. It is the highest attainable $\dot{V}_{O_2}$ for the patient, and when effort is truly maximal on an exercise test it will equal the peak $\dot{V}_{O_2}$. However because it is difficult to confirm maximal effort and therefore the attainment of $\dot{V}_{O_2}$ max, especially in patients rather than athletes, peak $\dot{V}_{O_2}$ is preferentially used.

Peak $\dot{V}_{O_2}$ can give us a surrogate for maximal cardiac output, with $\dot{V}_{O_2}$ and cardiac output related via the Fick Principle. This may explain why peak $\dot{V}_{O_2}$ has shown a great ability to determine prognosis in heart failure, which is largely thought of as a disease of reduced cardiac output. Peak $\dot{V}_{O_2}$ has been the most widely-used and accepted variable for the assessment of prognosis in patients with CHF and there have been numerous studies validating its ability to determine prognosis (Szlachcic et al 1985, Likoff et al 1987, Mancini et al 1991, Van den Broek et al 1992, Cohn et al 1993, Cohen-Solal et al 1997). However peak $\dot{V}_{O_2}$ falls in most disease states so is not useful when trying to discriminate causes of exercise intolerance. It is a strong prognostic marker in
respiratory disease as well (Oga et al. 2003, Cote et al. 2007) but what is interesting is that the relationship
between prognosis of these two disease groups and peak $V_{O_2}$ differs vastly. Considering exercise capacity, COPD
mortality remains markedly lower than similarly limited patients with CHF. In a cohort with COPD (mean age
of 67 years and a mean peak $V_{O_2}$ of 10.8 mL/min/kg) the 12 month mortality was only 4% (Cote et al. 2007). A
comparable study in CHF with a similar peak $V_{O_2}$ (mean 11.0 mL/min/kg) showed a 10-fold greater mortality
(40%) over the same period despite a mean age 10 years younger (Szlachcic et al. 1985) although it must be
noted that 2 decades of time separate these two studies. In patients with both conditions it appears to be the heart
failure that drives the prognosis as shown in a study by Mascarenhas et al where patients with CHF and
concurrent mild to moderate COPD (GOLD stages 1 and 2) had a similar prognosis compared to those patients
with CHF alone (Mascarenhas et al. 2008).

1.8.2 Anaerobic/ Lactate threshold
As described above there comes a point in incremental exercise when oxygen supply cannot meet demand,
pyruvate is reduced to lactate leading to the oxidation of NADH + H$^+$, to allow the regeneration of NAD for
further energy generation. Lactate liberates hydrogen ions (it is acidic) and an increase in circulating levels
prompts changes in ventilatory patterns. It is typically identified on cardiopulmonary exercise testing when there
is a change in the relationship between CO$_2$ and O$_2$ and is termed the anaerobic or lactate threshold. The merits
of various ways of identifying it will be discussed in the methods. It is believed that the early onset of anaerobic
metabolism is a specific sign of heart failure (Nery et al. 1983, Weber et al. 1984, Hansen et al. 1984) and it may
be a stronger prognostic marker in this population than peak $V_{O_2}$ (Gitt et al. 2002).

1.8.3 Ventilatory equivalents
An important concept in gas analysis is the relationship between ventilation and either the oxygen uptake or,
much more commonly, the carbon dioxide elimination. We will focus on the latter here. This is generally
described as the efficiency of ventilation to eliminate carbon dioxide. This $V_E/V_{CO_2}$ relationship can be described
throughout exercise, or at certain pre-specified points. One of the commonest forms is the $V_E/V_{CO_2}$ slope. Lower
values are better.
Throughout the majority of exercise $\dot{V}_E$ and $\dot{V}_{CO_2}$ are linearly related and it is the slope of this relationship, during the linear portion, that typically constitutes the $\dot{V}_E/\dot{V}_{CO_2}$ slope (Figure 1.3). As exercise continues the relationship deviates from linearity as ventilation increases disproportionately at high intensities.

**Figure 1.3: Minute ventilation and carbon dioxide elimination rise together during incremental exercise.**
Initially this is linear and the slope of this relationship is termed the $\dot{V}_E/\dot{V}_{CO_2}$ slope (22.68 in this example). Towards the end of exercise the buffering systems of bicarbonate are overwhelmed and respiratory compensation (by reducing arterial CO$_2$) for the metabolic acidosis occurs, decoupling the previously tight $\dot{V}_E$ to $\dot{V}_{CO_2}$ relationship.
Alternatively the $\dot{V}_E/\dot{V}_{CO_2}$ relationship can be calculated instantaneously; this is termed the $\dot{V}_E/\dot{V}_{CO_2}$ ratio and is measured at a number of points, such as the anaerobic threshold, or its lowest point. These are almost always very similar values, and similar to the $\dot{V}_E/\dot{V}_{CO_2}$ slope. Heart failure patients exhibit an elevated $\dot{V}_E/\dot{V}_{CO_2}$ relationship for reasons well described above; increased ventilation: perfusion mismatch, and primary hyperventilation during exercise leading to a lower arterial $CO_2$ partial pressure. Whilst the severity of this “ventilatory inefficiency” relates to peak $\dot{V}_{O_2}$, it gives us additional information regarding the pathophysiology of the patient beyond maximal exercise capacity. This may help explain why these measures largely outperform peak $\dot{V}_{O_2}$ when assessing for prognosis in heart failure (Chua et al 1997, Robbins et al 1999, Gitt et al 2002, Davies et al 2006). Because ventilatory equivalents are influenced by ventilation: perfusion mismatch, which is common in respiratory disease, abnormal equivalents may also seen in respiratory disease, so may not be useful when discriminating these 2 common conditions.

### 1.8.4 Breathing reserve

Our lungs have a theoretical maximum ventilation which they can attain; this differs from person to person and changes over time. An estimate of this maximum voluntary ventilation (MVV) is possible for each patient and, from this, the percentage remaining, which is termed the breathing reserve (BR). Typically healthy adults have a BR >30%, i.e. at peak exercise their lungs have used less than 70% of their potential- it is other organs that limit exercise. With lung disease the maximum voluntary ventilation is reduced and so for any given level of ventilation (and therefore exercise) the breathing reserve is reduced. When reaching low levels of breathing reserve during exercise, i.e. lungs working at or near maximum, breathing becomes mechanically limiting and intolerable. It has been suggested that breathing reserve can be used to differentiate patients limited by their lungs (where it will be low) from patients with CHF or other conditions (where it will be normal) (Milani et al 2004). The breathing reserve at the anaerobic threshold may be more discriminant than at peak exercise (Medoff et al 1998).

### 1.8.5 Oxygen efficiency

When oxygen uptake is plotted against $\dot{V}_E$ the resulting relationship is curvilinear, and so is less commonly described than the largely linear $\dot{V}_E/\dot{V}_{CO_2}$ relationship. However the relationship between $\dot{V}_{O_2}$ and $\dot{V}_E$ has been
described in 2 ways. Firstly when $\dot{V}_E$ is logarithmically transformed the resultant plot is linear and is termed the Oxygen Uptake Efficiency Slope (OUES) (Baba et al 1996) (Figure 1.4).

More recently the use of the highest value of the ratio of $\dot{V}_{O_2}/\dot{V}_E$ has been explored, this has been termed the oxygen uptake efficiency plateau (OUEP) (Sun et al 2012).

These variables are typically used for the assessment of prognosis in CHF (Davies et al 2006, Sun et al 2012) where they appear to be superior to peak $\dot{V}_{O_2}$ and $\dot{V}_E/\dot{V}_{CO_2}$ slope, but our experience with their use in respiratory disease is limited.

1.8.6 Oxygen Uptake: Work Rate relationship

As muscular work increases, oxygen uptake increases linearly in tandem. There is a slight lag in oxygen uptake, resulting in an oxygen deficit but generally the relationship of the two is linear. In previous studies this linear relationship has been shown to vary minimally in healthy adults, with negligible influence from patient characteristics such as height, weight, age and gender. CHF, but not respiratory disease, has been shown to reduce the slope of this relationship, and it may therefore be beneficial to help discriminate between a primary cardiac or respiratory abnormality in a breathless patient (Hansen et al 1987). Its use in identifying prognosis is less well established than for peak $\dot{V}_{O_2}$, $\dot{V}_E/\dot{V}_{CO_2}$ relationship and the OUES.
**Figure 1.4: Graphical representation of how OUES is calculated.** A plot of oxygen consumption against minute ventilation shows a curvilinear relationship. When minute ventilation is logarithmically transformed it assumes a linear relationship and the OUES is the gradient of the slope of the regression line through this data.
1.8.7 An algorithm approach

So there exist a number of potential variables which may reflect cardiac physiology. However rather than using any single variable described above, an algorithm has been proposed utilising these variables to aid in diagnosis in a patient who presents with exercise intolerance and/or breathlessness on exertion (Figure 1.5).

It can be seen that 3 of the variables I have described are used together in this algorithm: peak $\dot{V}_o_2$, the anaerobic threshold and the breathing reserve. In a patient with respiratory and cardiac disease the approaches described above cannot identify the relative contribution of each disease to the patient’s exercise limitation. If this patient were to be followed up serially, whilst we could identify a deterioration in overall function using peak $\dot{V}_o_2$, we could not identify whether this came about through a change in their heart or lungs.

Therefore CPX can distinguish cardiac from respiratory disease, but not easily through the use of a single variable. If a single variable exists can it quantitate cardiac dysfunction? Certain variables show a greater prognostic potential in heart failure than others; does this relate to their greater specificity for cardiac disease over respiratory disease.

These questions form the basis for this thesis. Can we identify, for patients with cardiac disease, a variable in CPX that is largely independent of respiratory disease (one of commonest co-morbidities) with a predictable linear course that can be used to measure the degree of cardiac limitation?
1.9 Aims

The aim of this thesis is to identify a single variable on cardiopulmonary exercise testing that can differentiate cardiac from respiratory limitation, and that can quantitatively identify the impact of cardiovascular disease on the patient’s exercise capacity.

We will call this the “Ideal” variable although I do not suggest that it will necessarily behave perfectly, merely that it will have greater specificity and sensitivity to cardiac disease and changes in cardiac function over other variables. In a patient undergoing regular follow-up for a known cardiac condition this “Ideal” variable would be the best for serial assessment of their cardiac condition, and would be largely independent of changes in comorbidities, especially respiratory disease.

An “Ideal” variable would have the following characteristics:

1. Be significantly different between various pathophysiological groups (i.e. cardiac vs respiratory) so can be used to identify, in a patient with exercise limitation, the principal limiting physiology. This specificity is the key determinant of an “Ideal” variable.

2. Have good test-retest reproducibility. Our aim is to identify a variable that can be measured serially to identify subtle changes in cardiac function; a poorly reproducible variable will not be able to perform this function.

3. Change predictably and appropriately over time with deterioration/improvement in the relevant organ system. Specifically our “Ideal” variable will reflect improvements in cardiac physiology and haemodynamics and so when the heart is intervened upon, assuming this intervention is successfully positive, our variable should improve appropriately.

4. Closely relate to an accepted measure of exercise capacity (e.g. peak \( \dot{V}_{O_2} \)) or symptoms (NYHA class, BNP) at baseline in patients with cardiac limitation. The variable should therefore be a continuous variable where the magnitude directly relates to the severity of disease limitation. An inherent assumption of this statement is that peak \( \dot{V}_{O_2} \) is not specific, and will be abnormal in both cardiac and respiratory patients. Should the variable closely relate to peak \( \dot{V}_{O_2} \) in the cardiac but not in the respiratory patients this is further evidence for its specificity. The relationship with peak \( \dot{V}_{O_2} \) and other markers of disease will not be expected to relate as closely following intervention. This is because we
know, as described earlier, that exercise capacity may not normalise quickly after intervention despite immediate changes to cardiac haemodynamics. The “Ideal” variable should improve quickly, as it more directly reflects the cardiac function. As time passes the close relation between peak $\dot{V}_{\text{O}_2}$ and the “Ideal” variable should improve again as the whole body takes time to respond to the improved haemodynamics with changes in muscle function, lung function and vascular physiology.

5. Remain unchanged over time with deterioration/improvement in other organ systems. This will not be directly tested within the experimental setting of the thesis, but previous and future work will be discussed.

Initially, to confirm the proof of concept, 2 data sets will be examined. Firstly on a retrospective cohort of patients with clinical heart failure, referred for treadmill CPX tests at St Mary’s Hospital between 2003-2007, I will show the differential effect of spirometric measures on CPX variables. Secondly I have collaborated with the investigators of the Study of Health in Pomerania, a large population study in over 1200 adults undergoing a single bicycle CPX test. Using this data I will show the influence of certain demographic variables (such as smoking) and spirometric measures on CPX variables within healthy adults as well as define normal limits for some of the less typically used variables.

In the main study for this thesis I will prospectively recruit patients with respiratory disease, and patients with cardiac disease undergoing interventions aimed at improving cardiac haemodynamics. The directed comparison of these 2 groups (prior to intervention) will be used to investigate criteria 1. Each patient will undergo 2 exercise tests at baseline; the comparison of these 2 tests will be used to investigate criteria 2. Criteria 3 will be investigated comparing the exercise tests before and after intervention in the cardiac patients. Criteria 4 will be investigated by comparing measures of disease severity such as symptoms and symptom scores, overall exercise capacity, BNP, and resting echocardiographic and spirometric abnormalities to the CPX variables, both at baseline, and comparing to the change following intervention.
2.0 Methods & Materials
2.1 Overview

The overall aim of this study is to better identify ideal variables for the serial follow-up of patients with cardiovascular disease by cardiopulmonary exercise testing (CPX). The 2 central components to this are an Observational study and an Interventional study. Both of these arms involve the recruitment of patients with 2 common cardiovascular diseases; chronic heart failure (CHF) and mitral valve disease (to include mitral regurgitation (MR) and mitral stenosis (MS)). For the purposes of the Interventional arm, patients were awaiting either the implantation of cardiac resynchronisation therapy (CRT) for CHF, or surgical correction of mitral disease. For the Observational study these cardiovascular patients were compared with patients diagnosed with a common respiratory disease; chronic obstructive pulmonary disease (COPD). All patients recruited underwent testing to ensure their diagnosis was correct and alternative diagnoses did not coexist. These methods included echocardiography and lung function testing with spirometry, lung volume subdivisions and gas transfer. These methods will be discussed in detail later. As abnormal renal function and anaemia can affect ventilation, venous blood sampling was performed on one day of testing to allow for consideration of these factors on analysis.

All patients then underwent 2 baseline CPX tests on a bicycle ergometer, and this will be discussed in detail later.

2.2 Patient Identification and Recruitment

Ethics: South East London Research Ethics Committee (REC) Ref 10/H0805/36. Documents from the Ethical committee are included in appendix 2.

Starting date following ethical approval was 29th September 2010. The following groups of patients were recruited by the methods outlined in each.

2.2.1 Mitral valve disease

Mitral regurgitation is a common clinical condition, typically detected through cardiac auscultation and later confirmed via echocardiography. It exists on a spectrum, and is generally believed not to limit exercise capacity until of severe magnitude. Whilst causes are manifold, two typical situations exist. Firstly there is organic mitral regurgitation, usually involving leaflet prolapse, where one or both of the anterior and posterior valve leaflets do not correctly coapt, leading to regurgitation of blood into the left atrium, often during late systole.
There is typically a structural problem with the mitral valve apparatus as well as the leaflets (which often have redundant tissue). It is often amenable to surgical repair and found in younger, otherwise healthy individuals. This was the group that we largely recruited from. The second commonest current cause in the UK is ischaemic MR. Generally the valve is normal, but changes in ventricular geometry or papillary muscle dysfunction lead to incomplete closure at the end of diastole. This is often termed functional ischaemic mitral regurgitation and often does not respond well to surgical correction (because the principal problem is with the ventricle, not the valve). Another cause is secondary to infective endocarditis.

At our institution (Imperial College Healthcare NHS Trust – ICHNT) mitral valve surgery is performed in house by 2 experienced surgeons. Following identification of a patient with MR, a cardiologist from our institution or another local trust without in house surgical expertise, will review the patient and severity of their MR, and if deemed appropriate will be referred to an individual surgeon, or the surgical team via the multi-disciplinary Joint Cardiovascular Cardiology Meeting at Hammersmith Hospital. Guidance for the need for surgical intervention is based on international guidelines as shown below (Figure 2.1).

Due to our local high success rate of early reparative surgery, many patients are operated in the presymptomatic period prior to the onset of LV dysfunction (Class IIa indication). Once surgery has been agreed between the cardiologist, cardiothoracic surgeon, and importantly the patient, they are placed on an outpatient waiting list, with wait times of typically a few weeks. The patient may be identified for my study at any point along the decision making process, but this typically occurs by the cardiologist (where the cardiologist is also involved in the research study) or once on the surgical wait list. The 2 surgeons were involved in the study, and aided with recruitment.

Mitral stenosis (MS) is a rarer condition within the UK. It is almost always a consequence of rheumatic heart disease. Surgery is encouraged in symptomatic patients with a reduced valve area (<1.5cm²), unfavourable valve structure for percutaneous balloon mitral commissurotomy and no contraindications to surgery. It presents with symptoms similar to mitral regurgitation but reparative surgery is much less successful and replacement is typically required.
Figure 2.1: European Society of Cardiology guidance 2012 for treatment of organic mitral regurgitation.
2.2.2 Heart failure

Patients with CHF were recruited from 3 main areas. Firstly, and so that overall recruitment could be optimised for both the Observational and Interventional arms of the study, patients awaiting CRT for heart failure were recruited. The role of CRT and the indications for implantation will be explained in detail elsewhere. A member of the heart failure or electrophysiology team was involved in decisions regarding CRT and, after discussion with the patient, they were placed on a waiting list for CRT. Either following identification from the referring cardiologist or from the CRT waiting list, patients for the study were highlighted. The second cohort within the CHF group were patients with pre-existing CRT. These were typically identified from the pacing clinic, where I was clinically involved. However to maximise numbers of appropriate individuals for recruitment, a second London based hospital trust (Royal Brompton & Harefield NHS Foundation Trust) was involved. Ethical approval for the identification and recruitment of patients was granted by the South East London Research and Ethics Committee. Patients were recruited by the cardiology team at this trust and referred to our study team for further information and potential inclusion in the study. Finally because more patients with heart failure were required for the Observational study than the Interventional study a number of patients without CRT were recruited. These patients were under the care of the routine Heart Failure Clinic at the St Mary’s hospital site of ICHNT, and a community cardiology site at Maida Vale. Both myself and other members of the research team were members of these clinics, routinely involved in the care and identification of these patients.

2.2.3 COPD

Patients with COPD were only recruited from ICHNT. Involvement of the respiratory clinic and a principal consultant chest physician occurred regularly. Patients with symptomatic COPD were identified by members of this clinic and, after agreement with the patient, details were passed to me. A number of stable COPD patients were also reviewed by the pulmonary rehabilitation team, and so meetings between myself and this group occurred regularly.
2.2.4 Inclusion and exclusion criteria

The principal inclusion criteria were as below:

1. Patients with symptomatic left ventricular systolic dysfunction. Ideally patients either had or were awaiting the implantation of CRT. Left ventricular systolic dysfunction was defined as a left ventricular ejection fraction (LVEF) < 50%. Cut-offs of normality for LVEF are debatable; multiple previous heart failure studies used an LVEF < 40% as an inclusion criteria, but more recent guidelines define systolic impairment as an LVEF ≤ 54% (Lang et al 2006). Studies of patients with preserved ejection fraction have traditionally used 45% or 50%, results from the CHARM study show that LVEF is a strong predictor of outcome below 45%, but above this there was no correlation (Solomon et al 2005). 50% has been proposed as a reasonable cut-off (Mahadevan et al 2008) and I believe this has a reasonable degree of specificity without compromising sensitivity. Patients must have been symptomatic, either with exercise intolerance or breathlessness on exertion, but were allowed to be stable with minimal symptoms at enrolment, as it is well established that asymptomatic patients with left ventricular dysfunction typically still have a reduction in exercise capacity (LeJemtel et al 1994, Mahon et al 2000).

2. Patients with mitral valve disease awaiting mitral valve surgery. Mitral regurgitation or stenosis must have been severe, and ideally reparative surgery (not replacement) was considered as the first line surgical option. Mitral stenosis is not usually amenable to repair, so a valve replacement was not a contraindication. No limits were placed upon symptom status or ventricular function.

3. Symptomatic COPD or a condition with symptoms of COPD (e.g. bronchiectasis) with obstructive spirometry. COPD must have been at least mild, and ideally moderate or above; moderate corresponds to GOLD stage 2 (FEV₁:FVC <0.70, FEV₁ 50-80% predicted and typical symptoms of breathlessness on exertion, cough and sputum production).

4. A mixed cardiac and pulmonary disease group. Patients with known left ventricular or mitral valve disease awaiting surgery with concomitant COPD of at least GOLD stage 2 (FEV₁:FVC <0.70, FEV₁ 50-80% predicted and typical symptoms of breathlessness on exertion, cough and sputum production).
The principal exclusion criteria included:

- Inability to perform exercise and/or significant neurological or musculoskeletal limitation
- Renal impairment (eGFR <30 as calculated by the Modified Diet in Renal Disease formula), as severe renal dysfunction can affect ventilatory variables.
- Symptomatic coronary disease, or significant ischaemia noted on cardiopulmonary exercise test.
- Recent (recovery within <1 month) decompensation of their underlying cardiac or respiratory disorder.
- Anaemia (Hb <11).
- Significant obesity (BMI >40) which could lead to limitation independent of the cardiovascular or respiratory condition.
- Poor echocardiographic windows so that an assessment of left ventricular function and presence of significant valvular pathology cannot be reliably determined.
- Inability to perform spirometry.
- Inability to consent/make decisions (lack of competence)
- Vulnerable adult/current detainee in prison/elsewhere.
- Inability to complete all the investigations at the time intervals agreed in the protocol.
- Permanent Pacemaker unless in the biventricular pacemaker group in the Interventional Study.
- Atrial fibrillation if within the COPD group (this arrhythmia was allowed in the cardiac groups and mixed group because of the very high incidence).
- Uncontrolled hypertension.
- Untreated life threatening arrhythmia.
- Recent surgery (within 10 weeks).
- All standard absolute contra-indications to exercise testing.
2.3 Equipment and Testing

2.3.1 Cardiopulmonary exercise testing

2.3.1.1 Overview of exercise testing and terminology

Every cardiopulmonary exercise test was performed on a COSMED Quark CPET System (COSMED S.r.l. Rome, Italy). This employs a 28mm diameter bidirectional flow turbine for the measurement of flow, and a <1mm diameter sampling line to deliver expired gases to the Quark unit. This measures oxygen concentrations using a paramagnetic analyzer with a response time <130 ms. Carbon dioxide concentrations are measured using a digital infrared analyzer with a response time <100 ms. Recent American Heart Association (AHA) guidance into CPX testing was adhered to (Balady et al 2010). During testing breath-by-breath measurements through the turbine are recorded, giving a real-time view of gas exchange. Both inspiratory and expiratory flow volumes can be measured, this is tidal volume ($V_T$). The time delay between the onset of a given expiration and the preceding one is also measured, allowing respiratory frequency ($R_f$) in minutes$^{-1}$ to be calculated. The product of $V_T$ and $R_f$ is the minute ventilation ($\dot{V}_E$), which is the volume of air expired within one minute (measured in litres/min).

The two other main variables measured breath-by-breath are the uptake of oxygen ($\dot{V}_{O_2}$) and elimination of carbon dioxide ($\dot{V}_{CO_2}$). As described in Chapter 1 these are ventilatory surrogates for the consumption of oxygen ($\dot{Q}_{O_2}$) and production of carbon dioxide ($\dot{Q}_{CO_2}$) in the tissues (during exercise this is almost entirely the exercising skeletal muscle). In simple terms they are measured by calculating the product of the concentration of oxygen difference between expired air and room air and the expiration volume (and similarly for CO$_2$). In reality, because oxygen concentration in expired air changes throughout each breath, more complex techniques using integration are used.

As described earlier, CO$_2$ is the product of cellular respiration. When glucose is the fuel the ratio of CO$_2$ to O$_2$ is 1:1, however this ratio is different when the fuel source contains fats or proteins as shown below:

Glucose: $C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$ The ratio here is 6:6 or 1.00

Fat: $C_{16}H_{32}O_2 + 23O_2 \rightarrow 16CO_2 + 16H_2O$ The ratio here is 16:23 or 0.696

Protein: $C_{72}H_{112}N_{18}O_{22}S + 77O_2 \rightarrow 63CO_2 + 38H_2O + SO_3 + 9CO(NH_2)_2$ The ratio here is 63:77 or 0.818

These are just typical examples for fats and proteins. This ratio of CO$_2$:O$_2$ is termed the Respiratory Quotient or $RQ$. At rest the RQ lies below 1 as a combination of glucose, fats and protein make up the fuel source. As
exercise starts a greater percentage used is glucose (the preferential substrate of skeletal muscle) so this RQ will rise.

The RQ at the muscle can be seen to be a ratio that must be <1, as no energy source produces more CO₂ than it consumes O₂. However we cannot measure RQ at the muscle, only a surrogate at the mouth. The \( \frac{V_{CO₂}}{V_{O₂}} \) ratio measured is termed the *respiratory exchange ratio* (RER). At rest and low levels of exercise it resembles the RQ, although typically the action of measuring it elicits some hyperventilation of the patient, with an excess of CO₂ elimination, raising the RER above the RQ. This hyperventilation usually settles during very early in exercise, so that the RER falls to a baseline level somewhere around 0.75. As exercise continues the RER slowly climbs as proportionally more glucose is used by the skeletal muscle. As described in Chapter 1 when aerobic exercise can no longer keep up with demand, anaerobic metabolism starts with the production of lactate and hydrogen ions, which is initially buffered by CO₂ production. This CO₂ is expired with a rise in the RER. As lactate production increases and hydrogen ion concentration rises there is a second method of compensation, hyperventilation, with the result of resetting arterial CO₂ concentration at a lower level; this is respiratory compensation. If exercise is allowed to progress to this stage, these last 2 processes will almost always allow the RER to climb above 1; values of 1.2-1.4 at peak exercise are suggestive of near maximal effort.

With this understanding of these key measurements in exercise testing in place detailed description of the many variables I measured, and how they were measured, can be made.

*2.3.1.2 Calibration*

The Quark CPET system was calibrated prior to each patient’s exercise test. Two forms of calibration were performed.

Firstly calibration of gas analysis was performed (Figure 2.2). The sampling line was connected to the Quark device and a two-point calibration involved the controlled introduction of two gases down the sampling line was performed. The first gas was air, to allow for correct calibration of the current atmospheric conditions. The second gas was from a premixed cylinder of 16% O₂ and 5% CO₂ (approximately the concentrations in expired air). This allowed the calibration of both O₂ and CO₂ within the normal physiological range. Abnormalities in calibration allowed identification of problems that could have led to inaccurate gas analysis during CPX.
Secondly calibration of turbine flow was performed using a 3L syringe. This was attached to the turbine and 12 full emptying and filling manoeuvres of the syringe were performed. Acceptable turbine calibration involved obtaining values within the accepted limits of normality, as shown below for a typical patient (Figure 2.3).

No tests were performed unless both successful calibrations had occurred.

Figure 2.2: Calibration of the Gas Analyzer Results Table. This involved a two-point calibration with room air and a 16:5 mixture of 16% oxygen and 5% carbon dioxide (to roughly simulate expired air). The oxygen analyzer and carbon dioxide analyzer were both calibrated during this test, and the presence of an abnormality was shown with a red highlighted figure within the results table; this prompted the operator to look for, identify, and correct a problem, and re-calibrate before starting the exercise test.
Figure 2.3: Calibration of turbine flow. A 3L calibration syringe (COSMED, Italy) was attached to the turbine, and 12 full emptying and filling manoeuvres to simulate expiration and inspiration were performed. The box to the left shows each individual manoeuvre, before successful calibration is shown on the right.
2.3.1.3 Ergometer and test set-up

I performed each CPX test as principal operator.

Each test was performed on an Ergoselect 100, bicycle ergometer (Ergoline GmbH, Bitz, Baden-Württemberg, Germany) (Figure 2.4). This is a computer-controlled electronic ergometer capable to eliciting work rates of up to 999 watts, independent of pedal speed within the range 30-130 revolutions per minute (rpm).

Another operator was present for safety.

Bicycle ergometry rather than treadmill was chosen for its principal advantages; work rate can be measured, upper-body movement and therefore artefact is reduced, and it is safer than treadmill exercise. Work rate measurement was particularly important as it is necessary for the calculation of an important variable, the oxygen uptake: work rate relationship. Limitations of bicycle over treadmill testing include a lower muscle mass utilised (typically reflected in lower peak oxygen uptake) and more technically challenging in those unused to cycling.

Each patient had the ergometer adjusted to allow for optimal pedalling comfort. Following sitting on the ergometer the seat height was adjusted so that the patient’s legs were slightly bent at the knee (typical angle 10-15%) as recommended, and the feet secured using the Velcro foot straps. The handlebars were also adjusted towards or away from the patient as required. A few pedal turns were encouraged to ensure a comfortable pedalling position.

The Quark CPET turbine was fitted to the patient using a silicon mask with quick-release headgear straps (Hans Rudolph Inc, Shawnee, KS, USA). A tight seal was attempted for each patient by the use of a range of sizes of masks, and confirmed with an attempted moderately forced expiration against a blocked turbine (air could be heard escaping in the event of a poor seal, and the patient was able to make a maximal expiration). A poor seal led to further attempts to ensure optimal positioning to eliminate/minimise mask leak.
Figure 2.4: Ergoselect bicycle ergometer. Seat height and handlebar position were adjustable for patient comfort. There is a panel at the front, visible only to the operator, displaying work rate, time and cadence. Another panel to the rear of this, visible to the patient, only displays the cadence.
Prior to exercise testing, baseline spirometry was performed for each patient through the turbine, using the manufacturer’s software System (PFT Suite, COSMED S.r.l. Rome, Italy) (Figure 2.5). I described each manoeuvre with 3 parts: firstly a gentle expiration through an open mouth; immediately into a full, controlled inspiration; and finally, without a pause, a maximal forced expiration until a plateau in expiration had occurred (patients were encouraged to expire for at least 6 seconds). This was repeated until an accurate reproducible effort was achieved. The main role of this spirometry was for safety rather than diagnosis. Spirometry performed at the time of full lung function tests was defined as the patients’ formal results (discussed later). The measurement of an FEV\textsubscript{1} prior to exercise testing allowed for breath by breath evaluation of ventilation during exercise to identify when a patient was approaching a possible respiratory maxima (described in detail later).

A 12 lead ECG was attached to the patient with the standard configuration of the 10 electrodes (Figure 2.6). The limb electrodes were, when necessary, moved if the pedalling motion led to noise artefact.

Continuous 12 lead ECG monitoring occurred throughout the test (Stress PC ECG Application, Version 5.0.522, COSMED S.r.l. Rome, Italy). The ECG monitoring also performed the action of measuring and transmitting heart rate data to the gas exchange analysis software (PFT Ergo, COSMED S.r.l. Rome, Italy).

Continuous arterial oxygen saturations were measured through one of two non-invasive mechanisms; either a finger or an earlobe pulse oximeter. These measure the change in absorbance of two different wavelengths of light passed through either the finger or earlobe to a photodetector on the other side. They exclude interference from non-pulsatile sources such as venous blood and fat.

Finally a non-invasive blood pressure cuff was attached to the right arm of each subject. Blood pressure was measured manually by one of the operators using a sphygmomanometer and stethoscope with the arm in the fully extended position after removal from the ergometer handlebars. Both systolic and diastolic blood pressure readings were recorded at rest. A blood pressure of >200/100mmHg at rest led to cancellation of the test and recommendation of a review with a physician.
Figure 2.5: Software visualisation of spirometry. The curve on the left is a flow-volume curve with expiration above the x-axis. The maximum values for volume and flow as well as the pattern can be distinguished. In the right hand plot volume in plotted against time. A plateau in volume can be seen to occur indicating full expiration.
Figure 2.6: Positioning of ECG electrodes on the subject’s chest. A traditional 12 lead ECG positioning was used with 4 limbs electrodes and 6 chest electrodes.
2.3.1.4 Testing

Each subject underwent at least two CPX tests on the bicycle. Each was given instructions on how to perform the test, following which each test started with 3 minutes of complete rest once I was happy the patient was in a true resting state ($\dot{V}O_2$ approximately 3.5 mL/min/kg, RER 0.70-0.85). The patient was advised not to talk (during this and any subsequent stage of the test) unless necessary (for example if they were experiencing pain). Following this 3 minutes I prompted the patient to begin cycling at a cadence of 55-64 rpm. This cadence (ideal for the ergometer to control work rate) was visible to the patient and the operators and gentle encouragement was given to ensure this cadence was achieved and sustained. The patient continued to cycle, initially through a 3 minute “unloaded” phase where the cycle did not offer resistance. In keeping with most hospital based ergometers, this did not have a flywheel, so was not fully unloaded. However the energy requirements to turn the pedals at zero watts is low and so the term unloaded will be used throughout the thesis. Following on from this initial 3 minutes of exercise, the ergometer started to increase load. The patient was prompted that it would get more strenuous to pedal but they should not alter their approach (i.e. a cadence of 55-64 was maintained throughout). The work rate continued to increase at a linear rate determined prior to the test, in increments of 12-15 seconds. Encouragement was given to the patient regularly to ensure they exercised as close to their maximum as possible.

Blood pressure was measured manually, in the same manner as pre-test, at least every 3 minutes, and as close to peak exercise as possible, an exercise reading of >240/120 led to immediate cessation of the test.

The operators continuously monitored both the ECG and the principal gas analysis variables in real time to identify any problems early and avoid any adverse clinical events. Standard criteria for cessation of a test were adhered to (Balady et al 2010). Unless an arrhythmia was present the exercise test was not stopped due to heart rate, and values over 100% predicted by the patient’s age were allowed.

Unless an abnormal clinical finding occurring during testing (i.e. arrhythmia, ischaemic ECG changes, respiratory distress) prompted me to stop the test, the patient was in control of when to stop, however an inability to maintain a cadence of at least 50rpm despite encouragement did lead me to advise that a maximum had been achieved.

Following cessation of the test the patient entered an active recovery phase. The work rate on the bike reverted to zero immediately, and the patient, following a few seconds rest, was encouraged to begin cycling again, this time at whatever cadence they felt comfortable at. The purpose of active recovery was to allow for continued
flow of blood through the exercising muscles but at a rate below the anaerobic threshold, to washout the metabolites such as lactate to reduce muscular discomfort later; and secondly to allow for a slower, more controlled fall in blood pressure. Gas analysis was measured for 5 minutes of recovery, at which point the test was stopped and the patient disconnected (unless variables such as ECG, heart rate, BP required a longer period of monitoring for safety purposes).

Each patient’s first test was performed on a 10 watts/minute ramp. This, following the 3 minute unloaded section, increased work rate of the ergometer by 2 watts every 12 seconds, or 10 watts per minute, linearly.

The purpose of the first test was two-fold:

1. To familiarise the patient with the test. CPX and cycling are not commonly performed and it has been suggested that there may be a learning effect to this test. A study on the effect of familiarisation showed a significant 0.8mL/min/kg increase in peak $\dot{V}_{O_2}$ on test 2 in heart failure patients on treadmill exercise (Elborn et al 1990). These results were in conflict with a reproducibility sub-study of the HF-ACTION trial, where 88% of patients exercised on the treadmill, and the remainder on the bicycle, which did not show a familiarisation effect between tests 1 and 2 for patients with CHF in peak $\dot{V}_{O_2}$ (Bensimhon et al 2008). However, in this study other variables, such as exercise time, $\dot{V}_E/\dot{V}_CO_2$ slope, peak HR and peak RER did show improvements in the second test. Peak $\dot{V}_{O_2}$ was similarly unchanged between serial treadmill tests in a further study (Russell et al 1998), although exercise time was prolonged in the second test compared to the first, which was interpreted by the authors to show evidence of a learning effect for greater efficiency on the treadmill without increases in oxygen consumption. Marburger et al and Meyer et al also showed no significant differences between test 1 and test 2, however all of the patients in these studies had either previously undergone CPX, or were given a practice test prior to the first recorded test (Meyer et al 1997, Marburger et al 1998). The discrepancies noted between these studies are difficult to explain. The original study by Elborn et al was small, and much earlier than the others, perhaps at a time when familiarity with treadmills *per se* was uncommon, especially amongst a group with an average age of 69. It may be possible therefore that it was not a lack of familiarisation with CPX so much as a lack of familiarisation with walking on a treadmill. This hypothesis would be difficult to confirm without the authors having specifically asked the study subjects their experience with treadmills.
What is relatively uncertain is whether the familiarity effect applies to bicycle ergometry. Bensimhon et al did not report whether a difference in the reproducibility results between the 88% on the treadmill and the 12% on the bicycle existed or not. Because the impact on cycle exercise is less well known it was felt that a familiarisation test would be beneficial to the current study.

2. To identify the optimal protocol. It is largely accepted (although this advice appears to be largely experiential rather than evidence-based) that exercise tests should aim for maximal effort to be obtained by approximately 10 minutes (Balady et al 2010). Significantly longer tests are believed to lead to undue fatigue and boredom and may lead to cessation prior to a physiological maximum. Shorter tests may lead to a lack of data and difficulty measuring certain variables such as the anaerobic threshold. Every patient’s first test was performed with incremental work of 10W/minute. This was found in practice to be a common appropriate protocol for many patients, and had the mathematical advantage that the number of minutes performed signified the incremental protocol required to perform 10 minutes of work (e.g. a patient pedals for 15 minutes achieving 150watts; the ideal protocol for him would be 150W/10 minutes = 15W/min, which was the time achieved). Protocols of 6, 8, 10, 12, 15 and 20W were felt to be sufficient to allow for the spectrum of exercise capacities. Where a patient’s ideal protocol lay between two of these values the upper protocol was chosen, as it is generally believed that a slightly shorter, more intense ramp is superior to a longer one.

The second test was conducted identically to the first test, except that the incremental section may have been at a different increase in work rate, i.e. 6, 8, 12, 15 or 20 rather than 10.

The second test was performed at least 2 hours after the first test to allow for recovery. In some patients these two tests were performed on different days (either due to fatigue or time constraints) but these patients were not chosen randomly, so in some analyses this as a potential confounding variable will be considered.

2.3.1.5 Data analysis and key variables

Each test was performed using the same software package (Quark CPET, COSMED) with identical default settings of measurement. Following each test’s cessation the data was outputted in two formats, firstly an .xpo file, to be used for analysis within the software, and an .xls format for further analysis within the spreadsheet package Excel (Microsoft Corporation). All analysis was performed in Excel. An example patient’s data is included in Appendix 1 (Chapter 12) to show graphically how it was handled. The data were recorded breath-
by-breath, so that for every test there were many rows of data (typically between 200-400 rows). There were multiple variables generated as well displayed as columns, some directly measured such as oxygen uptake ($\dot{V}_{O_2}$) and minute ventilation ($V_e$), and others derived from multiple variables. No averaging was performed initially.

Initially for each test the row numbers were identified that corresponded to the four stages of the exercise test; rest, unloaded cycling, incremental exercise and recovery. This allowed calculation of certain variables to only contain information within a certain portion of the test, for example if incremental exercise occurred between rows 100-300, then the OUES was calculated using only data between and including these two rows.

Peak $\dot{V}_{O_2}$, peak minute ventilation ($V_e$), peak respiratory frequency ($R_f$) and peak $O_2$ pulse, were all calculated in the same manner. The pre-determined limits for peak exercise variables were 20 second averages within the final minute of exercise and including 10 seconds of recovery. Rolling sequential averages were manually calculated starting 40 seconds before cessation of exercise to include all breaths within the preceding 20 second period as displayed in Figure 2.7. I have no knowledge of an identified ideal average time span within the literature; there will be a compromise between high precision but high variability with very short averaging intervals (or no average at all) with the low precision but low variability with very long intervals. Guidelines suggest averaging of data between 20-30 seconds (Balady et al. 2010). I arbitrarily chose 20 seconds with rolling averages largely based on clinical experience of how exercise data is reported. Although there is some evidence that in heart failure delayed attainment of peak exercise ventilatory variables may occur up to 45 seconds after the end of exercise (Cohen-Solal et al. 1997), practically this is not yet part of clinical guidance, and certain CPX software only allow for detection of peak variables during exercise itself.

The raw data was also assessed to ensure that there were no spurious readings. Peak $\dot{V}_{O_2}$ was defined as the highest 20 second rolling average of $\dot{V}_{O_2}$ measurements throughout the final minute of exercise. $\dot{V}_e$ and $R_f$ in analysis were the highest 20 second rolling average of $V_e$ and $R_f$ measurements throughout the final minute of exercise. The oxygen, or $O_2$, pulse is defined as the $\dot{V}_{O_2}$/ heart rate. As $\dot{V}_{O_2}$ is to cardiac output, so the $O_2$ pulse is to stroke volume. It is the product of stroke volume and $[CaO_2 - CvO_2]$ (where $[CaO_2 - CvO_2]$ equals the oxygen content difference between the arterial and venous blood) and is often considered as a surrogate for stroke volume (because oxygen extraction at the muscle is said to be similar between adults, in my opinion a false assumption). I measured the $O_2$ pulse on a breath-by-breath basis, and then recorded the highest 20 second rolling average of measurements throughout the final minute of exercise (i.e. it was not the ratio of peak $\dot{V}_{O_2}$ to
peak heart rate). Unfortunately a limitation of the ECG software I use was that heart rate was not always accurately measured in the presence of a broad QRS complex or pacing spikes. In patients where heart rate measurement was unreliable, the O₂ pulse was not calculated to avoid error.

Maximal voluntary ventilation (MVV), is the predicted maximum ventilation in a minute for a given patient. When directly measured the patient is encouraged to ventilate as hard and fast as they can for a set period of time (often 12 seconds) and the corresponding total volume then corrected for one minute; this is an uncomfortable manoeuvre and technically challenging. Alternatively a patient is asked to perform a forced expiration as described in section 2.3.1.3. The volume expired within the first second (FEV₁) is multiplied by a coefficient (I used 40) to allow an indirect calculation for the MVV (Campbell 1982, Hansen et al 1984). For example a patient with an FEV₁ of 2.5L will have an MVV of 100L/min. The formula to calculate the indirect MVV for my patients was therefore: MVV=FEV₁ x 40.

Knowing the MVV allows you to calculate the breathing reserve (BR), which shows the remaining potential for the lung to increase ventilation. It is typically greater than 30% in healthy adults, which means that the lungs still have 30% of their ventilatory potential remaining, showing the redundancy that exists within our lungs to perform exercise. They are not the organ maximally stressed upon typical exercise.

The BR was calculated using the following formula: \[ BR = 100 \times \left( \frac{MVV - \dot{V}_E}{MVV} \right) \]

For safety during the test the FEV₁ measured during spirometry performed immediately prior to the test was used (this automatically calculated and displayed BR on a breath by breath basis to allow the operator to view). However for the purposes of the analyses the FEV₁ obtained at the time of full lung function tests was used (discussed in section 2.3.3).
Figure 2.7: How peak data was manually averaged in Excel. Data from the final 40 seconds of exercise and the first 10 seconds of recovery have a rolling average of the preceding 20 seconds (as can be seen being calculated for one 20 second interval in the top panel) for each breath during this time period. The highest 20 second average for $V_{O_2}$, minute ventilation, breathing reserve, respiratory frequency and $O_2$ pulse are highlighted.
The final variable calculated using averaged raw data was the oxygen uptake efficiency plateau (OUEP) which was in the original paper describing it defined it as the highest 90 second average of the $\dot{V}_{O_2}/\dot{V}_E$ ratio (Sun et al 2011). In line with this study, which showed the strong predictive value of OUEP in the prognosis of heart failure patients, I also defined the OUEP as the highest 90 second average. OUEP was plotted for each patient against time so that a region of interest where the OUEP was highest could be identified, and the peak value (and that it was correctly averaged over 90 seconds) could be ensured manually.

Following on from the identification of these peak variables, all strictly defined in their calculation to eliminate observer bias/error, I calculated the only 2 variables in each test that were assessed subjectively. These were the anaerobic threshold, and the ventilatory compensation point. As discussed within Chapter 1, exercise can be divided into phases based on the responses to the production and elimination of lactate. The first stage, aerobic exercise, is a period where lactate does not form (or more correctly only at a very low basal level, easily managed by the body). Following this, as exercise becomes more intense, lactate production increases, but the hydrogen ions produced are buffered by bicarbonate within the system, yielding an increase in carbon dioxide. This is the second phase, anaerobic isocapnic buffering, and the anaerobic threshold (AT) delineates this phase from the aerobic phase. With increasing intensity of exercise the bicarbonate in the system is overwhelmed, ventilation must then increase relative to carbon dioxide to bring about a respiratory alkalosis to compensate for the metabolic acidosis and maintain body pH as close to physiologic neutral (7.35) as possible. This is the third phase, or respiratory compensation, and the ventilatory compensation point (VCP) delineates this from the second phase. The ventilatory mechanisms that occur at these two thresholds are too complex to discuss in detail here, but put simply the AT is the point at which oxygen consumption and ventilation become “decoupled” but carbon dioxide production and ventilation remain tightly coupled. This can be viewed graphically as an increase in the $\dot{V}_{CO_2}/\dot{V}_{O_2}$ ratio (the RER), a rise in the $\dot{V}_E/\dot{V}_{O_2}$ ratio without a concomitant rise in the $\dot{V}_E/\dot{V}_{CO_2}$ ratio, and a plateau in the end-tidal CO$_2$ with increasing end-tidal O$_2$ partial pressures. However the commonest clinical method for identifying the AT is to use the V-slope method (Beaver et al 1986), which displays the oxygen uptake (x-axis) against the carbon dioxide elimination (y-axis). It should assume the appearance of two straight lines, with the left-most line $< 45^\circ$ (i.e. $\dot{V}_{O_2}$ rising faster than $\dot{V}_{CO_2}$) and the right-most line $> 45^\circ$ (i.e. $\dot{V}_{CO_2}$ rising faster than $\dot{V}_{O_2}$). This is graphically represented in Figure 2.8 where the gradients of the 2 lines can best be observed easily with the addition of a line of identity ($y=x-c$). The intercept of these two lines is the AT.
Figure 2.8: The identification of the $\dot{V}O_2$ at the anaerobic threshold. 2 subjects are demonstrated (left and right). $\dot{V}O_2$ is plotted against $\dot{V}CO_2$ with a line of identity ($y=x-c$). The data should closely approximate 2 lines, one with a gradient $< 45\%$ (flatter than the line of identity) and one $> 45\%$ (steeper than the line of identity). The overlap point is the $\dot{V}O_2$ at the anaerobic threshold. The patient on the right (C & D) managed a maximal test so his $\dot{V}O_2$ starts to trend vertically at the end, where further increases in $\dot{V}O_2$ do not occur but continued increases in $\dot{V}CO_2$ does. This data should be ignored as can be seen in plot D (red line ignores highest points).
The $\dot{V}_{O_2}$ at this value can be directly measured from this plot, and the corresponding breath identified for all further variables dependent on the AT. Practically however, I feel that the AT identification can only be strengthened by the inclusion of the graphs of the RER against time, ventilatory equivalents against time, and the end-tidal partial pressures against time. The equivalent 3 plots from the second patient in Figure 2.8 are shown in Figure 2.9.

Measurements made at the anaerobic threshold include the $\dot{V}_{O_2}$, the $\dot{V}_E/\dot{V}_{CO_2}$ ratio, RER and end-tidal CO$_2$ (these variables will be discussed in detail below).

In clinical medicine this first threshold in changing ventilatory patterns, the AT, is the most important in helping us to delineate disease aetiology and severity. However should exercise continue for long enough, lactate production causes a significant drop in blood pH, beyond the bicarbonate system’s ability to buffer, which as described above causes the lungs to hyperventilate, in order to reduce arterial CO$_2$ concentration (PaCO$_2$), and compensate for the metabolic acidosis with a respiratory alkalosis. This second threshold is termed the ventilatory compensation point or VCP; at this point the linear relationship between $\dot{V}_E$ and $\dot{V}_{CO_2}$ is lost. Whilst in clinical medicine the occurrence of this second threshold is largely irrelevant, one important clinical variable does require its identification. The slope of the minute ventilation to carbon dioxide elimination relationship, the $\dot{V}_E/\dot{V}_{CO_2}$ slope, is traditionally agreed to be measured using data only from the linear portion of the slope (Figure 1.3). This therefore only applies to the portion of the relationship prior to the VCP. I identified for each patient if the VCP occurred, and if so, at which time point (Figure 2.10), to aid in the accurate measurement of the $\dot{V}_E/\dot{V}_{CO_2}$ slope.
Figure 2.9: Identification of the anaerobic threshold using 3 supplementary plots. The plot on the left shows the ventilatory equivalents, the $\dot{V}_E/\dot{V}_{O_2}$ ratio (red) and the $\dot{V}_E/\dot{V}_{CO_2}$ ratio (blue) against time. The middle plot shows the respiratory exchange ratio against time. The final plot shows the end-tidal CO$_2$ (blue) and end-tidal O$_2$ (red) against time. The AT (indicated by the vertical dashed line) can be seen to occur at a time when: the $\dot{V}_E/\dot{V}_{O_2}$ ratio either reaches a plateau with further decreases in the $\dot{V}_E/\dot{V}_{CO_2}$ ratio, or rises with a plateau in $\dot{V}_E/\dot{V}_{CO_2}$; there is a consistent steep rise in RER with time; and end-tidal CO$_2$ reaches a plateau with further increases in the end-tidal O$_2$. The dotted lines in these plots all correspond to the AT identified in this patient in Figure 2.8. All of these conditions may not be met at the same time point in each patient due to other factors that may impact ventilation (for example anxiety) so some interpretation is necessary along with the V-slope method to allow for an accurate estimation of the time point and oxygen uptake at the AT.
Figure 2.10: Identifying the Ventilatory compensation point (VCP). The VCP occurs when the lactic acidosis overwhelms the local buffering capacity of bicarbonate, potentially leading to acidaemia. This requires hyperventilation with respect to $\dot{V}_{CO_2}$, to elicit a “respiratory compensatory alkalosis” and maintain pH close to the physiological range. This leads to a lowering of the arterial CO₂ set point. Graphically the VCP occurs when the $\dot{V}_E/\dot{V}_{CO_2}$ ratio starts to increase towards the end of exercise, as shown in the first plot at 810 seconds, and when $\dot{V}_E$ and $\dot{V}_{CO_2}$ lose their tight linear relationship, as shown in the second plot with a $\dot{V}_{CO_2}$ of approximately 3L/min. The traditional measurement of the $\dot{V}_E/\dot{V}_{CO_2}$ slope will be calculated as the slope of the line of best fit in the right hand plot through the linear portion of the curve, i.e. with data after the VCP ignored.
A potentially important feature of the AT and VCP as variables in this analysis is that they are the only variables where subjective decision-making from the operator is necessary. Hence they may be more susceptible to bias than other markers, which may influence their reproducibility.

I will now discuss how I calculated the variables of the $\dot{V}_E$ to $\dot{V}_{CO_2}$ relationship. The first measure is the $\dot{V}_E/\dot{V}_{CO_2}$ slope as just discussed, which was calculated in 2 ways. As shown in Figure 2.10, $\dot{V}_E$ can be plotted against $\dot{V}_{CO_2}$. Firstly the line of best fit using least squares regression is plotted within Excel for data starting at the onset of unloaded exercise and ending at the VCP; the slope of this line is defined in my analyses as the $\dot{V}_E/\dot{V}_{CO_2}$ slope, or slope 1. Secondly the same method is used but including all data from unloaded exercise until the end of incremental exercise, i.e. ignoring the VCP, I termed this the full slope, or slope 2. There has been some data suggesting that calculating the slope using all data points, rather than censoring at the VCP, has greater prognostic power compared with limiting data to the linear portion (Ingle et al 2007), although a conflicting study showed no difference in prognostic ability when the data included progressively shorter blocks of data from the same exercise test (Arena et al 2003). Conceptually the evidence from Ingle et al is difficult to explain; in healthy adults, ventilation post-VCP (which is an uncommon occurrence in patients with heart failure and COPD) is uncomfortable and often leads to prompt cessation of the test. Greater time spent in this phase will mean a proportionally greater contribution to the slope from this non-linear segment which will increase the magnitude of the slope. Therefore subjects who push themselves harder and go further into anaerobic metabolism in the test will appear to be “less fit” by virtue of having a steeper slope than if they had stopped earlier (this makes it the only variable where values worsen as exercise continues in healthy adults). Differences in protocol may also affect this slope to a greater degree than the slope measured sub-VCP, as they may alter the proportional time spent in this ventilatory compensation phase of exercise. Therefore physiologically, the inclusion of data post-VCP does not make sense when calculating the slope. However, it is possible that patients with heart failure who display a non linear portion to the $\dot{V}_E/\dot{V}_{CO_2}$ slope have a different physiological mechanism than healthy adults to explain these ventilatory patterns. I therefore felt it would be appropriate to calculate both slopes separately and allow independent analysis of both.

The ratio of $\dot{V}_E$ to $\dot{V}_{CO_2}$ was also calculated, at 3 time points. Firstly at the anaerobic threshold and secondly, the ventilatory compensation point, when they occurred. These values were taken as the instantaneous ratios at the time points identified as previously described. Thirdly the lowest $\dot{V}_E$ to $\dot{V}_{CO_2}$ ratio during the test was also
determined; this is termed the $\dot{V}_E/\dot{V}_{CO_2}$ at nadir. This was, simply, the lowest instantaneous ratio at any time point during unloaded and incremental exercise.

The respiratory exchange ratio (RER) was described briefly earlier and is the instantaneous ratio of $\dot{V}_{CO_2}/\dot{V}_{O_2}$ per breath. $\dot{V}_{CO_2}$ alone is rarely described during CPX but through this ratio we are able to see the impact of exercise on the metabolic systems. At rest it should almost equal the RQ, with fluctuations relating to the variable nature of ventilation. At complete rest ventilation normally matches demand, but with awareness (the presence of a facemask or mouthpiece often impacts on this) this tight relationship may be lost with short periods of relative hyperventilation. However as exercise starts RER settles to match RQ, and ventilation adopts a more regulated pattern, before it slowly starts to rise (as skeletal muscle metabolises an increasing proportion of carbohydrate over fats and proteins). At the onset of anaerobic metabolism the RER rises more rapidly (although the rate of this increase is highly variable and dependent on the duration of the exercise) as excess CO$_2$ is produced from buffering, and then later as more hyperventilation expels yet more CO$_2$. Typically RER should reach its highest values at peak exercise (although recovery values are typically even higher). For this study I determined the RER at 3 time points. Firstly all values during the 3 minute rest period were assessed for the lowest value; this was chosen to represent resting RER. Although there are criticisms of picking a single value in this way, there are also conceptual concerns with averaging RER. It is very common during a rest period for subjects to hyperventilate (a totally flat RER profile never occurs), and therefore an average of the RER over these 3 minutes will be falsely elevated. The counter argument is that the RER will fall to levels lower than equilibrium after a period of hyperventilation (as CO$_2$ stores require replenishment, and CO$_2$ elimination reduces relative to O$_2$ uptake at the mouth). However this phenomenon appears to be less pronounced than the elevation in RER with hyperventilation, so whilst not a perfect measure, I feel it is reflecting resting muscle respiratory quotient (RQ) the closest. Secondly RER was recorded instantaneously at the AT (as previously described). Finally the RER was recorded at peak exercise, as the final RER value before the patient stopped exercising.

Gas analysis measures concentrations of oxygen and carbon dioxide in expired gas throughout the whole expiration. Values for $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ therefore use an average of the whole expiration. There is, however, benefit to looking only at the final instant before inspiration as well – this is termed end-tidal. The final portion of expired gas has been within the lung the longest and has had the greatest time to equilibrate with the pulmonary capillary system. The end-tidal partial pressure of oxygen and carbon dioxide in expired gases closely resembles the partial pressures within the arterial system and so is a potential surrogate for arterial values minimising the
requirement for invasive arterial blood sampling. This has a number of unique uses. Probably the most relevant is the identification of hyper or hypoventilation. Hyperventilation leads to loss of CO\textsubscript{2} from the blood stream, so arterial PaCO\textsubscript{2} (partial pressure of CO\textsubscript{2} in the artery) and P\text{ET}CO\textsubscript{2} (end-tidal partial pressure of CO\textsubscript{2} in the expired gas) are both lower than anticipated. The opposite happens in patients who relatively hypoventilate (can be seen in obesity). End-tidal partial pressures are also of use in the diagnosis of intra-cardiac shunts (Barron et al 2012) with a mirror image to the typical response indicative of a right-to-left shunt. The main limitation of end-tidal partial pressures is that they are only a surrogate for arterial values. The close relationship between these values and the corresponding arterial values (typically +2 to −4mmHg) can be lost in patients with respiratory disease and/or ventilation: perfusion mismatch. In these cases (invariably the patients in which we wish to use them) they must be interpreted cautiously. Whilst arterial measurement of O\textsubscript{2} and CO\textsubscript{2} are more accurate, they are time consuming, cannot be done in real-time and are invasive. End-tidal values are calculated routinely with each breath, and can act as a safe, surrogate for arterial values, so long as their limitations are respected. For each CPX test values of end-tidal CO\textsubscript{2} (P\text{ET}CO\textsubscript{2}) and end-tidal O\textsubscript{2} (P\text{ET}O\textsubscript{2}) were measured breath-by-breath. The only measurement recorded for data analysis was the P\text{ET}CO\textsubscript{2} at the anaerobic threshold. This typically represents the highest value and a plateau in P\text{ET}CO\textsubscript{2}.

Heart rate was recorded by the software continuously, but was recorded for analysis every 3 minutes. Resting heart rate was the average of all values within the 3 minute rest period. All other values were single, non-averaged, values on the 3 minute intervals (or the closest reading to this). Blood pressure was recorded as close to every 3 minute interval as possible and recorded alongside this time point (so for example if blood pressure was being taken between 8:40 and 9:10 it would be recorded as the BP at 9:00). A peak blood pressure was recorded alongside peak heart rate, and the product of the systolic blood pressure and the heart rate at this point was calculated as the Double Product (DP). The circulatory power was defined as the product of the peak $\dot{V}_{O_2}$ (highest 20 second average as described earlier) and the peak systolic blood pressure.

Exercise duration including the 3 minutes of unloaded cycling and maximal work rate (in watts) were also recorded, along with the subjects’ principal symptom for stopping the test.

Finally 3 other slopes were measured. Firstly the oxygen uptake efficiency slope (OUES), which was defined as the slope of the line of best fit when $\dot{V}_{O_2}$ is plotted against $\log_{10}$ minute ventilation (logarithmic transformation of $\dot{V}_k$ leads to a close linear relationship with $\dot{V}_{O_2}$) (Baba et al 1996). There remains a point of contention as to the units of OUES. Some believe it unitless, others L per 10-fold increase in minute ventilation. I largely favour the
latter, however it is an unruly unit for tabular and diagrammatic data, so will largely be referred to with no units.

The OUES was calculated in Excel using least squares regression utilising data within incremental data only. A proposed strength of the OUES is its resilience to foreshortening of data; we therefore further calculated the OUES using data between 25-75% of incremental exercise, 0-50%, 0-70% and 0-90%; these periods of exercise foreshortening have commonly been used in studies before. To calculate the time points 25, 50, 70, 75 and 90% a chart for each test was constructed as shown in Figure 2.11.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Corresponding Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>25%</td>
<td>((PE-360) * 0.25) +360 ≥a</td>
</tr>
<tr>
<td>50%</td>
<td>((PE-360) * 0.50) +360 ≤b</td>
</tr>
<tr>
<td>70%</td>
<td>((PE-360) * 0.70) +360 ≤c</td>
</tr>
<tr>
<td>75%</td>
<td>((PE-360) * 0.75) +360 ≤d</td>
</tr>
<tr>
<td>90%</td>
<td>((PE-360) * 0.90) +360 ≤e</td>
</tr>
</tbody>
</table>

**Figure 2.11: How submaximal OUES limits were constructed.** PE indicates peak exercise time in seconds. 360 seconds (total time for rest and unloaded exercise) was subtracted and then the 25th, 50th, 70th, 75th and 90th percentiles of this “incremental exercise time” were calculated. 360 seconds was then added to these percentile times so that the corresponding time within the whole exercise test for each percentile was found. The corresponding row for that time point within the excel sheet was then identified for each of the 5 percentiles. OUES25-75 was then calculated as the slope of data between rows a and d, OUES50 was calculated as the slope of data between onset of incremental exercise and row b, OUES70 between onset of incremental exercise and row c and OUES90 between onset of incremental exercise and row e.
The next slope to be calculated was the oxygen uptake-work rate slope ($V_{\dot{O}_2} - WR$ slope). $V_{\dot{O}_2}$ rises linearly with work, and the slope of this relationship approximates 10 mL/min O$_2$/Watt in healthy subjects (Wasserman et al 1975, Hansen et al 1984). It was calculated as the slope of the line of best fit (using least squares regression) in Excel when $V_{\dot{O}_2}$ in mL/min (y-axis) is plotted against work rate in watts (x-axis). I arbitrarily chose to measure the slope including data from 10W upwards. Work rates below 10W were not subjectively distinguishable from unloaded cycling when subjects were asked (Chapter 5 goes into this in more detail), and so data prior to this time point was ignored. The effect of the delay from the onset of incremental exercise to the start of measuring the slope (termed the physiologic time constant) has been previously examined (Hansen et al 1987) with the recommendation that the slope is not measured until at least 35 seconds of incremental exercise has passed. My choice of 10W was loosely based upon this recommendation.

Finally we calculated the heart rate – oxygen uptake slope ($HR - V_{\dot{O}_2}$ slope) as the slope of the line of best fit (by least squares regression) in Excel when heart rate in bpm (y-axis) is plotted against $V_{\dot{O}_2}$ in mL/min (x-axis) using data from incremental exercise only. The y intercept for this line of best fit was also recorded.
2.3.2 Echocardiography

Each study subject for the Observational and Interventional studies underwent transthoracic echocardiography. This was for the purpose of identifying severity of left ventricular dysfunction or mitral regurgitation in the cardiac groups, and ensuring there was no undiagnosed structural cardiac disease in the respiratory group.

For each patient the IE33 imaging system was used (Philips, Amsterdam, The Netherlands). A S5-1 probe, with a 1-5 MHz extended operating frequency range, was used for all cases, and all echocardiographic studies were conducted by me in keeping with the British Society of Echocardiography guidelines (Chambers et al 2010). A three lead ECG (corresponding to leads I, II and III) was attached to each patient. Main gain settings were optimised for endocardial border delineation. Each subject underwent the same echocardiographic study, however echocardiography is highly dependent on subject characteristics and not every view and measurement was possible in every patient. Principal measurements (highlighted later) were performed in triplicate and a mean of these 3 values recorded.
2.3.2.1 Views

Standard echocardiographic views were attempted in every patient. These are shown below:

Parasternal long axis (PLAX) – Probe placed on the left sternal edge pointing towards the right shoulder. The left ventricular dimensions can be measured here.

Tricuspid inflow view – Probe tilted inferiorly to view the tricuspid valve, right atrium and right ventricle. Tricuspid regurgitation can be measured here.

Right ventricular outflow – Probe is tilted superiorly to see the pulmonary valve and pulmonary artery.

Parasternal short axis (PSAX) – From PLAX the probe is rotated 90°, then tilted to different levels.

Apical 4 Chamber (A4C) – Probe is placed where the cardiac apex is located, pointing superiorly. Both ventricles and atria can be seen in this view.

Apical long-axis (A2C) – Probe is rotated approximately 60° anticlockwise from A4C to see the left ventricle and atrium in another plane.
Other views included:

Apical 5 Chamber (A5C) – Probe is tilted anteriorly from A4C to open up the left ventricular outflow tract (the so-called 5th chamber).

Apical 3 Chamber (A3C) – Probe is rotated approximately a further 30° anticlockwise from A2C to see the ventricle in a further plane.

Subcostal 4 Chamber view – Probe is placed on the abdomen looking superiorly to the heart. Intra-atrial septum visualised well in this view.

Subcostal IVC view – The probe is rotated to view the IVC as it passes through the liver to the right atrium.
2.3.2.2 Two-Dimensional measurements

Left ventricular dimensions – In the PLAX view the septum, internal diameter of the left ventricle and the posterior wall were measured at the level of the mitral valve tips in both diastole and systole (i.e. at the maximal and minimal sizes for internal diameter). Fractional shortening was then calculated as the difference between internal diameter in systole and diastole divided by the diameter in diastole.

Left ventricular volumes and areas – In the A4C and A2C views the endocardium was traced at the end of diastole and maximal systole to measure volume and area (Figure 2.12). A technique called Simpson’s biplane method calculated volumes based on the assumption of small disks stacked.

Left atrial dimensions – The left atrial diameter was measured in the PLAX view at the beginning of diastole. Left atrial area was measured in the A4C view along with left atrial length, also at the beginning of diastole, to calculate left atrial volume.

Right ventricular dimensions – In the A4C view the right ventricular dimensions were measured at the level of the tricuspid annulus, at the level of the papillary muscles and the length from apex to tricuspid at the end of diastole.

Right atrial dimensions – Right atrial area was measured at the beginning of diastole in the A4C view along with atrial length, to allow for the calculation of right atrial volume.

Left ventricular outflow dimensions – 3 measurements were made at the beginning of systole; the left ventricular outflow tract (LVOT) diameter, aortic valve annulus diameter and the sinus of Valsalva diameter. These measurements were all made in the PLAX view with zoom focused on the left ventricular outflow region.

TAPSE – The tricuspid annular plane systolic excursion is a simple measure of right ventricular systolic function. In the A4C view a single one-dimensional view down through the right ventricle and the lateral tricuspid annulus is imaged against time. This is termed M-mode echocardiography, and was one of the original modalities. The movement towards or away from the transducer for points along a single line can be measured. When correctly aligned the annulus of the tricuspid valve can be seen to move towards the transducer in systole. The distance moved relates to systolic function, with values over 1.6 cm considered normal.

Inferior vena cava (IVC) – This is imaged in the subcostal view. The IVC was viewed throughout at least 1 full respiratory cycle for assessment of inspiratory collapse. The diameter of the IVC was measured at its maximum and minimum, corresponding to expiration and inspiration.
Figure 2.12: Measurement of ventricular volumes using the Simpson’s method. The endomyocardial border is traced around at end-diastole (left) and end-systole (right). Using a model based on a series of stacked disks the volume within the cavity is calculated. Because the left ventricle is not symmetrical (i.e. the same width in all planes) it is recommended that this process is also performed in a plane perpendicular to this one, such as the A2C view. This Simpson’s biplane method will combine these two views to calculate the end-diastolic and end-systolic volumes, and from them an ejection fraction.
2.3.2.3 Standard Doppler measurements

Doppler echocardiography uses the Doppler principle to identify blood flow, and the direction and velocities of this within the heart. The Doppler effect describes how the wavelength or frequency of a wave is altered when there is relative motion between the source/receiver of the wave (these are the same for echocardiography), and the wave reflector (red blood cells). The change in frequency (or wavelength) is termed the Doppler shift and relates to blood flow velocity in the following way, termed the Doppler equation:

\[ \pm \Delta f = \frac{2 f V \cos \theta}{c} \]

Where \( \Delta f \) = change of frequency; \( f_t \) = transducer frequency; \( V \) = velocity of blood flow (unknown), \( \Theta \) = incident angle between the beam and direction of blood flow; and \( c \) = velocity of sound in soft tissue (assumed to be 1540 m/s).

Measuring the velocities of blood is important as it allows us to calculate the difference in pressure between two locations within the heart. Through the Simplified Bernoulli equation pressure difference between two locations relates to the velocity of blood passing between them in the following way: \( \Delta \text{Pressure} = 4 \times (\text{Velocity})^2 \)

Doppler derived measurements included:

Transmitral filling pattern – Blood typically fills the left ventricle during diastole from the left atrium in 2 movements, an initial passive filling denoted as the E wave, and atrial contraction denoted as the A wave. A sample volume is placed at the tips of the mitral leaflets in the A4C and pulsed wave Doppler displays the Doppler signals coming from this area of interest to measure the E and A waves. The time taken for the E wave to return to the baseline, the deceleration time, was also recorded (Figure 2.13).

Aortic valve flow – Continuous wave Doppler of the LVOT is sampled in the A4C view. Unlike Pulsed wave Doppler, Continuous wave Doppler samples all points along a line allowing for identification of maximum velocity, but without identifying where that maximal velocity has originated from (range ambiguity). This generates an aortic flow pattern. Maximal and mean velocities (and therefore pressure gradients) can be measured. By tracing around the signal a velocity-time integral is measured. This is the distance that blood passes in one cycle, and when multiplied by the cross-sectional area of the structure, a volume of blood passing that point per cycle is calculated. A similar flow profile is measured using Pulsed wave Doppler with the sample volume placed in the left ventricular outflow tract (LVOT), proximal to the aortic valve. The two profiles along with the LVOT diameter measured in the PLAX view allow for calculation of the aortic valve area.
Figure 2.13: Transmitral filling pattern. Following optimisation of the signal to ensure a clean spectral trace, the pattern of two distinct waves can be seen, both moving towards the subject’s cardiac apex (transducer position in the A4C view). The E wave occurs following systole. The A wave occurs shortly before the following systolic phase. The peak velocities of both waves are measured and displayed on the image. The line of descent of the E wave is drawn to the baseline to measure the deceleration time.
Pulmonary valve flow – Pulsed wave Doppler of the right ventricular outflow tract (similar to the LVOT Pulsed wave Doppler) is viewed in the PSAX or the right ventricular outflow tract view of the PLAX view. The VTI can be measured, and specific to this profile, the time from onset to peak velocity is measured. This is termed the Pulmonary Acceleration Time and values <100 ms have been shown to identify patients with elevated pulmonary pressures.

Mitral and Tricuspid regurgitation – Small, physiological quantities of regurgitation through these 2 valves is not uncommon. A continuous wave Doppler signal can be identified which measures the pressure difference between the ventricle and the atrium in systole. If atrial pressure is known this allows calculation of the ventricular pressure. This is principally used to calculated pulmonary pressure (working with the assumption that right ventricular systolic pressure equals pulmonary systolic pressure is the absence of pulmonary stenosis which is a very rare condition). IVC vessel calibre and respiratory variation predicts right atrial pressure, so together with tricuspid regurgitation velocity, pulmonary artery systolic pressure can be accurately estimated.

2.3.2.4 Tissue Doppler measurements (TDI)
In the same way that Doppler can measure blood velocities, it can also measure the velocity of tissue. Filters are introduced to exclude high velocities seen with blood so that only the lower velocity tissue is seen. Using pulsed wave Doppler, we can see movement of the myocardium throughout the cardiac cycle, however it is highly dependent on the quality of the echocardiographic view of the region of interest. The TDI measurements typically obtained are: e’, a’, s’ waves at the lateral and septal walls – These 3 waves are measured with the sample volume of the Pulsed wave Doppler placed on the lateral and septal aspects of the mitral annulus in the A4C view. The e’ and a’ waves correspond to the timings seen with the transmtral filling pattern, and are directed away from the transducer. The s’ wave corresponds to systolic motion, and is towards the transducer (Figure 2.14).

Right ventricular S’ – Similarly to the left ventricle, the sample volume is placed over the lateral tricuspid annulus to view movement of the right ventricle. Only the S’ wave is of significant interest when interrogating the right ventricle, it represents systolic function of the RV and is towards the transducer.
Figure 2.14: Tissue Doppler imaging of the septal wall at the mitral annulus. The e’(†), a’(‡) and s’(‡) waves can be seen throughout the cardiac cycle and correspond to early passive diastolic filling, late diastolic filling from atrial contraction and systolic contraction respectively.
2.3.2.5 Specific measurements for mitral regurgitation

Traditionally the determination of MR severity has largely been a qualitative exercise, where the operator views the size and turbulent nature of a regurgitant jet, the strength of the regurgitant signal relative to the forward signal, and the impact on the LV and LA. More recently semi-quantitative and quantitative methods for severity determination have become popular, including jet area: left atrial area ratio, vena contracta width, and the proximal isovelocity surface area (PISA) technique. For jet area measurements the atria and the regurgitant jet at its maximum are traced to give two corresponding area measurements. The vena contracta is the narrowest diameter of the regurgitant jet as it passes through the mitral valve.

In contrast to these simple measurements the Proximal Isovelocity Surface Area (PISA) method, first described for MR over 20 years ago (Recusani et al 1991), is a complex calculation requiring multiple measurements. How PISA calculates mitral regurgitant volumes, fractions and the regurgitant orifice area is well described elsewhere (Simpson et al 1995) and uses the principles of the continuity equation and conservation of mass.

Flow has been noted to accelerate towards a regurgitant orifice, a so-called convergence zone, and has the appearance of a hemisphere. The diameter of this hemisphere and the velocity of blood at its outer limits (which is the velocity at which the echocardiogram cannot distinguish direction of flow – the aliasing velocity - and is clearly shown on the display) allows for calculation of the flow rate (Figure 2.15). The flow rate through the regurgitant orifice must be the same (conservation of mass), thus generating the formula:

\[ 2\pi r^2 \times V_N = EROA \times V_{max} \]

Where \(2\pi r^2\) = surface area of the hemispheric shell,
\(r\) = diameter of this shell;
\(V_N\) = velocity at the radius of the shell, the aliasing velocity;
\(EROA\) = effective regurgitant orifice area (unknown) and \(V_{max}\) = maximal velocity of regurgitant blood through the mitral valve measured by Continuous wave Doppler.

If the VTI of the regurgitant jet is also calculated then regurgitant volume can be calculated:

\[ RV = EROA \times VTI_{Rj} \]

Where \(RV\) = regurgitant volume, and \(VTI_{Rj}\) = VTI of the regurgitant jet.

Finally if the LVOT VTI and diameter were measured then the regurgitant fraction (proportion of blood ejected from the left ventricle that goes “backwards”) can be calculated:
Severity of mitral regurgitation as assessed by these mechanisms was graded as per the British Society of Echocardiography guidelines.

\[ RF = \frac{RV}{RV + SV_{LVOT}} \times 100 \]

Where \( SV_{LVOT} \) = stroke volume through the LVOT and is calculated using the VTI and diameter.

**Figure 2.15: The calculation of mitral regurgitation severity via the PISA method.** The baseline of the colour Doppler scale is moved in the direction of the regurgitant jet. The resultant low velocity (seen circled on the right) allows aliasing of the colour jet proximal to the valve (on the ventricular side as seen here in the A4C view). A hemisphere is seen where there is a clear distinct edge of colour aliasing (yellow into blue). The distance from the valve tips to the edge of this hemisphere is measured (upper circle). Along with knowledge of the mitral regurgitant VTI and LV forward stroke volume (calculated using the LVOT VTI and diameter) regurgitant volume, fraction and effective regurgitant orifice area can all be calculated.
2.3.2.6 Specific measurements for mitral stenosis

2 measurements, proven to correlate to orifice area in mitral stenosis were measured. The VTI of the transmitral flow in the A4C view was measured using continuous wave Doppler (similar to the measurement of the E and A waves). This gives a mean pressure gradient, which increases with worsening stenosis. The pressure half time (PHT) of the slope of the E wave of this Doppler also accurately estimates mitral valve area via the following equation: mitral valve area = 220/ PHT.
2.3.3 Pulmonary function tests

Full lung function testing was performed on the Spiro Air (Medisoft, Sorinnes, Belgium), and included spirometry, diffusion and full lung subdivisions. These tests were all performed by an experienced and trained respiratory physiologist at St Mary’s Hospital Chest and Allergy Clinic. Following machine calibration 3 groups of tests were performed on each patient: Spirometry, full lung subdivisions and diffusion capacity.

2.3.3.1 Calibration

At the beginning of every day, and after every change of pneumotachograph, calibration was performed on the equipment. Similarly to the CPX machine and software, flow is calibrated using a 3L syringe. Secondly the carbon monoxide-helium analyser performs an automatic calibration, measuring a known concentration of both these gases found in a standardised cylinder, and then using as a second reference gas, air (where the concentration of both these gases will be zero).

2.3.3.2 Spirometry

Flow is measured using a pneumotachograph, and volume measured by true volume displacement. This measures a pressure difference across a multitude of small bore tubes which act to control resistance and direct flow in a laminar manner, so that the resulting pressure difference can be used to calculate flow (and therefore volume) using the Hagen-Poiseuille formula:

\[ Q = \frac{P \pi r^4}{8 \eta l} \]

Where \( P \) = pressure drop, \( r \) = radius of the tube, \( \eta \) = viscosity, and \( l \) = length of the tube.

The variables recorded during spirometry include the Forced Expiratory Volume in 1 second (FEV₁), Forced Vital Capacity (FVC), the ratio of these two variables (FEV₁:FVC ratio), the Peak Expiratory Flow rate (PEFR), and the Maximum Expiratory Flow when 50% of the FVC has been exhaled (MEF₅₀).

2.3.3.3 Full lung subdivisions

The closed circuit helium (He) dilution method (with CO₂ absorption and O₂ compensation) is used to measure subdivisions. Following a short period of stable tidal volume breathing, the subject is connected via the spirometer through the closed circuit which contains a known concentration and volume of helium. This reaches
equilibrium within the lungs (and does not cross the alveolar membranes) and settles at a new concentration allowing calculation of the lung volumes by the following equation:

\[ SD = \frac{C_1 \times V_1}{C_2} - V_1 \]

Where SD = subdivision, C1 = Initial He concentration, C2 = Final He concentration, and V1 = volume of gas in the spirometer.

Initially the subject continues to breathe to their tidal volume and the concentration can be measured at the point of a normal expiration (functional residual capacity, FRC). Maximal inspiratory and expiratory efforts then allow calculation of total lung capacity (TLC) and residual volume. The full breakdown of lung volumes is shown in Figure 2.16.

### 2.3.3.4 Diffusion capacity

The diffusion capacity or transfer factor of the lung for carbon monoxide (D\(_{LCO}\) or T\(_{LCO}\)) is a measure of how well oxygen passes from the alveoli into the blood. D\(_{LCO}\) and alveolar volume (V\(_A\)) were measured using the single breath helium trace gas method. The subject is attached to a bag collection system and after stable tidal breathing occurs, they are told to take a maximal slow expiration. At this point the test gas is introduced and the subject takes a rapid inhalation to reach, as near as possible, TLC, at which point they hold their breath for 10 seconds. The gaseous mixture is Helium (14%), carbon monoxide (0.28%), oxygen (21%), and the remainder nitrogen. The 10 second pause allows for the carbon monoxide (but not the helium which is termed the tracer gas and is used to measure V\(_A\)) to move from the alveoli into the blood. On exhaling, the first 500mL of gas is discarded (the washout volume) as this constitutes gas that was present in the airways, and was not in contact with the blood. The following volume of gas is termed the sample volume and from this the D\(_{LCO}\) is measured.

Following measurement of the D\(_{LCO}\) it is corrected in 2 ways: firstly for haemoglobin, which was measured on every patient using a capillary blood sample and the HemoCue Hb 201+ analyzer (HemoCue, Dronfield, UK). Secondly, as it is determined by the surface area of the alveolar membranes (V\(_A\)) it is corrected for this to give the transfer coefficient (K\(_{CO}\)) \(K_{CO} = D_{LCO}/V_A\).

Each patient underwent full lung function testing as described above on recruitment to the study. For patients undergoing mitral valve surgery, the tests were repeated at reassessment following the surgery, to identify any changes that may have been caused by the process of surgery. Alternatively the reduction in pulmonary artery pressures following correction of a regurgitant mitral valve may actually improved lung mechanics.
Figure 2.16: The subdivisions of the lung. Most of these subdivisions can be measured using simple spirometry, however RV, FRC and TLC require more advanced tests, such as the helium dilution method.
2.3.4 Venous blood sampling

Each patient underwent sampling of their venous blood on recruitment to the study, and also on the second visit for patients after mitral valve surgery and CRT insertion. I performed all the venous blood sampling for the study.

Following application of a tourniquet to the upper left arm, a visible/palpable vein, typically in the antecubital fossa, was cleaned with a sterile alcohol swab. Then using an aseptic technique a sterile needle attached to a Vacutainer holder (BD Medical, Franklin Lakes, New Jersey, USA) was passed into the chosen vein. 3 vials of blood were then sequentially introduced into the Vacutainer and 2-5mL of blood was used to fill each. The vials collected were:

2x Lavender top – EDTA coated tubes to prevent blood clotting, used for whole blood haematology determinations including haemoglobin concentration, white cell concentration and platelet concentration. The second EDTA tube was used to measure B-type natriuretic peptide (BNP). BNP is an amino acid secreted by the heart in response to chamber stretch, and regulates body salt and water composition. It is a sensitive marker of heart failure, being elevated in patients with decompensation, and to a lesser extent in patients with stable heart failure with a moderate inverse relationship to echocardiographic indices of ventricular dysfunction (Cowie et al 2002, Atisha et al 2004).

1x Gold top – Clot activator and gel for serum separation, used for biochemistry determinations. The specific biochemical recordings made included: creatinine, urea, sodium, potassium, bicarbonate, bilirubin, albumin, alanine transaminase and alkaline phosphatase.

Most of these blood tests have normal ranges that do not differ with patient demographics, but for creatinine, age, gender and race can influence the normal values. Estimated glomerular filtration rate (eGFR) can be calculated using creatinine and these 3 variables via the Modification of Diet in Renal Disease (MDRD) formula, which I performed in each patient.

\[
GFR = 175 \times (Creatinine/88.4)^{1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})
\]
3.0 Retrospective Pilot analysis – the differential influence of spirometry on CPX variables
Heart failure and ventilatory disease often coexist and both cause abnormalities in cardiopulmonary exercise test variables. I analysed the relative reliance of peak $\dot{V}_{O_2}$, the $\dot{V}_{O_2}$ at the anaerobic threshold (AT), the $\dot{V}_{E}/\dot{V}_{CO_2}$ slope and the Oxygen Uptake Efficiency Slope (OUES), upon measures of abnormal respiratory function.

The data of 168 patients (median age 65, IQR 22-87) with clinical heart failure who underwent treadmill CPX testing with spirometry and echocardiography were retrospectively analysed. 43 patients had heart failure with preserved ejection fraction.

Median peak $\dot{V}_{O_2}$ was 16.8 mL/min/kg (IQR 12.0-21.1), $\dot{V}_{O_2}$ at the AT 11.1 (6.9-14.0), $\dot{V}_{E}/\dot{V}_{CO_2}$ slope 44.0 (36.4-55.7), OUES 1.38 L/min/10-fold increase in $\dot{V}_{E}$ (0.97-1.83), and FEV$_1$ 1.78 L (1.43-2.30).

FEV$_1$ was significantly related to peak $\dot{V}_{O_2}$ (3.2 mL/min/kg difference between above- and below-median percentage of predicted FEV$_1$), with a weak relationship to $\dot{V}_{O_2}$ at the AT. FEV$_1$ was not related to the $\dot{V}_{E}/\dot{V}_{CO_2}$ slope or OUES. When patients were separated into normal, obstructive, restrictive or mixed patterns of spirometry peak $\dot{V}_{O_2}$ was the only CPX variable to significantly differ between groups.

Although it is the pre- eminent variable for cardiac dysfunction on CPX testing, peak $\dot{V}_{O_2}$ is very sensitive to abnormalities in spirometry, and is highly dependent on FEV$_1$. In contrast $\dot{V}_{E}/\dot{V}_{CO_2}$ slope and OUES do not show any relation to FEV$_1$. It is possible that the accepted greater prognostic power of $\dot{V}_{E}/\dot{V}_{CO_2}$ slope and OUES, over peak $\dot{V}_{O_2}$, may be in part due to a greater specificity for the impact of cardiac- rather than ventilatory-dysfunction on exercise capacity. This pilot data shows proof of concept that various CPX variables appear to be differentially affected by respiratory disease.
3.2 Introduction

Cardiopulmonary exercise testing (CPX) is an excellent test for distinguishing cardiac from respiratory limitation, but in a patient with both aetiologies of disease it can be difficult to evaluate the relative contributions from both the cardiovascular and respiratory systems. Both cardiovascular and respiratory diseases are common, and commonly coexist (Mascarenhas et al 2008). Abnormal exercise physiology is a central feature of chronic heart failure (CHF) perhaps more so than with respiratory disease, but in both groups quantifying exercise capacity is a vital step in risk stratification and estimation of prognosis; CPX is the gold standard investigation to assess exercise capacity.


It is unclear why certain variables perform more strongly than others, but peak $\dot{V}_{O2}$ can be limited by its dependence on maximal effort. Peak $\dot{V}_{O2}$ is calculated using data from the last portion of exercise; the assessment of cardiovascular function in CHF patients by peak $\dot{V}_{O2}$ may be confounded if a secondary ventilatory restriction becomes the major limiter as peak $\dot{V}_{O2}$ is approached, i.e. the peak $\dot{V}_{O2}$ in this instance does not reflect the impact to exercise capacity from their cardiac limitation.

Variables that have shown increased prognostic power over peak $\dot{V}_{O2}$ include the $\dot{V}_{E}/\dot{V}_{CO2}$ relationship (Chua et al 1997, Robbins et al 1999, Gitt et al 2002), the oxygen uptake efficiency slope (OUES) (Davies et al 2006) and the anaerobic threshold (Gitt et al 2002). A strength of these 3 variables is their greater independence from the degree of effort when compared to peak $\dot{V}_{O2}$. In principle they may also be more independent of abnormal respiratory function, since respiratory limitation typically leads to an early peak in $\dot{V}_{O2}$ but not necessarily with an abnormality in physiology in the lead up to this point.
In this retrospective pilot study I examine how abnormal respiratory function impacts on certain CPX variables in a group of patients with CHF. If ventilatory abnormalities differentially relate to various variables on CPX testing then this may explain the differences in prognostic power, and gives further evidence for the validity of the thesis concept that a CPX variable exists which is specific for cardiac dysfunction.
3.3 Methods

3.3.1 Subjects
All patients undergoing CPX with spirometry and echocardiography as part of outpatient assessment for CHF at St Mary’s Hospital, between July 2003 and June 2007, had their raw data analysed retrospectively. The patients had been referred to the heart failure clinic based on an appropriate history of exercise intolerance or peripheral or pulmonary oedema and it was not necessary to have a reduced left ventricular ejection fraction to be diagnosed with clinical heart failure. CPX was performed for assessment of functional capacity. Echocardiography was performed on all patients at St Mary’s Hospital. Patients with abnormal spirometry were not excluded.

3.3.2 Cardiopulmonary exercise testing
Patients underwent exercise testing on a motorised treadmill in an air-conditioned room after familiarisation with the equipment. Spirometry and gas analysis was performed using the Medical Graphics CardioCP2 which was calibrated before each test (Medical Graphics UK Limited, Gloucester, UK). Percentage predicted values for forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC) were calculated by the spirometer. Breathing reserve (BR) was calculated retrospectively using the formula BR=100 x (MVV-\dot{V}_E)/MVV (%), where MVV was maximal voluntary ventilation, calculated as 40 x FEV₁ and \dot{V}_E was the minute ventilation at peak exercise. They exercised to symptom-limited exhaustion using a smoothed modified Bruce protocol (Scott et al 2001). Peak exercise oxygen consumption was measured in mL/min/kg (except when graphically presented alongside OUES where absolute values in L/min are shown) and defined as the highest 30 second average during the last 60 seconds of exercise, and \dot{V}_E/\dot{V}_{CO_2} slope and oxygen uptake efficiency slope were measured across the full duration of exercise. \dot{V}_{O_2} at the AT was calculated using the V-Slope method.

3.3.3 Echocardiographic assessment of left ventricular function
All patients had undergone a routine transthoracic echocardiogram using conventional 2-dimensional grey scale and Doppler techniques. Patients were imaged using a S5-1, 3.5 MHz transducer (Model 7500, Philips Medical Systems, Andover, Massachusetts, USA). Where possible systolic function assessment was performed using the Simpson’s modified ellipsoid method to compute the end-diastolic and end-systolic volumes in both apical 4-chamber and 2-chamber orthogonal imaging planes. In some patients the Simpson’s method was not possible to
measure left ventricular systolic function; in these instances fractional shortening as measured in the parasternal long axis was used.

### 3.3.4 Statistical analysis

Statistical analysis was performed using Stata version 11.1 for Windows (StataCorp LP, College Station, Texas). Patient demographics and CPX variables were assessed for normality using the Shapiro-Wilk test. If an individual variable was found to be non-normally distributed, a logarithmic transformation of the data was performed and used if it then displayed a normal distribution. If this logarithmic transformation did not render the data normally distributed, the original data was used and non-parametric tests employed where possible.

To assess for the relationship between different CPX variables pairwise correlation coefficients were calculated.

Groups were compared using ANOVA for parametric data and Wilcoxon rank test and Kruskal-Wallis test for non-parametric data. The Kruskal-Wallis test is a non-parametric equivalent to one-way analysis of variance; it is like the Mann-Whitney test where the number of groups ≥ 3, however it is still appropriate when there are only 2 groups. Linear regression was used when a continuous variable was assessed for its impact on another continuous variable.

A p value of < 0.05 was considered statistically significant throughout.
3.4 Results

3.4.1 Patient characteristics

From 481 patients undergoing CPX during the period between July 2003 and June 2007, 168 patients had undergone echocardiography with ejection fraction estimation, and cardiopulmonary exercise testing with FEV\textsubscript{1}, FVC, peak $\dot{V}_{O_2}$, $\dot{V}_{E}/\dot{V}_{CO_2}$ slope, OUES and $\dot{V}_{O_2}$, at the AT measurements. No patient was limited by ischaemia. These 168 patients make up the study cohort.

Age, weight (but not height), body mass index (BMI) and body surface area (BSA) were not normally distributed but the data was not rendered normal by logarithmic transformation. The median age of the study group was 65 years (range 22-87). 126 (75%) were male.

The mean height was 169.8 ± 9.0 cm. Median weight was 76.9 kg (range 46-183.3), BSA was 1.89 m\textsuperscript{2} (range 1.46-2.98) and BMI was 27.3 kg/m\textsuperscript{2} (range 16.9-49.7).

Mean ejection fraction (EF) was measured in 101 patients and was 38.5 ± 15.8%. In the remaining 67 patients the fractional shortening (FS) was 21.9 ± 9.2%. 74.4% of patients had systolic impairment as defined by an EF < 55% or, in the absence of a measureable EF, an FS < 25%. Therefore 25.6% of patients had heart failure with preserved ejection fraction. The aetiology of heart failure was IHD in 49.4% (44 patients had undergone previous coronary artery bypass surgery), dilated cardiomyopathy in 33.9%, valvular disease in 7.1%, hypertension in 5.4%, and other/unknown in 4.2%. NYHA status was Class I in 23%, Class II in 50%, Class III in 20% and not documented in 6%.

25.6% of patients had a low FEV\textsubscript{1}:FVC ratio suggestive of obstructive airways disease. However because I wasn’t specifying the type of respiratory disease, FEV\textsubscript{1} was chosen as the discriminating variable of lung function, as it will be reduced in both obstructive and restrictive disorders. 59.5% of patients had a reduced FEV\textsubscript{1} as defined by < 70% of predicted values. 50 patients had low saturations (SaO\textsubscript{2} < 95%) either at rest or during exercise, of which 26 had a fall of at least 5%. Echocardiographic, CPX and spirometry measures are shown in Table 3.1. Echocardiographic variables were normally distributed so are displayed as mean ± SD. CPX and spirometry measures were not normally distributed so are shown as median, interquartile range. In 2 patients an AT was not determined.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean/ Median</th>
<th>± SD/ 25th-75th percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ventricular diastolic dimension (mm)</td>
<td>54.4</td>
<td>± 9.3</td>
</tr>
<tr>
<td>Left ventricular systolic dimension (mm)</td>
<td>43.4</td>
<td>± 11.1</td>
</tr>
<tr>
<td>Resting ejection fraction (%)</td>
<td>38.5</td>
<td>± 15.8</td>
</tr>
<tr>
<td>Heart rate (bpm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resting</td>
<td>79</td>
<td>63-88</td>
</tr>
<tr>
<td>Peak</td>
<td>136</td>
<td>115-156</td>
</tr>
<tr>
<td>Systolic BP (mmHg)</td>
<td>120</td>
<td>110-140</td>
</tr>
<tr>
<td>Peak</td>
<td>150</td>
<td>130-180</td>
</tr>
<tr>
<td>Diastolic BP (mmHg)</td>
<td>80</td>
<td>70-80</td>
</tr>
<tr>
<td>Peak</td>
<td>80</td>
<td>70-80</td>
</tr>
<tr>
<td>RER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resting</td>
<td>0.76</td>
<td>0.69-0.83</td>
</tr>
<tr>
<td>Peak</td>
<td>0.98</td>
<td>0.86-1.06</td>
</tr>
<tr>
<td>Peak $\dot{V}_{O_2}$ (mL/min/kg)</td>
<td>16.8</td>
<td>12-21.1</td>
</tr>
<tr>
<td>Percentage predicted peak $\dot{V}_{O_2}$ (%)</td>
<td>60.8</td>
<td>45.5-74.4</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope</td>
<td>44</td>
<td>36.4-55.7</td>
</tr>
<tr>
<td>$\dot{V}_{O_2}$ at the AT (mL/min/kg)</td>
<td>11.1</td>
<td>6.9-14</td>
</tr>
<tr>
<td>OUES (L/min/10-fold ventilation increase)</td>
<td>1.38</td>
<td>0.97-1.83</td>
</tr>
<tr>
<td>FVC (L)</td>
<td>2.36</td>
<td>1.85-2.96</td>
</tr>
<tr>
<td>Percentage Predicted FVC (%)</td>
<td>65.1</td>
<td>51.9-75.8</td>
</tr>
<tr>
<td>FEV$_1$ (L)</td>
<td>1.78</td>
<td>1.43-2.30</td>
</tr>
<tr>
<td>Percentage Predicted FEV$_1$ (%)</td>
<td>65.8</td>
<td>50.4-79.6</td>
</tr>
<tr>
<td>Quotient FEV$_1$/FVC (%)</td>
<td>81.0</td>
<td>69.0-88.4</td>
</tr>
<tr>
<td>Breathing Reserve (%)</td>
<td>22.0</td>
<td>7.5-39.6</td>
</tr>
</tbody>
</table>

**Table 3.1: Baseline characteristics.** Echocardiographic, spirometric and cardiopulmonary exercise test variables in 168 patients with chronic heart failure. Results are presented as mean ± SD, or median (IQR).
3.4.2 Relationship between exercise variables

Peak $\dot{V}_{O_2}$, $\dot{V}_E/\dot{V}CO_2$ slope, OUES, $\dot{V}_{O_2}$ at the AT and BR were all assessed by pairwise correlation. These variables all correlated with one another, with the exception of peak $\dot{V}_{O_2}$ and BR, and $\dot{V}_E/\dot{V}CO_2$ slope and BR. The correlation coefficients ($r$) for four of these pairs are shown in Figure 3.1.

3.4.3 Effect of ventilatory measures on CPX variables

The influence of FEV$_1$ as a continuous variable, when divided into 2 groups based on the median value, and the pattern of respiratory disease (none, obstructive spirometry, restrictive spirometry and mixed spirometry) on peak $\dot{V}_{O_2}$, OUES, $\dot{V}_E/\dot{V}CO_2$ slope and the $\dot{V}_{O_2}$ at the AT were all assessed. On linear regression the percentage of predicted FEV$_1$ is related to peak $\dot{V}_{O_2}$ ($R^2=0.11$, $p<0.001$) and to a lesser extent the $\dot{V}_{O_2}$ at the AT ($R^2=0.04$, $p<0.01$), but not to the $\dot{V}_E/\dot{V}CO_2$ slope ($R^2=0.02$, $p=0.07$) or the OUES ($R^2=0.02$, $p=0.09$).

The study population was then divided at the median value of the percentage of predicted FEV$_1$ (65.8%) into 2 groups. Peak $\dot{V}_{O_2}$ differed significantly between the groups (Kruskal-Wallis; $\chi^2=9.3$, $p=0.002$) with a median peak $\dot{V}_{O_2}$ of 14.5 mL/min/kg in the group with lower FEV$_1$, and a median of 17.7 mL/min/kg with the higher FEV$_1$, an 18% reduction in the former group. $\dot{V}_{O_2}$ at the AT (Kruskal-Wallis; $\chi^2=3.2$, $p=0.07$), $\dot{V}_E/\dot{V}CO_2$ slope (Kruskal-Wallis; $\chi^2=0.39$, $p=0.53$) and OUES (Kruskal-Wallis; $\chi^2=1.58$, $p=0.21$) were not significantly different between the patients with the lowest and highest predicted FEV$_1$ (Figure 3.2).

Finally the group was divided into four groups: The first group “normal spirometry” included patients with percentage of predicted FVC $>$ 70% and FEV$_1$/FVC ratio $>$ 70% (n=43); the second group, “obstructive”, had a ratio $<$ 70%, and percentage of predicted FVC $>$ 70% (n=19); the third group, “restrictive”, had a ratio $>$ 70% and percentage of predicted FVC $<$ 70% (n=82); and the final group, “mixed”, had a ratio $<$ 70% and percentage of predicted FVC $<$ 70% (n=24). Peak $\dot{V}_{O_2}$ differed significantly between the groups (Kruskal-Wallis; $\chi^2=11.3$, $p=0.01$). $\dot{V}_{O_2}$ at the AT (Kruskal-Wallis; $\chi^2=3.1$, $p=0.37$), $\dot{V}_E/\dot{V}CO_2$ slope (Kruskal-Wallis; $\chi^2=1.28$, $p=0.74$) and OUES (Kruskal-Wallis; $\chi^2=1.97$, $p=0.58$) were not significantly different between the four groups. The median peak $\dot{V}_{O_2}$ was 18.3, 17.6, 15.9 and 12.8 mL/min/kg in the 4 groups respectively (Figure 3.3).
Figure 3.1: Correlation between the 4 CPX variables; peak $\dot{V}_O_2$ (L/min), OUES (L/min/10-fold increase in ventilation), $\dot{V}_E/\dot{V}_{CO_2}$ slope, and $\dot{V}_O_2$ at the AT (L/min). The correlation coefficient for each pair is shown on each plot.
3.4.4 Impact of ventilatory measures on CPX variables in patients not limited by lungs

A breathing reserve of $\geq 30\%$ is typically believed to indicate a test that is not limited by the lungs. 67 patients had a BR $\geq 30\%$ and within this group the percentage of predicted FEV$_1$ remained related to peak $\dot{V}_{O_2}$ ($R^2=0.15$, $p=0.001$) and to the $\dot{V}_{O_2}$ at the AT ($R^2=0.10$, $p=0.01$), but not to the $\dot{V}_E/\dot{V}CO_2$ slope ($R^2=0.03$, $p=0.19$) or the OUES ($R^2=0.06$, $p=0.05$).

3.4.5 Effect of ejection fraction on CPX variables

Similarly to the analyses with percentage of predicted FEV$_1$, the study population was divided into 2 groups based on left ventricular systolic function in 2 ways. Firstly I divided into normal ejection fraction (heart failure with preserved ejection fraction (HF-PEF)) and abnormal (heart failure with reduced ejection fraction (HF-REF)) with a cut-off for normality of $\geq 55\%$. Where an ejection fraction was not directly measured it was calculated using Teichholz formula from the internal dimensions in the parasternal long axis view. 125 patients had a reduced EF and 43 a preserved EF. There were no significant differences in peak $\dot{V}_{O_2}$ (Kruskal-Wallis; $\chi^2=1.9$, $p=0.66$), $\dot{V}_{O_2}$ at the AT (Kruskal-Wallis; $\chi^2=0.63$, $p=0.43$), $\dot{V}_E/\dot{V}CO_2$ slope (Kruskal-Wallis; $\chi^2=2.6$, $p=0.11$) or OUES (Kruskal-Wallis; $\chi^2=0.03$, $p=0.87$) between these 2 groups.

Secondly I divided into 2 groups by median EF (42.5%). There were no significant differences in peak $\dot{V}_{O_2}$ (Kruskal-Wallis; $\chi^2=3.2$, $p=0.07$), $\dot{V}_{O_2}$ at the AT (Kruskal-Wallis; $\chi^2=1.9$, $p=0.17$) or OUES (Kruskal-Wallis; $\chi^2=2.1$, $p=0.15$) between these 2 groups (Figure 3.2). However $\dot{V}_E/\dot{V}CO_2$ slope showed a borderline significant difference, with higher values in the group with the lower EF (Kruskal-Wallis; $\chi^2=4.2$, $p=0.04$).

To assess if EF as a marker of heart failure severity was confounded in this analysis by the presence of patients with HF-PEF, the relation between EF and the 4 CPX variables was assessed only in patients with a reduced EF. On linear regression ejection fraction is not related to any of the 4 variables (peak $\dot{V}_{O_2}$ ($R^2=0.02$, $p=0.14$); $\dot{V}_{O_2}$ at the AT ($R^2=0.02$, $p=0.11$); $\dot{V}_E/\dot{V}CO_2$ slope ($R^2=0.00$, $p=0.88$); OUES ($R^2=0.01$, $p=0.41$)).
Figure 3.2: Box and whisker plot of the impact of FEV₁ (divided into 2 groups by median percent predicted (65.8%)) and ejection fraction (divided into 2 groups by median (42.5%)) on the four CPX variables. The whiskers represent the lowest data point within 1.5 IQR of the 25th and 75th percentiles. Kruskal-Wallis statistic and p values are shown for each comparison.
Figure 3.3: Box and whisker plot of the impact of spirometric abnormalities on the four CPX variables when patients are divided into normal, obstructive, restrictive or mixed pattern. The whiskers represent the lowest data point within 1.5 IQR of the 25th and 75th percentiles. Kruskal-Wallis statistic and p values are shown for each comparison. See text for definitions of the four groups.
3.5 Discussion

In this pilot study, to help establish a proof of concept for the thesis, I took a retrospective cohort of patients with a diagnosis of heart failure (independent of echocardiographic evidence of a low ejection fraction) and assessed the impact of spirometric abnormalities on 4 commonly reported CPX variables; peak \( \dot{V}_O_2 \), the \( \dot{V}_E/\dot{V}_CO_2 \) slope, the \( \dot{V}_O_2 \) at the anaerobic threshold and the OUES. Peak \( \dot{V}_O_2 \), and to a lesser extent \( \dot{V}_O_2 \) at the AT, related to respiratory abnormalities as measured by spirometry, whereas \( \dot{V}_E/\dot{V}_CO_2 \) slope and OUES did not. None of the variables related closely to ejection fraction.

3.5.1 The relation of CPX variables to one another

It is an important concept and central to the hypothesis of the thesis, that CPX variables behave differently from one another and are not just different ways of portraying the same information. If they all closely predicted one another it is unlikely any single variable could outperform the rest when assessing for either prognostic purposes or disease discrimination. Within this pilot cohort OUES and peak \( \dot{V}_O_2 \) were the most closely correlated \((r=0.77)\) suggesting that the value of either could be reasonably predicted from the value of the other one. The \( \dot{V}_E/\dot{V}_CO_2 \) slope was the least closely correlated to the other 3 variables. The OUES is known to correlate closely to peak \( \dot{V}_O_2 \) in healthy adults with correlation coefficients \( \geq 0.79 \) (Baba et al 1996, Hollenberg et al 2000, Pichon et al 2002, Van Laethem et al 2009, Williamson et al 2012), and in heart failure with correlation coefficients \( \geq 0.80 \) (Van Laethem et al 2005, Davies et al 2006, Straburzyńska-Migaj et al 2010). In patients with heart failure the correlation of OUES to the anaerobic threshold and \( \dot{V}_E/\dot{V}_CO_2 \) slope are not as close as with peak \( \dot{V}_O_2 \) (Davies et al 2006, Straburzyńska-Migaj et al 2010). Within one study of healthy individuals the AT showed a closer correlation to peak \( \dot{V}_O_2 \) than OUES (Van Laethem et al 2005). Our results are similar to those seen in these studies, peak \( \dot{V}_O_2 \), AT and OUES are reasonably well correlated; with \( \dot{V}_E/\dot{V}_CO_2 \) slope to a lesser extent. It is important that these variables relate to one another, but also important that they do not relate too closely. As shown by Pichon et al 2002, the OUES cannot be reliably used to predict peak \( \dot{V}_O_2 \) from submaximal data despite a close correlation. Whilst this might be seen as a limitation to using OUES in an athletic cohort (where arguably only peak \( \dot{V}_O_2/\dot{V}_O_2 \) max, and to a lesser extent AT, are relevant) it may well be a strength in clinical medicine. The presence of a correlation with peak \( \dot{V}_O_2 \) suggests that the magnitude of the variable may be important and relates to disease severity, whilst the lack of a perfect correlation shows that there may be extra
information regarding the physiological response to exercise shown by the AT, $\dot{V}_E/\dot{V}_{CO_2}$ slope and OUES beyond what we can gain from the peak $\dot{V}_{O_2}$. May this extra information relate to the heart or lungs, or another organ?

### 3.5.2 The impact of spirometric abnormalities on CPX

Within our pilot study, however respiratory abnormalities were identified, peak $\dot{V}_{O_2}$ was significantly affected. It showed a relatively strong relation ($R^2=0.11$) to percentage of predicted FEV$_1$, was significantly lower in patients with the lowest percentage of predicted FEV$_1$ when divided into 2 groups, and significantly affected by the grouping into normal spirometry and obstructive, restrictive and mixed patterns. In contrast the $\dot{V}_E/\dot{V}_{CO_2}$ slope and OUES were not affected by spirometric abnormalities, and the AT was weakly related ($R^2=0.04$) to percentage of predicted FEV$_1$ but not when the patients were separated according to their FEV$_1$ or into disease patterns. This differential effect of respiratory physiology on CPX variables may explain some of the heterogeneity between them regarding their correlation to one another, to their prognostic power and potentially to their discriminatory power. But do these results mean that peak $\dot{V}_{O_2}$ is more sensitive to respiratory disease? Spirometric abnormalities are very common in patients with CHF (Hosenpud et al 1990, Naum et al 1992, Dimopoulou et al 1998, Hughes et al 1999, Daganou et al 1999), especially an isolated restrictive defect which was seen in almost half of the patients in this study. Abnormal spirometry within this cohort may be a sign of more significant heart failure, and therefore the lower peak $\dot{V}_{O_2}$ seen when patients are separated according to their FEV$_1$ relates to their more severe heart failure, and is therefore actually more significantly correlated to cardiac function than $\dot{V}_{O_2}$ at the AT, $\dot{V}_E/\dot{V}_{CO_2}$ slope or OUES. However the relation between peak $\dot{V}_{O_2}$ and spirometry in patients who do not appear to be limited by their lungs is evidence against this hypothesis. A cut-off of 30% for the breathing reserve has typically been used to identify patients who are limited by their lungs versus another organ. Whilst heart failure patients have abnormalities within their lungs that impair both maximal voluntary ventilation and ventilatory efficiency, the cardiac limitation still generally halts exercise with a significant breathing reserve remaining. Therefore we can say that for heart failure patients with or without a respiratory limitation to exercise (as evidenced by a BR cut-off of 30%) there remains a strongly significant relationship between peak $\dot{V}_{O_2}$ and percentage of predicted FEV$_1$. Peak $\dot{V}_{O_2}$, in contrast to $\dot{V}_{O_2}$ at the AT, OUES and $\dot{V}_E/\dot{V}_{CO_2}$ slope, is affected by the added presence of respiratory abnormalities to CHF.
3.5.3 The impact of ejection fraction on CPX variables

It has long been established that resting markers of cardiac dysfunction such as ejection fraction relate poorly to exercise capacity (Szlachcic et al 1985, Cohn et al 1993). My data agrees with these previous studies; peak $\dot{V}_{O_2}$, OUES and $\dot{V}_{O_2}$ at the AT do not relate and $\dot{V}_E/\dot{V}_{CO_2}$ slope shows a borderline significance between 2 groups when separated by median EF. There are also no significant differences in the 4 variables between patients with clinical heart failure and preserved ejection fraction, and those with reduced ejection fraction. When excluding patients with preserved ejection fraction, and analysing EF as a continuous variable, there remains no relationship between EF and any of the 4 variables.

3.5.4 Clinical implications of these results

In a general population, often with a combination of respiratory and cardiac disease, it is, from a multitude of CPX variables, typically peak $\dot{V}_{O_2}$ that is used to make clinical decisions. However our interpretation of peak $\dot{V}_{O_2}$ is based upon information extrapolated from studies with very different populations. The patients in the studies which support peak $\dot{V}_{O_2}$ as the principal prognostic variable have typically been a highly selected group referred to specialist centres for evaluation of severe CHF and assessment for potential heart transplantation. These patients were invariably young, with a mean age of 50 years in the landmark paper on the role of measuring peak $\dot{V}_{O_2}$ in heart transplant candidates (Mancini et al 1991). The patients were less likely to have coexisting respiratory disease because of these younger ages and because there was a higher prevalence of idiopathic dilated cardiomyopathy (DCM) than that seen typically in clinical practice. Patients with an ischaemic cause of CHF are much more likely to have non-desirable risk factors, such as smoking, that would increase the prevalence of coexistent lung disease in this group. In young patients with idiopathic DCM, peak $\dot{V}_{O_2}$ is an accurate measure of functional capacity and prognosis. However, does this apply to a more typical hospital heart failure population, with the greater likelihood of co-morbidities?

Generally prognosis is significantly worse in patients with cardiovascular disease than respiratory disease at similar levels of functional impairment, as discussed in Chapter 1. The addition of mild to moderate COPD does not alter prognosis in CHF (Mascarenhas et al 2008), yet it should depress exercise capacity. Any CPX variable influenced by both heart and lung disease may have its ability to predict prognosis attenuated by its sensitivity to respiratory abnormalities. In our mixed population with cardiac and respiratory abnormalities peak $\dot{V}_{O_2}$ was significantly affected by FEV$_1$. 

112
Therefore I suggest that peak $V_{O_2}$ is not an ideal variable when assessing a general population with heart failure. Its sensitivity to respiratory abnormalities means that its ability to detect changes in exercise capacity relating specifically to cardiac limitation, and in determining the impact of cardiac disease on prognosis is only strong for selected patients with normal respiratory function; an uncommon finding.

Other variables may better reflect cardiac physiology in all patients, because they are relatively unaffected by respiratory disease.

3.5.5 Implications for the thesis
The aim of this pilot data was to show a proof of concept for the thesis. Peak $V_{O_2}$ is known to be reduced in respiratory and cardiac disease, and so it is unlikely that it can be used to specifically follow changes in cardiac over respiratory function. This pilot data confirms the strong association between respiratory abnormalities and peak $V_{O_2}$ and shows that other variables do not display this association. Therefore the hypothesis that an ideal variable on CPX exists, that can distinguish cardiac from respiratory limitation and track changes in cardiac function, seems valid. The prospective study aims to confirm the results seen here and extend them to identify if a variable does exist that is minimally influenced by respiratory physiology and strongly influenced by cardiac physiology.

3.5.6 Limitations
This pilot data collection is purely retrospective; prospective data collection is needed to confirm the findings. Because I was not involved in the exercise tests or echocardiography I cannot vouch for the validity of the original datasets. Certain results suggest that spirometry may have been poorly performed in some patients, with extremely low breathing reserves noted. 27 patients had negative breathing reserves with 16 having values under minus 10%. In these cases the MVV must therefore be spuriously low, due to an under-representative FEV$_1$ on spirometry. The direct measure of MVV (patient voluntarily ventilates at their maximum for 12 seconds) is not routinely performed in clinical medicine and is quite uncomfortable; the indirect method (FEV$_1$ x 40) is therefore generally preferred, although this can be susceptible to error through poor spirometric technique. It is likely however that my outliers do not affect the results too significantly. It is known that poorly measured independent variables will reduce the power of statistical relationships between them and the dependent
variable; therefore the relationship with peak \( \dot{V}_{O_2} \) will probably be even stronger in a cohort with highly reproducible spirometry through more accurate testing, than the already statistically significant values I show here.

All patients were referred for cardiopulmonary exercise testing and an echocardiogram based on functional limitation and clinical CHF, but not all patients show echocardiographic evidence of left ventricular systolic dysfunction, as measured by ejection fraction. The majority of these patients will have heart failure with preserved ejection fraction. The diagnosis of heart failure has been made by an experienced cardiologist, but it is possible that a small proportion of these patients will not in fact have cardiac disease. I have used ejection fraction as the measure of cardiac function, although it is recognised, and I accept, that EF is not the best measure of severity of cardiac failure. However EF is easily measurable, widely recognised, and in general it is accepted that a lower EF identifies patients with more impairment in their systolic function. Within this retrospective cohort EF was readily available. There is also a suggestion that many patients in this analysis did not perform maximal tests. The low peak RER (median 0.98, IQR 0.86-1.06) from these patients, lower than values seen in patients in previous CHF studies, is likely to be explained by the cohort, i.e. real world patients with multiple co-morbidities. The values seen in this analysis lie somewhere between pure CHF (RER= 1.1 in a study by Davies et al 2006) and pure COPD (RER= 0.82 in a study by O’Donnell et al 1992). In reality submaximal tests are common and should not invalidate the results of the test or this analysis.

Finally it must be remembered that all these tests were treadmill exercise. Although this isn’t a limitation to this analysis, the prospective study will investigate bicycle ergometer testing and potential differences in results between these 2 forms of exercise may be possible.

### 3.5.7 Conclusions

Peak \( \dot{V}_{O_2} \), OUES, the \( \dot{V}_E/\dot{V}_{CO_2} \) slope and \( \dot{V}_{O_2} \) at the AT were all differentially influenced by respiratory function, as identified by spirometric findings, prior to CPX testing in a typical out-patient CHF population. Peak \( \dot{V}_{O_2} \) related closely to FEV\(_1\) and was significantly different between groups when patients are categorised by pattern of respiratory disease. Even in patients unlimited by their respiratory function (as defined by a normal BR \( \geq 30\% \)) peak \( \dot{V}_{O_2} \) was still related to FEV\(_1\). In contrast OUES and \( \dot{V}_E/\dot{V}_{CO_2} \) slope, and on most analyses \( \dot{V}_{O_2} \) at the AT, were not related to FEV\(_1\) or disease grouping. Although these 4 variables correlate reasonably closely to
one another it is possible that the lack of complete agreement between them is, in some part, due to the differential effect of cardiac and respiratory pathophysiology on each variable.
4.0 Study of Health in Pomerania - Physiology of oxygen uptake kinetics
4.1 Abstract

Cardiopulmonary exercise testing (CPX) allows for assessment of cardiac and respiratory limitation, but is often affected by patient effort. Indices of oxygen kinetics, including the Oxygen Uptake Efficiency Slope (OUES), oxygen uptake to work rate slope ($\dot{V}_{O_2} - WR$ slope) and the heart rate to oxygen uptake slope ($HR - \dot{V}_{O_2}$ slope) are relatively independent of effort but it is unknown how much they are affected by patient characteristics. They may be more resistant to confounding from respiratory abnormalities than traditional variables identified at peak exercise.

1708 volunteers from the population-based Study of Health in Pomerania (SHIP) underwent an incremental bicycle exercise protocol. In each, markers of oxygen kinetics were calculated throughout exercise only (rest and recovery excluded). Participants with structural heart disease, echocardiographic or lung function pathology were excluded. The final study sample consisted of 577 males and 625 females. Age, height, weight, smoking, and beta-blockers were analysed for their influencing power in each sex. Reference equations for each variable were determined by quantile regression analysis.

Age, gender, height, weight, FEV$_1$ (but not percent predicted FEV$_1$) and other variables of lung function strongly related to OUES and OUES/kg (although height related to OUES/kg much less strongly than to OUES). Smoking was weakly related to OUES in males only. $\dot{V}_{O_2} - WR$ slope was minimally affected by age, gender, weight and FEV$_1$. Gender, height, weight and beta-blocker use, and smoking status, were related to the $HR - \dot{V}_{O_2}$ slope whilst age was only related in females and spirometry in males.

Markers of oxygen kinetics are differentially affected by age, gender, height, weight, lung function, beta-blocker use and smoking status. This allows the generation of unique reference equations for these variables, which aid physicians when exercise effort may be submaximal.
4.2 Introduction

Understanding why CPX variables change, and to varying degrees in each variable, both between and within patients is difficult. Humans are complex organisms and whilst we have excellent mechanisms designed to maintain homeostasis there are still marked differences between any two heart beats or two breaths. Largely through the identification of reference equations for the principal CPX variables previous researchers have identified factors that contribute to the between-patient variation. Within this chapter I will aim to expand this understanding, to look at factors influencing some of the lesser described CPX variables; those involved in oxygen kinetics. In Chapter 5 I will then look closely at factors contributing to the within-patient variation of many variables.

Variables of oxygen kinetics differ from peak $\dot{V}_{O_2}$ in that they assess the change in $\dot{V}_{O_2}$ in relation to different denominators throughout exercise, rather than focus only on a specific portion, for example peak exercise. Variables measured at peak exercise ignore all the potential data gathered during the minutes preceding this, which may be extremely valuable. It is believed that variables of oxygen kinetics relate to cardiovascular function, the delivery and utilisation of oxygen for which an adequately functioning heart and circulatory system is key. Early examples of these variables included $\tau$, otherwise called the time constant, which described, through an exponential function, the time taken for $\dot{V}_{O_2}$ to plateau after the beginning of exercise. However this required a constant work rate exercise test and these are largely no longer performed, certainly not in clinical medicine where incremental tests are now standard. A potential estimation of $\tau$, measurable on an incremental test is possible using the oxygen uptake to work rate ($\dot{V}_{O_2} - WR$) relationship. This is the slope describing the generally linear relationship between oxygen uptake and work rate in Watts. Even with an initial unloaded period, there is still a time lag before $\dot{V}_{O_2}$ starts to rise following the start of the incremental exercise portion, and this appears to correlate to $\tau$. However it is the gradient of this slope that has become the principal measure described from this relationship. The $\dot{V}_{O_2} - WR$ slope appears to show a certain degree of specificity for cardiovascular dysfunction, and varies little between healthy adults, regardless of age, height and weight (Hansen et al 1987).

If, instead of work rate, we plot the change in $\dot{V}_{O_2}$ against minute ventilation ($\dot{V}_E$) we find an exponential relationship. In 1996 it was first shown that logarithmic transformation of $\dot{V}_E$ describes a linear relationship; the oxygen uptake efficiency slope (OUES) was created (Baba et al 1996). The slope of this relationship is largely unchanged when data is shortened in patients with heart failure (Van Laethem et al 2005), it appears to be
highly predictive of prognosis in heart failure (Davies et al 2006) and may, like the $\dot{V}_O_2 - WR$ slope, signal some specificity for cardiac dysfunction.

Finally we can view the change in $\dot{V}_O_2$ against heart rate. At peak exercise the instantaneous relationship of these two variables is commonly described as the $O_2$ pulse. As peak $\dot{V}_O_2$ can be considered a surrogate for cardiac output so too can the $O_2$ pulse be considered a surrogate for stroke volume. But peak exercise variables fail to appreciate how they have evolved during the test, and are highly reliant on patients performing until truly symptom limited. When plotting heart rate (y axis) against $\dot{V}_O_2$ (x axis) the relationship rises roughly linearly.

The regression line of this relationship will be steeper (greater change in heart rate relative to $\dot{V}_O_2$) in conditions where stroke volume or oxygen extraction are limited, for example valvular regurgitation or peripheral myopathy (Flaherty et al 2001) and shallower (lesser change in heart rate relative to $\dot{V}_O_2$) in conditions where heart rate is constrained, for example conducting system disease leading to chronotropic incompetence or the use of beta-blockers.

All 3 of these variables have the potential to give us more information than we currently extract when we limit our focus on the final few seconds before test termination. They may be crucial in patients unable to exercise to a cardiorespiratory maxima because of an alternative limiting condition, for example an orthopaedic complaint.

Despite the potential benefits of routinely reporting these variables they are largely under-researched compared with other CPX variables such as peak $\dot{V}_O_2$, the $\dot{V}_E/\dot{V}_CO_2$ slope, the $O_2$ pulse and the anaerobic threshold.

Before variables can be used for diagnosis and management decisions in patients, it is vital to understand how they behave in health, and to identify factors that may confound a variable. Large scale population cohort studies give us an excellent way to identify how variables behave in health and typically, because recruitment is randomised, results are less susceptible to bias.

The Study of Health in Pomerania (SHIP) is a large population cohort study of healthy volunteers and recently published results on the influence of factors such as age, gender, body size, smoking and beta-blockade on some of the more commonly used CPX variables such as peak $\dot{V}_O_2$ (Gläser et al 2010). It generated contemporary predictive equations for a European population, and, unlike many previous studies, also provided the normal distribution of data across a wide age range and genders so that limits of health and normality could be appreciated.
Collaboration between my department and the authors of SHIP has allowed me a unique opportunity to investigate the oxygen kinetics of over one thousand largely healthy volunteers from a typical European population. My aim within this Chapter is to identify the impact of factors, such as age, gender, body size, respiratory function, smoking and beta-blockade, on these variables of oxygen kinetics - OUES, the $\dot{V}_O_2 - WR$ slope and the $HR - \dot{V}_O_2$, as well as generate predictive equations, both a measure of the average and range of normality. A further aim of this study is to investigate the reproducibility of OUES and the $\dot{V}_O_2 - WR$ slope following cropping of data prior to exercise termination, to mimic foreshortened exercise, and therefore to assess the statement that they are more robust measures than peak variables.
4.3 Methods

4.3.1 Study population

The Study of Health in Pomerania (SHIP) is a large cross-sectional, population-based survey in a region in the north-east of Germany. It recruited volunteers as study participants. The overall design and objectives of SHIP are well described (John et al 2001, Völzke et al 2010). Following reunification, it was noted that residents of former East Germany had worse health outcomes than those from former West Germany; this study was initiated to help explain these differences.

212 157 inhabitants were living in the region at the time of the study start date, a random selection of these residents were identified from the population registration offices (this captures all German inhabitants). A two-stage cluster sampling method was adopted from the World Health Organization (WHO) Monitoring Cardiovascular Disease Project in Augsburg, Germany. A representative sample, comprising 7 008 adults aged 20–79 yrs with 292 persons of each sex in each of the twelve 5-yr age strata, was identified. Following removal of migrated or deceased people the net sample consisted of 6 267 eligible subjects. 4 310 individuals participated in the baseline study of SHIP (SHIP-0), with data collection between October 1997 and March 2001. A 5-yr follow-up examination was performed (SHIP-1) between March 2003 until July 2006, which then comprised 3 300 subjects aged 25–85 yrs (the remainder had migrated, were deceased or were non-responding). Of those, 1708 individuals (834 males and 874 females) volunteered for a standardised progressive cycle incremental exercise test.

The study was funded by the Community Medicine Net of the University of Greifswald. The study conformed to the principles of the Declaration of Helsinki as reflected by an a priori approval of the Ethics Committee of the University of Greifswald (Greifswald, Germany) and all participants gave written informed consent.

Aside from CPX a number of other investigations and questionnaires were performed on participants at either of these time points or a third time point - SHIP-2 (between 2008-12). My analysis is limited to socio-demographic information and CPX results.

4.3.2 Pre-exercise diagnostics and exclusion criteria

Socio-demographic and behavioural characteristics, smoking status, and medical histories, were assessed using computer assisted interviews, administered by trained and certified staff. The previous history of diseases was based on self-reported physician’s diagnosis. The definition of cardiopulmonary disorders was based on self-
reported physician’s diagnosis, use of specific medication, electro- or echocardiographical pathological findings, and lung function abnormalities measured by spirometry and body plethysmography. Normal lung function was defined according to the recommendations of the European Coal and Steel Community (Quanjer et al 1993).

Information regarding participants underlying cardiopulmonary disorders was not released to me; however a list of participants felt to have an underlying contributory cause to exercise intolerance was supplied (without specific causes).

External observers regularly participated in certification procedures to facilitate comparability between SHIP and other population-based studies. Data collection was monitored by a Data Safety and Monitoring Committee (Greifswald).

Height and weight were measured for the calculation of the body mass index (BMI; body weight in kg divided by square of height in metres). Full lung function tests were performed at rest using a body plethysmograph equipped with a pneumotachograph (VIASYS Healthcare, Jaeger, Hoechberg, Germany). A minimum of three lung function manoeuvres were performed to obtain at least two acceptable readings.

**4.3.3 Exercise testing**

CPX variables were analysed breath-by-breath using a VIASYS Healthcare system (Oxycon Pro, Rudolph’s mask, JAEGER/VIASYS Healthcare system; Hoechberg, Germany). Prior to each test, the equipment was calibrated in standard fashion with reference gas and volume calibration. A single symptom-limited exercise test was performed using the same ergometer model as used for my Observational and Interventional studies (Ergoselect 100; Ergoline, Bitz, Germany). The protocol, which differed from the one described in my methods section and was not individualised, was modified from Jones et al with a stepwise increase in work load of 16 W/min, following unloaded cycling (Jones et al 1985). Each test was preceded by a resting period of ≥ 3 min to reach steady-state conditions. The test was continuously monitored by a physician. All tests were continued as symptom-limited (volitional exertion, dyspnoea or fatigue) unless the participant developed chest pain or ECG abnormalities. Participants were encouraged to reach maximal exhaustion prior to starting; during exercise no further motivation was given. A standard 12-lead ECG was obtained at rest, at every minute during exercise, and for ≥ 5 min during recovery; blood pressure was measured with a standard cuff sphygmomanometer.
Measures of \( V_E \) (L/min), tidal volume \( (V_t; \text{ in litres}) \), \( V_{O_2} \) and \( V_{CO_2} \) (both mL/min) were acquired on a breath-by-breath basis and averaged over 10-second intervals.

### 4.3.4 CPX data analysis

Peak \( \dot{V}_{O_2} \), \( \dot{V}_{O_2} \) at the AT, \( O_2 \) pulse, \( \dot{V}_E/\dot{V}_{CO_2} \) slope and ratios at rest and at the AT, and end-tidal CO\(_2\) measures were all previously calculated by the SHIP authors.

Participant data (10 second averages) was made available to me for post-hoc calculations of further variables.

These calculations included OUES, \( \dot{V}_{O_2} - WR \) slope and \( HR - \dot{V}_{O_2} \) slope, which were conducted using Matlab (Mathworks, Natick, Mass, USA) using only data points during exercise (unless otherwise described). OUES is defined as the slope of the regression line when \( \dot{V}_{O_2} \) (mL/min) (y-axis) is plotted against \( \log_{10} \) minute ventilation (x-axis). Similarly the \( \dot{V}_{O_2} - WR \) slope is defined as the slope of the regression line when \( \dot{V}_{O_2} \) (mL/min) (y-axis) is plotted against work load (Watts) (x-axis). Finally the \( HR - \dot{V}_{O_2} \) slope is defined as the slope of the regression line when heart rate (bpm) (y-axis) is plotted against \( \dot{V}_{O_2} \) (mL/min) (x-axis).

### 4.3.5 Test foreshortening

OUES and the \( \dot{V}_{O_2} - WR \) slope were further analysed so that a number of pre-specified data ranges within exercise were used to assess for the within-test reproducibility of these variables and their susceptibility to foreshortened exercise. Participants were grouped by their final workload achieved and the OUES for different data ranges compared by pairwise correlation within each group. Different ranges of foreshortened exercise were therefore not compared against different groups of participants which would affect the results (for example, the OUES calculated using data between 20-68W was compared against the OUES calculated from 20-196W, but only in participants achieving exactly 196W - if all participants were used then the larger numbers in the 20-68W cohort would include more exercise-limited people with typically lower OUES values). For the \( \dot{V}_{O_2} - WR \) slope the average \( \dot{V}_{O_2} \) at each workload was calculated and the slope calculated using this data from the start of exercise up until various points (52W, 100W, 148W, 196W and 244W) for 3 groups of participants based on their final workloads (100-116W, 148-164W, 196-212W and 244-276W) to view how foreshortening of data affects the slope.
4.3.6 Statistical analysis

Age, height, weight and BMI and lung function variables were assessed as continuous variables using regression analysis. The cohort was further divided into 5 age ranges; 25-34, 35-44, 45-54, 55-64, > 64 years and into 2 BMI groups with the cut-off at 25 kg/m$^2$ for graphical and tabular representation of the data only. Furthermore because the impact of age on CPX variables may not behave linearly at the extremes, polynomial regression was used to visually assess this. ANOVA was used for categorical potential confounders such as gender, smoking and beta-blocker use. To construct median, 5th and 95th percentile predictive equations, quantile regression was used. Quantile regression uses data within prespecified quantiles to draw regression lines. Its strength is that outliers do not influence the results; in this respect it is similar to linear regression in the way that the median is similar to the mean. Based on the results of the regression analysis for each variable individually, age, gender, height and weight, beta-blocker use and smoking status (graded as current or non-smoker – former smokers are graded with non-smokers) and FEV$_1$ could be included as co-variates in the model (in a similar manner to the complementary reference ranges published for peak $V_o_2$, anaerobic threshold and $O_2$ pulse). Each variable is only added to the regression equation for each variable individually if it showed a significant contribution in the simple univariate analyses. OUES and OUES/kg were interrogated with the same model.

Furthermore, for OUES, the current predictive equations, once identified, were compared to those previously published (Hollenberg et al 2000, Sun et al 2012) using the median participant within 3 height, age and BMI brackets for males and females.

A p value of <0.05 is considered significant throughout.
4.4 Results

4.4.1 Baseline characteristics

Following exclusions (prior to my involvement with the data) there remained a population of 1203 participants. These participants were not deemed to have a cardiopulmonary condition likely to limit exercise. They could be considered “healthy adults”, although exclusion did not mandate the absence of every chronic disease; the phrase “general population” is probably more appropriate. Hypertension, for example, was not a reason for exclusion. From these 1203 a single participant did not have raw data of sufficient quality to perform calculations of OUES, $\dot{V}_{O_2} - WR$ slope and $HR - \dot{V}_{O_2}$ slope. This left a final study population of 1202.

625 of these participants were female, 577 male. Within males the age range was 25-84 years, and within females 25-80 years. Median height was 170 cm (range 145-197 cm) and weight was 76 kg (range 42-150 kg).

Table 4.1 shows demographic data by age and gender.

159 males and 160 females were current smokers. 39 males and 54 females were taking regular beta-blockers.

4.4.2 Influence of demographics on CPX

Gender was a significant determinant of OUES (by ANOVA, $R^2=0.38$, $p<0.0001$), OUES/kg ($R^2=0.11$, $p<0.0001$), and $HR - \dot{V}_{O_2}$ slope ($R^2=0.30$, $p<0.0001$), with a significant but weaker interaction with $\dot{V}_{O_2} - WR$ slope ($R^2=0.04$, $p<0.0001$). Further determinants were thereafter assessed for males and females separately. The strength of the relationship between each determinant and each variable is shown in Table 4.2.

In both males and females, age was strongly inversely related to OUES and OUES/kg. The $\dot{V}_{O_2} - WR$ slope was only weakly, though significantly, related to age; in males this is a positive relationship and in females an inverse relationship. The $HR - \dot{V}_{O_2}$ slope was not related to age in males, and had a weak, albeit significant, relation to age in females. It can be seen in Figure 4.1 that the relationship of peak $\dot{V}_{O_2}$ and OUES with age remains linear throughout the majority of adult life (approximately 30-70 years of age) but outside of this range these variables behave in a less linear fashion. However given the linear portion of these plots is the age range where the majority of our patients are drawn from, and that polynomial equations are complex in clinical practice to understand and use, linear regression was used for reference range distribution.
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Subjects (n)</th>
<th>BMI (kg/m²)</th>
<th>Smokers (%)</th>
<th>Beta-blockers (%)</th>
<th>Peak $\dot{V}_{O_2}$ (mL/min/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-34</td>
<td>170</td>
<td>25.7 ± 3.2</td>
<td>41.8</td>
<td>0</td>
<td>34.5 ± 6.3</td>
</tr>
<tr>
<td>M</td>
<td>79</td>
<td>25.7 ± 3.2</td>
<td>41.8</td>
<td>0</td>
<td>34.5 ± 6.3</td>
</tr>
<tr>
<td>F</td>
<td>91</td>
<td>24.3 ± 4.8</td>
<td>35.1</td>
<td>1.1</td>
<td>28.1 ± 6.2</td>
</tr>
<tr>
<td>35-44</td>
<td>307</td>
<td>26.8 ± 3.5</td>
<td>31.2</td>
<td>1.9</td>
<td>32.2 ± 7.0</td>
</tr>
<tr>
<td>M</td>
<td>154</td>
<td>26.8 ± 3.5</td>
<td>31.2</td>
<td>1.9</td>
<td>32.2 ± 7.0</td>
</tr>
<tr>
<td>F</td>
<td>153</td>
<td>24.8 ± 4.0</td>
<td>38.6</td>
<td>1.3</td>
<td>26.1 ± 5.0</td>
</tr>
<tr>
<td>45-54</td>
<td>286</td>
<td>27.8 ± 3.7</td>
<td>35.9</td>
<td>3.9</td>
<td>28.4 ± 6.4</td>
</tr>
<tr>
<td>M</td>
<td>129</td>
<td>27.8 ± 3.7</td>
<td>35.9</td>
<td>3.9</td>
<td>28.4 ± 6.4</td>
</tr>
<tr>
<td>F</td>
<td>157</td>
<td>27.2 ± 5.2</td>
<td>25.4</td>
<td>7.0</td>
<td>22.8 ± 4.1</td>
</tr>
<tr>
<td>55-64</td>
<td>257</td>
<td>27.6 ± 3.4</td>
<td>19.7</td>
<td>8.5</td>
<td>27.5 ± 5.6</td>
</tr>
<tr>
<td>M</td>
<td>117</td>
<td>27.6 ± 3.4</td>
<td>19.7</td>
<td>8.5</td>
<td>27.5 ± 5.6</td>
</tr>
<tr>
<td>F</td>
<td>140</td>
<td>27.6 ± 4.6</td>
<td>15.7</td>
<td>11.4</td>
<td>22.0 ± 3.9</td>
</tr>
<tr>
<td>≥ 65</td>
<td>183</td>
<td>27.7 ± 4.5</td>
<td>9.1</td>
<td>21.2</td>
<td>23.7 ± 5.4</td>
</tr>
<tr>
<td>M</td>
<td>99</td>
<td>27.7 ± 4.5</td>
<td>9.1</td>
<td>21.2</td>
<td>23.7 ± 5.4</td>
</tr>
<tr>
<td>F</td>
<td>84</td>
<td>26.9 ± 4.1</td>
<td>8.3</td>
<td>28.6</td>
<td>19.7 ± 3.6</td>
</tr>
<tr>
<td>Total</td>
<td>1202</td>
<td>27.2 ± 3.8</td>
<td>27.6</td>
<td>6.8</td>
<td>29.3 ± 7.1</td>
</tr>
<tr>
<td>M</td>
<td>577</td>
<td>27.2 ± 3.8</td>
<td>27.6</td>
<td>6.8</td>
<td>29.3 ± 7.1</td>
</tr>
<tr>
<td>F</td>
<td>625</td>
<td>26.3 ± 4.8</td>
<td>25.6</td>
<td>8.6</td>
<td>23.8 ± 5.3</td>
</tr>
</tbody>
</table>

Table 4.1: Baseline characteristics of the 1202 participants undergoing CPX testing with measurement of oxygen kinetic variables. Participants are grouped by gender and 10-year age deciles. For peak $\dot{V}_{O_2}$ and BMI data are displayed as mean ± SD.
Height had a strong, positive relationship to OUES in males and females but, although remaining significant, the strength of this relationship was weaker when OUES was corrected for weight. The $\bar{V}_{\text{O}_2} - WR$ slope was not significantly related to height. The $HR - \bar{V}_{\text{O}_2}$ slope was weakly, inversely related to height in both males and females. Weight was strongly positively related to OUES in males and females, and strongly, but inversely, related to OUES/kg. The $\bar{V}_{\text{O}_2} - WR$ slope was weakly related to weight in both genders. The $HR - \bar{V}_{\text{O}_2}$ slope was strongly, inversely related to weight in males and females. Weight, when categorised by BMI, related, by ANOVA, to all 4 variables.

Importantly BMI did not relate as strongly to any of the 4 variables as height and weight as co-variates, and the addition of BMI to a simple model of height and weight did not increase the predictive power of the model (i.e. no change in $R^2$) suggesting that height and weight, but not BMI should be used in the more complex regression models. BMI was weakly but significantly related to the $\bar{V}_{\text{O}_2} - WR$, and strongly inversely related to the $HR - \bar{V}_{\text{O}_2}$ slope.

The impact of gender, age (divided into decades) and body size (divided into 2 groups, $25 \leq \text{BMI} < 25 \text{ kg/m}^2$) can be viewed in Figures 4.2-4.5.
Figure 4.1: The distribution of peak $V_{O_2}$ and OUES with age fit a polynomial distribution. The line represents the median, with 95% confidence intervals in grey. Throughout the majority of adult life the relationship between these variables and age is linear. In adults before 30 and after 70 years of age, the relationship deviates from this linearity.
<table>
<thead>
<tr>
<th>Determinant</th>
<th>OUES</th>
<th>OUES/kg</th>
<th>$\dot{V}_{O_2} - WR$ Slope</th>
<th>$HR - \dot{V}_{O_2}$ Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>(-) 0.17</td>
<td>&lt;0.001</td>
<td>(-) 0.13</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>M Height (cm)</td>
<td>0.18</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>0.003</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>0.12</td>
<td>&lt;0.001</td>
<td>(-) 0.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/ m$^2$)</td>
<td>0.02</td>
<td>&lt;0.001</td>
<td>(-) 0.16</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>(-) 0.12</td>
<td>&lt;0.001</td>
<td>(-) 0.18</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>F Height (cm)</td>
<td>0.14</td>
<td>&lt;0.001</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>0.10</td>
<td>&lt;0.001</td>
<td>(-) 0.20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/ m$^2$)</td>
<td>0.03</td>
<td>&lt;0.001</td>
<td>(-) 0.27</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 4.2: The influence of age, height, weight and BMI on the four variables of oxygen kinetics in both males and females. The $R^2$ value indicates the rough contribution of a determinant to the variability of that variable. A $p < 0.05$ indicates statistical significance. The $R^2$ value was calculated using linear regression analysis. $R^2$ values with a preceding (-) indicate a significant inverse relationship (i.e. an increase in the magnitude of a continuous determinant causes a reduction in the variable).
Figure 4.2: Reference ranges for OUES for males and females with relation to age and subdivided according to body mass index (BMI) groups. Lower lines represent the 5th percentiles, centre lines the medians and upper lines the 95th percentiles.
Figure 4.3: Reference ranges for OUES/kg for males and females with relation to age and subdivided according to body mass index (BMI) groups ——BMI ≤ 25kg/m², - - - - BMI > 25kg/m². Lower lines represent the 5th percentiles, centre lines the medians and upper lines the 95th percentiles.
Figure 4.4: Reference ranges for $\dot{V}_{O_2} - WR$ slope for males and females with relation to age and subdivided according to body mass index (BMI) groups. Lower lines represent the 5th percentiles, centre lines the medians and upper lines the 95th percentiles.
Figure 4.5: Reference ranges for $HR - V_{O2}$ slope for males and females with relation to age and subdivided according to body mass index (BMI) groups. --- BMI $\leq 25\text{kg/m}^2$, ---- BMI $> 25\text{kg/m}^2$. Lower lines represent the 5th percentiles, centre lines the medians and upper lines the 95th percentiles.
4.4.3 Influence of respiratory variables and smoking on CPX

The separate influence of FEV₁, FVC, the FEV₁:FVC ratio and K_{CO} on males and females can be seen in Table 4.3. Whilst a number of the interactions reached statistical significance it can be seen that only a few had a reasonable strength of relationship (i.e. R² value). Whilst the \( \dot{V}_{O_2} - WR \) slope was related to most of the measures in males and females it can be seen that this relationship was weak for all. The \( HR - \dot{V}_{O_2} \) slope was related to FEV₁ and FVC, but only in males, and these relationships were very weak.

OUES and OUES/kg were both related to FEV₁ and FVC, but not the ratio of the two. These relationships were stronger for OUES than for OUES/kg. Generally a much stronger relationship was also noted in males than females. A multivariate regression model including age, height and weight showed that in males both OUES and OUES/kg remained related to FEV₁, FVC and K_{CO} (p<0.001 for all). However for females neither of the CPX variables related to FEV₁, FVC or K_{CO}. Because it is difficult to identify the contribution to a regression model from one co-variante over another, the percent of predicted FEV₁ was also calculated. OUES related significantly but weakly in males (R²=0.02, p=0.002) and not in females (R²=0.00, p=0.19).

Furthermore the R² values for the variables of oxygen kinetics were compared against the R² values for other variables. Peak \( \dot{V}_{O_2} \) had, for both males and females, a higher R² value for both measures of FEV₁ (R²=0.34 & 0.21 respectively), FVC (R²=0.34 & 0.21 respectively), and K_{CO} (R²=0.21 & 0.06 respectively) than OUES. Similarly peak \( \dot{V}_{O_2} /kg \) had higher R² values than OUES/kg. In the same multivariate model used above there was a significant relationship between peak \( \dot{V}_{O_2} \) and FEV₁ and FVC in both males and females. The \( \dot{V}_{E} / \dot{V}_{CO_2} \) slope and \( \dot{V}_{O_2} \) at the AT both were significantly related to FEV₁, FVC, and K_{CO}, although the R² values were lower than for both peak \( \dot{V}_{O_2} \) and OUES.

Smoking did not relate to OUES/kg, \( \dot{V}_{O_2} - WR \) or the \( HR - \dot{V}_{O_2} \) slope, and had a weak inverse relationship with OUES in males (R²=0.01, p=0.01) and a weak positive relationship with females (R²=0.01, p=0.01).
<table>
<thead>
<tr>
<th>Determinant</th>
<th>OUES</th>
<th>OUES/kg</th>
<th>$V_{O_2} - WR$ Slope</th>
<th>$HR - V_{O_2}$ Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
<td>Strength of relationship ($R^2$)</td>
<td>p-value</td>
</tr>
<tr>
<td>FEV$_1$</td>
<td>0.22</td>
<td>&lt;0.001</td>
<td>0.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FVC</td>
<td>0.22</td>
<td>0.002</td>
<td>0.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV$_1$;FVC</td>
<td>0.01</td>
<td>0.006</td>
<td>0.01</td>
<td>0.054</td>
</tr>
<tr>
<td>$K_{CO}$</td>
<td>0.16</td>
<td>&lt;0.001</td>
<td>0.07</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV$_1$</td>
<td>0.12</td>
<td>&lt;0.001</td>
<td>0.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FVC</td>
<td>0.12</td>
<td>&lt;0.001</td>
<td>0.08</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV$_1$;FVC</td>
<td>0.00</td>
<td>0.18</td>
<td>0.00</td>
<td>0.22</td>
</tr>
<tr>
<td>$K_{CO}$</td>
<td>0.05</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Table 4.3: The influence of four respiratory variables on the four variables of oxygen kinetics in both males and females. The $R^2$ value indicates the rough contribution of a determinant to the variability of that variable. A $p < 0.05$ indicates statistical significance. The $R^2$ value was calculated using linear regression analysis. $R^2$ values with a preceding (-) indicate a significant inverse relationship (i.e. an increase in the magnitude of a continuous determinant causes a reduction in the variable).
4.4.4 Influence of beta-blockade on CPX variables

Age, height and weight were all significantly different between those who did and didn’t take beta-blockers; on average regular beta-blocker users were older (mean difference 13.8 years, p<0.001), shorter (mean difference 3.7cm, p<0.001) and heavier (mean difference 3.9 kg, p=0.01). Therefore the effect of beta-blockade on the CPX variables was performed in a multivariate model with these co-variates.

As shown in Table 4.4, beta-blocker use did not significantly reduce OUES, OUES/kg or the $\dot{V}_{O_2} - WR$ slope in males, although there were borderline significant reductions in OUES and OUES/kg in females. The use of beta-blockers significantly reduced the $HR - \dot{V}_{O_2}$ slope in both males and females.

4.4.5 Reference distribution

Using quantile regression reference distributions for median predicted and prespecified percentiles can be calculated. I chose the 5th and 95th percentiles to delineate the upper and lower limits of normality. The effect of age, gender, height, weight, beta blocker use, current smoking, and FEV$_1$ on the predictive equations for these four variables can be seen in Table 4.5.
### Table 4.4: Multivariate regression analysis adjusted for age, height and weight in males and females for the effect of taking beta-blockers on CPX variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean difference (95% CI)</th>
<th>SE</th>
<th>P</th>
<th>Mean difference (95% CI)</th>
<th>SE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES</td>
<td>-127.3 (-346, 92)</td>
<td>111.5</td>
<td>0.25</td>
<td>-128 (-242, -14)</td>
<td>58.3</td>
<td>0.028</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>-1.12 (-3.8, 1.5)</td>
<td>1.34</td>
<td>0.41</td>
<td>-1.71 (-3.4, -0.1)</td>
<td>0.84</td>
<td>0.043</td>
</tr>
<tr>
<td>$\dot{V}_o_2 - WR$ Slope</td>
<td>-0.12 (-0.59, 0.36)</td>
<td>0.24</td>
<td>0.63</td>
<td>-0.17 (-0.53, 0.2)</td>
<td>0.19</td>
<td>0.38</td>
</tr>
<tr>
<td>$HR - \dot{V}_o_2$ Slope</td>
<td>-0.004 (-0.008, -0.001)</td>
<td>0.00</td>
<td>0.016</td>
<td>-0.005 (-0.009, -0.001)</td>
<td>0.00</td>
<td>0.024</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES</td>
<td>-128 (-242, -14)</td>
<td>58.3</td>
<td>0.028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES/kg</td>
<td>-1.71 (-3.4, -0.1)</td>
<td>0.84</td>
<td>0.043</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\dot{V}_o_2 - WR$ Slope</td>
<td>-0.17 (-0.53, 0.2)</td>
<td>0.19</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HR - \dot{V}_o_2$ Slope</td>
<td>-0.005 (-0.009, -0.001)</td>
<td>0.00</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Percentile</td>
<td>Value</td>
<td>Equation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1272.9</td>
<td>-727 -6.98A +11.82H +6.56W -47.65bb-10.65cs +15.29f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1876.3</td>
<td>-182.4 -8.89A +10.12H +10.51W -117.65bb-21.45cs +40.31f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>2613.6</td>
<td>36.5 -18.47A +15.55H +12.63W -317.24bb-91.98cs +13.47f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES/kg</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>18.22</td>
<td>13.8 -0.14A +0.16H -0.16W -0.19bb-0.64cs-0.34f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>26.65</td>
<td>23.1 -0.16A +0.16H -0.2W -1.77bb-0.13cs-0.06f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>38.62</td>
<td>33.5 -0.23A +0.19H -0.29W -2.11bb-0.4cs +1.3f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ Slope</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>7.66</td>
<td>7.28 -0.01A +0.02W -0.04f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>9.47</td>
<td>8.82+0.01A +0.01W -0.1f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>11.3</td>
<td>10.61+0.02A +0.01W -0.23f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Slope</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0320</td>
<td>0.0897 -0.0001A -0.0002H-0.0004W -0.0021bb +0.0027f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0544</td>
<td>0.1327 -0.0001A -0.0003H-0.0004W-0.0042bb +0.0027f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0796</td>
<td>0.2279 -0.0001A -0.0008H-0.0005W -0.01bb +0.0038f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1783.5</td>
<td>54.7 -9.82A +4.42H +11.74W -227bb-169cs +180.72f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>2703.5</td>
<td>907.7 -11.51A +5.67H +8.62W -49.99bb-214.53cs +172.97f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>3951.4</td>
<td>-2380 -13.97A +34.48H +9.54W -251.11bb-523.24cs +0.47f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUES/kg</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>21.82</td>
<td>21.44 -0.13A +0.07H-0.11W -2.4bb-1.88cs +1.6f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>31.43</td>
<td>43.12 -0.14A +0.05H-0.23W -0.57bb-2.68cs +2f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>46.88</td>
<td>18.25 -0.14A +0.34H-0.37W -2.96bb-6.1cs +1.38f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ Slope</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>8.06</td>
<td>9.31 -0.02A-0.01W +0.17f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>10.06</td>
<td>9.63 -0.01A +0.01W +0.07f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>11.47</td>
<td>9.98 -0.003A +0.01W +0.17f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Slope</td>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0236</td>
<td>0.0686 -0.0001A -0.0001H-0.0002W -0.0016bb +0.0004f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0368</td>
<td>0.0743 -0.0001A -0.0003W -0.0039bb-0.0006f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.0548</td>
<td>0.0795 -0.0001A +0.00023H-0.0005W -0.002bb-0.0051f</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5: Predictive equations of reference ranges for 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentiles derived from age (A) in years, height (H) in cm, weight (W) in kg, and FEV<sub>1</sub> (f) in L. Beta blocker intake (bb) is coded as ‘0’ for no and ‘1’ for yes. Current smoking (cs) is coded as ‘0’ for no and ‘1’ for yes. The values show the average magnitude of each variable at the 3 percentiles.
4.4.6 Comparison with previous predictive equations for OUES

There are no standardised references ranges for $HR - \dot{V}_{O_2}$ slope or $\dot{V}_{O_2} - WR$ slope. However for OUES previous reference equations have been published (Hollenberg et al 2000, Sun et al 2012). For 3 heights, 3 BMI values and 3 ages the median male and female participants were identified. The age, height, weight and FEV$_1$ were entered into my predictive equations from Table 4.5 and the predictive equations from these 2 studies. The results were plotted as a percentage of the values identified from Table 4.5 (Figure 4.6). Sun et al 2012 can be seen to under-predict considerably compared to both other ranges with shorter, older and heavier individuals. Hollenberg et al over-predicts in the younger, and then under-predicts in the older participants compared to my data.

4.4.7 Impact of foreshortened exercise on OUES

Alongside the full OUES, calculated using all exercise data, OUES was also calculated from 9 different sub-maximal ranges. I cropped progressively from the whole exercise duration to shorter durations removing ever increasing periods at the end of exercise. Because each participant underwent an identical bicycle protocol periods of data cropping were denoted by Watts achieved. The data was cropped sequentially from full exercise, from 292W downwards, then 244W, 196W, 148W, 116W, 100W, 84W, 68W, and finally 52W downwards.

Figure 4.7 shows how this cropping affects the OUES in three groups: those achieving 116W but then stopping; those achieving 196W but then stopping; and those achieving 292W but then stopping, as a percentage of OUES measured using full data. For comparison the $\dot{V}_{O_2}$ at these points, as a percentage of peak $\dot{V}_{O_2}$, was also calculated and shown in Figure 4.7. Table 4.6 shows the pairwise correlation coefficient for each different crop against 7 others. Almost all cropped data correlated to all other crops with the exception of the 20-292W crop. This only contained a small number of participants and a value that significantly correlated to the full value was obtained so long as a work rate of at least 148 was achieved (144W below maximum).
Figure 4.6: Comparison of my new reference equations with 2 pre-existing reference ranges. Hollenberg et al 2000 is shown with crosses and Sun et al 2012 with open circles. 3 different heights (varied between males and females because of the height distribution within these groups), 3 different BMIs and 3 different ages were arbitrarily chosen and the median participant within the SHIP population meeting that criterion (i.e. male 190cm) was identified. The predicted OUES for that person was calculated using all 3 predictive equations with comparison for Hollenberg et al and Sun et al made against those in Table 4.5.
Figure 4.7: The effect of data cropping on OUES and peak $\dot{V}_{O_2}$. Participants data was measured from the beginning of exercise until one of nine different work rates (dots) with the OUES calculated in the typical way. Both OUES and $\dot{V}_{O_2}$ are plotted as percentages of the value obtained using the full data for 3 groups of participants based on their final work rate. Data foreshortening affects peak $\dot{V}_{O_2}$ to a greater degree than OUES. In participants achieving the lower peak work rates (116W) foreshortening affects OUES less than in those with higher maximum work rates (292W).
<table>
<thead>
<tr>
<th></th>
<th>20-84</th>
<th>20-100</th>
<th>20-116</th>
<th>20-148</th>
<th>20-196</th>
<th>20-244</th>
<th>20-292</th>
<th>Full OUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-100</td>
<td>0.9518*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-116</td>
<td>0.9128*</td>
<td>0.9821*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-148</td>
<td>0.7990*</td>
<td>0.8947*</td>
<td>0.9422*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-196</td>
<td>0.6555*</td>
<td>0.8000*</td>
<td>0.8514*</td>
<td>0.9043*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-244</td>
<td>0.5804*</td>
<td>0.6409*</td>
<td>0.6754*</td>
<td>0.7705*</td>
<td>0.8454*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-292</td>
<td>0.3776</td>
<td>0.3749</td>
<td>0.3728</td>
<td>0.5576*</td>
<td>0.8705*</td>
<td>0.9719*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Full OUES</td>
<td>0.7888*</td>
<td>0.8443*</td>
<td>0.8826*</td>
<td>0.8981*</td>
<td>0.9212*</td>
<td>0.9807*</td>
<td>0.9954*</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.6: Correlation coefficient values for comparison of each pair of foreshortened OUES readings, calculated for all data present for the range stated. For example the value is 0.9518 when comparing OUES calculated from data between 20 to 100W, versus OUES calculated using data from 20 to 84W in the same participants, i.e. those attaining at least 100W. Full OUES is the OUES calculated for each participant using their full exercise data, i.e. no foreshortening. Starred values indicate a statistically significant correlation.
4.4.8 Impact of foreshortened exercise on the $\dot{V}_{O_2} - WR$ slope

In a similar manner the data was then grouped into four cohorts based on final achieved work rate (100-116W, 148-164W, 196-212W and 244-276W), and $\dot{V}_{O_2}$ plotted against work rate for each of these groups based on cropping the data at the lower limits of each of these groups. These arbitrary groups roughly corresponded to each quartile, but all participants within each quartile were not included because each group needed to contain a small range of maximum work rates achieved for consistency. So participants with final work rates between these 4 group’s ranges were excluded from this analysis.

It can be seen that cropping the data from the top causes small decreases in the slope for the 2 groups where greater work rates were achieved (Figure 4.8A). However for the 2 lower final work rate groups there was actually an increase in the $\dot{V}_{O_2} - WR$ slope as the data is cropped. Figure 4.8B displays the mean $\dot{V}_{O_2}$ at each work rate for each of these same four groups. It can be seen that for any given work rate the $\dot{V}_{O_2}$ is higher in the groups achieving higher maximum work rates but the change in $\dot{V}_{O_2}$ over time remains similar. There does not appear to be a clear point of change on any of the four lines away from a linear relationship.
Figure 4.8A: Multiple mean $\dot{V}_O_2 - WR$ slopes were calculated from all the participants within four groups; those achieving a maximum of 100-116W, those achieving 148-164W, those achieving 196-212W and those achieving 244-276W. These slopes were calculated using data from start of exercise to 5 different final cut-off points. For example the grey column in the left most set is the $\dot{V}_O_2 - WR$ slope calculated from the beginning of exercise to the final second of 52W only in participants stopping at between 244-276W.

Figure 4.8B: $\dot{V}_O_2$ was then plotted against work rate for all participants within these same four groups. It can be seen that the slope does not deviate from linearity and is similar between different work rates attained.
4.5 Discussion

As shown in Chapter 3, OUES and $\dot{V}_E/\dot{V}_{CO_2}$ slope are less affected by spirometry in a group of patients with heart failure. Here I demonstrate, for OUES and other important cardiopulmonary exercise test variables focusing on the kinetics of oxygen uptake, how they are affected by demographic characteristics and lung function in a general healthy population as well as generate a set of contemporary predictive equations based on quantile regression.

4.5.1 Study of Health in Pomerania

SHIP is a large ongoing population cohort study, in the North-East of Germany, established to answer some of the differences in health equality between areas of Germany. Although the study found higher rates of obesity, hypertension and left ventricular hypertrophy than the national rates within Germany (John et al 2001, Völzke et al 2010), because Germany typically has better indices of health than the European average (WHO Global Infobase 2013), it could be considered that this population is relatively representative of northern Europe. A random sample from over 200 000 inhabitants was drawn, of which almost 2000, aged 25-84, returned at 5 years and underwent CPX. A final study population of 1203 was identified, without significant cardiopulmonary disease. The participants may not have been fully healthy, in that certain chronic diseases, for example hypertension, were not excluded. Hypertension is so common within a general population (especially in older life) I felt it was important to retain these participants. Previous publications from these 1203 participants looked at the impact of individual characteristics on peak $\dot{V}_{O_2}$, O$_2$ pulse, $\dot{V}_{O_2}$ at the AT, and the $\dot{V}_E/\dot{V}_{CO_2}$ slope and ratio (Koch et al 2009, Gläser et al 2010). All variables were influenced by age, height, weight and gender. Current male smokers showed a significant impact on all variables as well, although in females only the $\dot{V}_E/\dot{V}_{CO_2}$ ratio was worse in smokers. Importantly they showed that for peak $\dot{V}_{O_2}$, $\dot{V}_{O_2}$ at the AT, and the $\dot{V}_E/\dot{V}_{CO_2}$ ratio, a strong linear trend of worsening values with increasing pack year histories of smoking (Gläser et al 2010).

I show similar results, with OUES (and OUES/kg) related to height and weight and inversely related to age, as has been shown before (Hollenberg et al 2000). The $\dot{V}_{O_2} - WR$ slope was largely unrelated to any of these variables. The $HR - \dot{V}_{O_2}$ slope was affected by gender, and significantly inversely related to weight. It was largely unrelated to age, with a weak inverse relationship to height. Smoking only weakly related to OUES, and in different directions in males and females, and did not relate to the other variables. FEV$_1$, FVC and K$_{CO}$ appear to relate to OUES and OUES/kg more strongly in males than females. Much of this appears to be due to the
dependent impact of age, height and weight on both lung function and OUES. Peak $\dot{V}_{O_2}$ appears more strongly related to spirometric variables than OUES. Finally beta-blockade largely only affected the $HR - \dot{V}_{O_2}$ slope, unsurprisingly in an inverse manner.

### 4.5.2 Predictive equations

Predictive equations and values for cardiopulmonary exercise variables are necessary in order to establish normality for patients. Previous predictive equations should be revisited as populations change. As CPX availability has grown, the number of commonly reported variables has also expanded. In many cases predictive values are either based on small cohort studies or on studies where participants recruited may not necessarily be representative of the whole population. Also many predictive equations focus largely on a measure of average without demonstrating the boundaries of normality. Within a clinical population it is expected that well over 50% of tested individuals will have values below the mean, therefore a measure of the spread of normality is more important than the average. I have used quantile regression because this allows the calculation of the 5th and 95th percentiles without any effect from significant outliers beyond these boundaries. This is similar to the advantages seen by using the median over the mean as a measure of the average, which I have also used as the single “average” measure for each variable. I believe these predictive equations (and the ones for peak $\dot{V}_{O_2}$ from Gläser et al 2010) are highly representative of a general European population, and I will therefore be using them throughout the rest of the thesis.

Another interesting result from this data is the relationship between age and CPX variables. Whilst a linear relationship can be fitted to almost any set of data, this does not mean that the relationship is truly linear. Polynomial regression will fit linearly if that best describes the data, but otherwise will form as a complex function of the 2 axes. We can see in Figure 4.1 that for most of adult life OUES, like peak $\dot{V}_{O_2}$, declines linearly with age. However this decline becomes less pronounced after about 70 years, and we can also see a plateau around 30-35 years of age. This is important when using these predictive equations, as certainly below 25 years they will be unreliable. I would suggest the possible entry of age 30 into the predictive equations for anyone younger than this (to avoid over-predicted misrepresentative reference ranges).

I will now look at the results of the interaction between demographic factors on each of the variables in turn, as well as this impact on the predictive equations.
4.5.3 Oxygen uptake efficiency slope

It has largely been accepted that OUES is an effort-independent variable (Baba et al 1996); this makes it useful for patients undergoing clinical exercise testing who may stop prematurely due to reasons beyond their heart or lungs, for example orthopaedic constraints. However, just like peak $\dot{V}_{O_2}$, $\dot{V}_{O_2}$ at the AT and $\dot{V}_E/\dot{V}_{CO_2}$ slope, participant characteristics such as age, gender, height and weight significantly influence the result. This means that OUES, as an absolute value, has limitations to its use, and should, therefore, wherever possible be corrected by predictive equations. What is interesting is the influence of lung function on OUES. Smoking affected OUES but with a very weak relationship in males, and in females was associated with an improvement in OUES.

Spirometry unsurprisingly related strongly to OUES in a general population, because they both are strongly related to age, height and weight. When corrected for these factors using predictive equations for FEV$_1$, OUES remained related to FEV$_1$ in males but not females, and the strength of this relationship was very weak. OUES was less strongly related to all spirometric variables than peak $\dot{V}_{O_2}$, which supports the results found in patients with heart failure as described in Chapter 3.

Until recently only a single set of adult predictive equations have been published for OUES (Hollenberg et al 2000), with 2 significant limitations. Firstly the adults were aged 53 years and above. Secondly, similarly to many other published reference equations, average values rather than normal ranges are shown. To generate their predictive equations, the effect of age on OUES was extrapolated backwards from 53 years, which may not make it reliable for younger patients. The difference between this predictive equation and mine can be seen graphically in Figure 4.6 (crosses). I have calculated, for the median person within three age, height and BMI groups, the predicted OUES based on both sets of equations. Unsurprisingly Hollenberg’s reference ranges appear to over-predict the average value in younger adults. However there are no consistent trends in how height and BMI influence the differences between these 2 equations; these two equations show a relative amount of agreement. Very recently a further set of predictive equations has been published from the USA (Sun et al 2012). Using this predictive equation shows startling differences at the extremes of height and age, more noticeable in females than males. It is difficult to reconcile these differences, but I feel that the equations from Sun et al markedly under-estimate predicted OUES, rather than mine over-estimate. For example for both an average female of 150cm height, and an average female aged 70 years (using average age, weight and FEV$_1$ of females of 150cm height, and average height, weight and FEV$_1$ of females of 70 years of age from within the SHIP’s large population to generate a predicted values) an OUES of only 0.8 L/10-fold increase in $\dot{V}_E$ are predicted from Sun et al; predicted values disproportionately lower than other groups and males.
I have also generated the predictive equation for OUES when adjusted for weight in kilograms. Correction for weight is easier than for body surface area (BSA) or BMI; these both require calculations and are less popularly performed in busy clinical practice. Correcting for weight allows for comparison with peak $\dot{V}_{\text{O}_2}$, which everyone typically describes in a weight adjusted (per kilogram) manner. Higher average values of peak $\dot{V}_{\text{O}_2}$ in mL/min are seen in heavier individuals, however when adjusted for weight peak $\dot{V}_{\text{O}_2}$ in mL/min/kg is generally lower. The same pattern is seen with OUES (Figures 4.2 and 4.3). By viewing OUES as both absolute and weight adjusted values we may be able to reduce the limitations that either value has when used alone.

Whilst it is believed that OUES is resistant to foreshortened exercise, and so is, therefore, an effort-independent variable, I show that foreshortening of exercise by cropping the data at a number of arbitrary points can affect results (Figure 4.7). The reliability of OUES despite foreshortening has been shown in numerous studies (Baba et al 1999, Hollenberg et al 2000, Van Laethem et al 2005) and whilst my data largely agrees, in the most fit individuals (those achieving at least 200W) variation can be seen. Reassuringly when data including at least 66% of final intensity is included, an OUES value within 10% of the overall is achieved. It is unlikely that a participant capable of achieving almost 300W would voluntarily stop prior to 200W for example. In people achieving lower final intensities (arguably the group in which foreshortening of data is most likely) the variation seen with cropping is even less noticeable. These volunteers will have exercise capacities similar to patients with heart failure, which supports the previous data showing no significant effect from submaximal data in heart failure patients (Van Laethem et al 2005). My data therefore provides further evidence to the argument that, whilst not completely effort-independent, OUES is an excellent variable in its robustness to foreshortened exercise.

**4.5.4 Oxygen uptake to work rate relationship**

The slope of the linear relationship between work rate (Watts) against $\dot{V}_{\text{O}_2}$ (mL/min) is believed to reflect aerobic work efficiency, and oxygen delivery and utilisation at the muscle. It is therefore believed to be primarily influenced by cardiac rather than respiratory function. Previous cohort studies identified almost identical average values of 10 mL/min/Watt in healthy controls with a standard deviation of 0.7-1.0 mL/min/Watt (Wasserman et al 1975, Hansen et al 1984, Hansen et al 1987). A reduced slope has been reported in patients with heart failure, peripheral vascular disease and myopathies. In contrast, extremely fit cyclists have been found to have an increased slope of 11.5 mL/min Watt (Riley et al 1996). However one of these previous
studies only examined males, and, as can be seen from Figure 4.5, females appear to have a slightly lower median value. In the SHIP cohort age does not appear to affect median values of the slope, however there does appear to be a divergence of values in the older age groups. Unsurprisingly BMI doesn’t have much effect on the slope; it has long been established that obesity displaces the $\dot{V}_{\text{O}_2}$-work rate relationship upwards without altering the slope (Wasserman et al 1975, Hansen et al 1984, Hansen et al 1987). Despite having generated a predictive equation there are conceptual concerns to generating a predicted value for a variable like this. Unlike a linear variable, such as peak $\dot{V}_{\text{O}_2}$ and OUES, the $\dot{V}_{\text{O}_2} - WR$ slope corrected as a percentage of predicted would be less meaningful as a value of 5 is not half as good as 10, and 20 is not twice as good. Figure 4.5 would be best employed to place a patient’s value on the graph and ensure it lies within the 5th and 50th percentile lines to be reassured about normalcy.

Amongst participants achieving higher final work rates there is a small but clear reduction in the slope as the data is cropped from the top (Figure 4.8). In contrast the $\dot{V}_{\text{O}_2} - WR$ relationship in those achieving lower intensities did not noticeably differ when cropped. It is conceivable that in stronger, larger individuals a minimum work rate exists below which it still requires the same amount of energy to turn the pedals, i.e. $\dot{V}_{\text{O}_2}$ remains similar from unloaded through to 20-40W, beyond which the normal increment of ~10 mL/min/W takes over. This would have the effect of reducing the slope as proportionally more of the data came from this plateau section. The reduced slope seen in foreshortened data from fitter individuals does argue against one of the explanations given for the low slope in heart failure patients; the proportional increase from anaerobically derived energy sources at all time points compared with healthy adults. If this was true we would still expect to see this pattern in healthy individuals following the anaerobic threshold, albeit to a lesser degree than heart failure patients, and arguably the group we would see this pattern in most would be fit individuals, who are able to push themselves further into anaerobic metabolism than sedentary individuals. In fit individuals we see a slight increase in the slope as exercise continues, suggesting that, certainly within this group, anaerobic sources do not replace, merely supplement, aerobic sources; the delivery of 10 mL/min/W of oxygen continues well into high intensity exercise. A confounding factor is that fit individuals are more likely to use accessory muscle groups towards peak exercise, which might necessitate an increase in energy requirements per increment of work rate. This phenomenon is less likely to occur in sedentary individuals and those with heart failure.
4.5.5 Heart rate to oxygen uptake relationship

There are 3 main factors influencing oxygen uptake; heart rate, stroke volume and oxygen extraction at the muscle. The former two factors comprise cardiac output. Many conditions can affect any number of these factors so it can be useful for the exercise physiologist to assess the individual interaction between them.

Typically we use the relationship of $\dot{V}_{O_2}$ / heart rate, the O$_2$ pulse. Peak muscular oxygen extraction is believed to be near equal in all adults under standard conditions, therefore the O$_2$ pulse, which is the product of oxygen extraction and stroke volume, can be used as a surrogate for stroke volume. However this has a similar limitation to peak $\dot{V}_{O_2}$, namely that sub-maximal effort will fail to produce a true reflection of peak exercise stroke volume. An alternative is the heart rate – $\dot{V}_{O_2}$ relationship or slope, where heart rate is plotted on the y axis, and $\dot{V}_{O_2}$ on the x axis, with the slope of the regression line calculated by least squares. A similar variable, the heart rate response (HRR) which is equal to (HR$_{peak}$ – HR$_{rest}$) / ($\dot{V}_{O_2peak}$ – $\dot{V}_{O_2rest}$), has occasionally been described in other studies utilising CPX. My measure of the slope of the regression line of HR versus $\dot{V}_{O_2}$ will incorporate all data points, whilst the HRR only utilises the peak and rest instantaneous data. I therefore believe it may be more robust (resting values are especially susceptible to error) and may be more easily calculable from modern CPX software packages where the measurement of a regression slope requires a single click. I also believe there may be multiple clinical utilities of this variable. A healthy adult stopping prematurely through non-cardiac limitation will display a normal slope - both maximum heart rate and peak $\dot{V}_{O_2}$ are similarly reduced from predicted maxima - and this makes it, in these cases, a more useful variable than the O$_2$ pulse to reassure there is a normal response. In patients with heart failure the slope may either be elevated (when cardiac conduction is maintained) or decreased (when there is associated chronotropic incompetence). The latter is more commonly seen, especially in the modern era of beta-blockade. In a group of patients with mitochondrial myopathies, proven on biopsy, Flaherty et al showed the HRR to be universally above 50, perhaps reflecting an increased dependency of oxygen uptake on heart rate to compensate for reduced oxygen extraction at the muscle (Flaherty et al 2001). This would reflect a HR – $\dot{V}_{O_2}$ slope of 0.050.

Amongst the SHIP population markedly higher median values in the HR – $\dot{V}_{O_2}$ slope were noted in females compared to males, with little variation attributed to age. The slope was reduced in heavier individuals, paradoxically suggesting lower heart rates throughout exercise. Height did not strongly affect this slope. Unsurprisingly the use of beta-blockers reduced the slope in males and females. Importantly a HR – $\dot{V}_{O_2}$ slope
of 0.05, equivalent to a HRR of 50 which is believed to be helpful in the diagnosis of mitochondrial myopathies, would be considered a healthy, normal finding based upon my findings especially in females.

4.5.6 Limitations

It is widely accepted that exercise testing should employ individualized protocols, so that exercise duration ranges from 8-12 minutes (Balady et al 2010). All participants in the SHIP study were assessed using the same protocol and it is currently unknown how an exercise test lasting significantly less than 8 or more than 12 minutes may affect the OUES, the $\dot{V}_{O_2} - WR$ slope or the $HR - \dot{V}_{O_2}$ slope. Simply cropping the data does not reproduce the differences from short, medium and long duration tests. 156 participants did not achieve 8 minutes; 135 participants achieved over 12 minutes. Within Chapter 5 I will look into the effect of test duration on CPX variables. However although this could be considered a limitation, for an exercise physiology laboratory or research study where individualised protocols are not practical, adherence to a 16W/min protocol will still allow these reference equations to be applicable. Whenever possible I believe individualised protocols should be performed.

As described earlier within this section these participants were not entirely representative of an overall German population, with greater levels of obesity and hypertension within Pomerania than the rest of Germany. Therefore it is possible that they are not reflective of a general European population. However this criticism could be levelled at all reference range studies; it isn’t practical for each country or geographical region to produce their own predictive equations.

Participants who don’t deem themselves fit enough would be less likely to volunteer for the exercise study, leading to a selection bias. This is likely to be an ever increasing problem as volunteer age increases. However it has been shown in a previous publication that the CPX population was largely similar to the general SHIP cohort of over 4000 participants with the only differences less hypertensive people and less smokers represented in the CPX group (Koch et al 2009).

I included ex-smokers with non-smokers. It could be argued that they should be included with current smokers, however I feel that the main impact on exercise testing from smoking will be a reduction in oxygen carrying capacity in the blood due to increased carbon monoxide. This is removed from the blood within 48 hours of
smoking, and so would not limit ex-smokers. The long-term detriment of smoking to lung function would be captured on the full lung function tests and so would not require consideration.

Finally prior to my involvement with the data, certain conditions had been selected for which participants from the original 1708 were excluded, leaving the remainder which was considered a cohort of healthy adults. It could be argued that certain conditions I elected to keep in, smoking, hypertension and obesity, could be considered disease states and should not therefore be involved in the identification of reference ranges. However these are common patient attributes that we see regularly in clinical practice and so I felt it was important to include them.

4.5.7 Conclusions
CPX variables such as the OUES, $\dot{V}_{O_2} - WR$ slope and $HR - \dot{V}_{O_2}$ slope have not been investigated to the same extent as variables such as peak $\dot{V}_{O_2}$ so questions remain as to how they are influenced in a typical population.

Within this chapter I have shown how age, height, weight and gender interact to affect these variables and from these interactions generate predictive equations for an average value and 5th to 95th percentiles.

OUES, whilst affected by lung function, compares favourably to peak $\dot{V}_{O_2}$ in that it is less affected by resting spirometry and is more resistant to data-foreshortening. $\dot{V}_{O_2} - WR$ slope and the $HR - \dot{V}_{O_2}$ slope are largely unrelated to lung function, although the $HR - \dot{V}_{O_2}$ slope is related to beta-blocker use.
5.0 Reproducibility/ Repeatability
5.1 Abstract

Cardiopulmonary exercise testing (CPX) has been shown to have good reproducibility. However the test-retest reliability for multiple CPX variables has not been compared in a single study and the influence of different diseases and study conditions on test-retest reliability has not been examined. I show different measures of test-retest reliability for multiple CPX variables and compare them by category of cardiac or respiratory disease.

Patients from the Observational study underwent two CPX tests as part of the study protocol. Separately eight healthy adults underwent four CPX tests, two on an individualised appropriate ramp, one on a very steep ramp and one on a shallow ramp. Comparison of between patient/participant intraclass correlation coefficients (ICC) and coefficients of variation between the two tests were calculated. Analysis of variance was used to calculate the influence of potential confounders on test-retest reliability.

Variables such as peak $\dot{V}_O_2$ (ICC 0.95; CI 0.94-0.97), OUES (ICC 0.93; 0.90-0.95), O$_2$ pulse (ICC 0.96; 0.94-0.97) and the $\dot{V}_E/\dot{V}_CO_2$ ratio at the nadir (ICC 0.92; 0.89-0.95) all showed excellent test-retest reliability, with within-subject coefficients of variation <0.12 in the patient study. Oxygen uptake at the anaerobic threshold (ICC 0.84; 0.78-0.89) and the $\dot{V}_E/\dot{V}_CO_2$ slope (ICC 0.88; 0.79-0.93) still showed good test-retest reliability, although significantly weaker than for peak $\dot{V}_O_2$. Overall test-retest reliability was similar in the healthy adults. A change in ramp protocol affected the reliability of the peak work load and $\dot{V}_O_2 - WR$ slope, but did not significantly affect the majority of CPX variables.

CPX shows high test-retest reliability; certain variables such as peak $\dot{V}_O_2$ and OUES outperform others. These results identify which variables are most suitable for serial testing of patients with 3 common disease aetiologies owing to their superior reproducibility.
5.2 Introduction

Any test for a chronic feature of a patient should, in principle, give the same value every time it is conducted. The more alike two results conducted on the same specimen are, the more reproducible they are said to be. The correct definition of reproducibility relates to the degree of agreement between measurements or observations on the same specimen/patient in different locations by different people. In clinical medicine and research true reproducibility is rarely assessed. Typically the criteria of one or both of different locations and different operators are not met; when this occurs it should correctly be termed repeatability, rather than reproducibility. However repeatability can be considered an important component of reproducibility.

There are many factors affecting a test’s reproducibility, broadly broken down into 2 groups; random error, which includes within-patient biological and temporal variations and operator/reporter variability; and systematic error, which includes phenomena which causes the result to be consistently under or over-estimated within a particular individual, in exercise testing examples would be fatigue or familiarisation. This second part is sometimes called bias or trueness as it relates to the ability of the test to identify the patient’s true result; any consistent deviation is bias. The magnitude of the random error every time a test is conducted cannot be predicted but it is possible to understand the likely spread of random error so that the final result can be reported with intervals of confidence. Both random and systematic errors are assessed with repeatability studies. For the purposes of this introduction I will largely use the term reproducibility, as this is the term typically used by authors publishing in this area, even if their studies actually report repeatability. The results I show within this chapter are however repeatability, not true reproducibility, data.


The precision of modern hardware at measuring the principal non-derived gas-exchange variables such as oxygen consumption ($\dot{V}_{O_2}$) and carbon dioxide production ($\dot{V}_{CO_2}$) is quoted to be around 0.01-0.03%, and for flow variables such as tidal volume ($V_T$) and respiratory frequency ($R_f$) around 1-3% (Balady et al 2010). The majority of derived variables are now calculated using computer software, minimising interpretation error. Therefore the majority of variation between two tests in the same patient will be temporal and biological variation in the patient and technical imperfections such as mask leak. This variation is largely what is measured in these reproducibility studies.
A key study into reproducibility was the HERITAGE family study (Skinner et al 1999) which, due to its study design, probably is the most accurate study at determining true reproducibility as a subgroup of their participants were assessed at each of 4 recruitment sites to look for the influence of different locations and different operators. Coefficients of variation (CoV) and intraclass correlation coefficients (ICC) for peak $\dot{V}_{O_2}$ did not differ when within site repeatability was compared with between site reproducibility. The total numbers of 390 healthy individuals makes it one of the largest reproducibility studies to date.

Prior to this, Elborn et al had, in a small group of patients with heart failure, shown evidence of a learning effect on the treadmill, with a consistent bias between the first two tests. When comparing tests 2 and 3 there was not a significant difference. Exercise time was 20% higher and peak $\dot{V}_{O_2}$ was 17% higher in the second test compared with the first (Elborn et al 1990). Bensimhon et al tried, on multivariate analysis, to find factors that altered the reproducibility of peak $\dot{V}_{O_2}$. The strongest predictor of a change in peak $\dot{V}_{O_2}$ on the second test was a low peak $\dot{V}_{O_2}$ on the first test. Reproducibility also appears to be relatively unchanged when looking at a more elderly population with heart failure (Marburger et al 1998) and when patients with severe heart failure are tested (Meyer et al 1997).

Largely the above studies have focused on a small number of cardiopulmonary variables, namely peak $\dot{V}_{O_2}$, peak $\dot{V}_{CO_2}$, $\dot{V}_{CO_2}$ at the AT, peak RER, blood pressure and heart rate. A study by Keteyian et al was perhaps unique in its broad approach to evaluating a number of variables for reproducibility, although somewhat strangely many variables were only assessed in a subgroup of study individuals, which would not allow for a visual qualitative comparison of the CoV between variables. Often following the introduction of a new variable into the CPX vocabulary, studies will report the reproducibility or repeatability of this variable, but often fail to compare with all other established variables in a methodological and statistical manner (Hollenberg et al 2000, Van Laetham 2009, Jakovljevic et al 2012). What remains unknown is whether a group of variables are consistently and significantly better or worse than others in reproducibility.

Another problem with quotations for reproducibility statistics in medicine is that studies often perform under the most optimal conditions. They take optimal patients and have experts in the field performing the tests. This is not the picture seen in day-to-day clinical practice. In CPX research studies participants are tested on standard protocols, yet in clinical practice protocols are often not standardised within an institution, let alone between institutes. Within the 11 studies described above, 10 used a standard protocol for all patients with no deviation. It is established that to obtain optimal results for your patient an individualised protocol, aiming for maximal
exercise within 8-12 minutes, should be employed (Balady et al 2010). The only study that appears to have definitively looked into the effect of test duration, did so on only 5 healthy adult males (Buchfuhrer et al 1983).

Only a single study (Agostoni et al 2005) has investigated the effect of work rate protocol on exercise test variables, and only in patients with heart failure. Peak $\dot{V}_{\text{O}_2}$ and the $\dot{V}_{\text{O}_2} - WR$ slope were lower in the shortest, steepest ramp protocol. However the reproducibility of various variables were not directly, statistically compared between their 3 levels of protocol difficulty. I do not believe a similar experiment has been published in healthy controls, and certainly there is no established data as to the contributing nature of protocol difficulty on a broad range of CPX variables.

Another problem is that research studies often adhere to strict policies that are rarely enforced in routine clinical practice, for example the fasting state. Many of the previously reported reproducibility studies utilised baseline data for an interventional study. In these cases it is appropriate to optimise and standardise testing conditions, to minimise biological variation and allow a true interventional effect to be seen. But it is highly likely with the increased use of CPX in research (often without direct input from CPX experts) that controlled, optimal conditions will be less rigorously adhered to. “True” or “day-to-day” reproducibility is yet to be described. In research a familiarisation test may be warranted, but to minimise time disruption it may need to be performed on the same day as the second test; none of the studies described above examined the role of inter-test interval on reproducibility.

Therefore I believe a number of questions remain regarding the reproducibility of CPX variables:

- What is the test-retest reliability of all commonly used CPX variables?
- How do the test-retest reliabilities of each variable compare with one another in the same study?
- Do the test-retest reliabilities perform similarly under typical clinical conditions when comparing to studies employing ideal research conditions?
- Does disease aetiology affect reproducibility?
- How does protocol affect reproducibility of the test?
- How does inter-test interval affect reproducibility?
There is an assumption that measurements of slopes, including the $\dot{V}_E/\dot{V}_\text{CO}_2$ slope and OUES, are more reproducible than instantaneous or averaged data variables such as peak $\dot{V}_O_2$ or $\dot{V}_E/\dot{V}_\text{CO}_2$ ratio, largely due to their relative independence from effort. However this issue has not been examined by formal statistical techniques. Measurement of the gradient of a slope is susceptible to variation, which may therefore affect reproducibility, if the start and finish of the data used to construct the slope are not standardised. There still remains doubt as to where many CPX slopes should be measured. For example the $\dot{V}_E/\dot{V}_\text{CO}_2$ slope should include exercise data up to, but not beyond, the ventilatory compensation point (the point in incremental exercise where the relationship between minute ventilation and carbon dioxide elimination becomes decoupled due to the metabolic acidosis secondary to lactate build-up). Recently evidence has suggested that measuring the slope using all data points increases its prognostic power in patients with heart failure (Ingle et al 2007) when compared to measuring only until the VCP. Could this be influenced by the subjectivity of VCP identification? Unlike most other CPX variables, the VCP requires subjective identification from the operator. Therefore it is possible that the slope up to the VCP has poor reproducibility because the identification of the VCP is difficult? Using healthy adult data and a sample from the heart failure, mitral valve surgery and COPD patients I will show graphically how the slopes may change when data is foreshortened to different degrees. There also remains a question as to when slope measurements should start? In cycle exercise an unloaded period of exercise typically occurs. This may be of different lengths of time between centres (although 3 minutes is a typical value used) but depending on how long a patient manages to cycle in the incremental phase, the proportional contribution of this unloaded section will vary from test to test, which may affect the value of the slope.

Within this study I performed 2 baseline tests on each patient. The initial test protocol was 3 minutes of unloaded cycling following 3 minutes of rest, and then a ramp of 10 watts per minute until exhaustion, and had two main aims: to act as familiarisation for the patient; and to identify the optimal protocol for their second test. Based on the peak work rate achieved on a 10W/min protocol, an appropriate protocol aiming to deliver 8-10 minutes of incremental cycling was chosen for the second test. As stated earlier, based predominantly on a small study (Buchfuhrer et al 1983) 10 minutes allows enough time for accurate data acquisition, whilst minimising fatigue. Most patients underwent this second test on the same day as the first after at least 2 hours of recovery, others on a separate day. The design of the study therefore allows the calculation of, for typical patients with COPD, heart failure and mitral regurgitation, the test-retest reliability of multiple CPX variables, and assessment of the influence of different disease aetiologies, protocol changes and inter-test time interval on the test-retest reliability.
Separate to the principal study I also recruited a group of healthy adults, who underwent four CPX tests to allow the test-retest reliability of variables in healthy adults to be shown and to act as a baseline for our protocol.

Similar to the study by Agostoni et al (which was in heart failure patients) participants performed one test with a ramp that would be considered too hard, one test with a ramp considered too easy and two tests with a ramp considered appropriate. The aim of this test was to compare, within a healthy population, the test-retest reliability of multiple variables under the conditions of an identical protocol (the two appropriate tests) versus the conditions of significantly different ramp protocols.
5.3 Methods

5.3.1 Study 1 – Healthy controls

A group of eight healthy participants was recruited, all employees of either Imperial College, or Imperial College Healthcare NHS Trust. These participants had no chronic cardiorespiratory disease, were of normal weight and did not smoke. They included some (n=3) with high levels of physical fitness (but no elite athletes) and some more sedentary individuals (n=5). They each undertook four CPX tests, none on the same day. Wherever possible I tried to perform each individual’s tests at a similar time of the day. Based on the participants’ described exercise capacity, an average protocol aimed at delivering maximal effort around 8-12 minutes was agreed upon, a further protocol was then designed aiming for maximal effort in approximately 6 minutes, and a further protocol aiming for maximal effort in approximately 15 minutes. Two tests were performed using the middle protocol (8-12 minutes). Test 4 was always one of these tests. Tests 1 to 3 were a random order of a too hard protocol, a too easy protocol and an appropriate protocol. Henceforth these will be referred to as steep, shallow and standard ramps respectively.

The tests were performed on an Ergoselect 100 bicycle ergometer (Ergoline GmbH, Bitz, Baden-Württemberg, Germany) in an air-conditioned room after familiarisation with the equipment. Participants were encouraged to abstain from caffeinated products. The fasting state was not adhered to, but only small meals were encouraged, remote from testing. Spirometry and exercise testing was performed using COSMED Quark CPX System (COSMED S.r.l. Rome, Italy) which was calibrated before each test in the same manner as for the principal cohort study. They were asked to refrain from talking throughout the test, unless clinically required. Gas exchange was monitored breath-by-breath. The protocol resembled that of the main study cohort: 3 minutes resting; 3 minutes unloaded cycling without any resistance from the cycle ergometer; a steady ramp protocol specific to the participant and which test they were performing; and 3 minutes of recovery. The ramp protocol, in the same manner as the patient cohort, increased in small increments every 10-15 seconds, rather than larger increments every minute. The participants were asked to soundlessly indicate when they felt the resistance of the pedals. Although they were not blinded to the length of the unloaded portion (3 minutes) no clock was visible to help influence their decision. The work rate on the bike was only visible to the operator (me), not to the participants.
5.3.2 Study 2 - The Observational cohort

The full methodology for the recruitment and testing of patients is described in the Methods Chapter. As part of the study design each patient undertook two CPX tests prior to any intervention. The first test acted both as familiarisation and to identify an optimal protocol for test 2. Therefore when comparing tests 1 and 2 they were not always on the same protocol. The change in protocol and its influence on reproducibility will be assessed as described in the statistics section. Many patients in the Observational cohort also underwent an intervention. These results are not included in the reproducibility analysis. Patients with prior CRT implantation undertook 3 tests: the familiarisation test with their CRT settings unadjusted as for the other groups; a further test with their CRT settings unadjusted on the patient’s optimal protocol; and another test with the CRT function of their pacemaker disabled (but basic pacing and defibrillator functions were not adjusted). These final two tests were performed on the same protocol as decided from the familiarisation test and were performed in a random order. These 3 tests were spread out over 2 visits, largely with the familiarisation test on visit 1, and the two subsequent tests in a random order on another day. In the same manner as for all other patients, at least 2 hours separated the 2 tests performed on the same day. For the purposes of the reproducibility analyses the 2 tests with the CRT settings unadjusted will be compared. Where the second of these 2 tests was performed on a separate day to the familiarisation test, but 2 hours after the test with CRT settings disabled (whose data is not considered in the reproducibility analysis), it will still be considered as being on the same day, as the principal aim of this analysis is to determine the role of fatigue on reproducibility.

5.3.3 Statistical Analysis

Statistical analysis was performed using Stata version 11.1 for Windows (StataCorp LP, College Station, Texas). All patient characteristics and CPX variables were assessed for normal distribution using Shapiro-Wilk test. Because of the problems with analysing multiple variables, and no a priori hypotheses regarding which variables may behave normally or non-normally, a Sidak correction was employed ($\beta = 1 - (1 - \alpha)^{1/n}$) where $\beta$ equals the p value for significance, and n equals the number of variables (correctly variables that are related for example peak $\dot{V}_{O_2}$ and peak $\dot{V}_{O_2}$/kg count as 1). For those variables showing a non-normal distribution only if logarithmic transformation significantly normalised distribution was this transformation used. For all other non-normal distributions the raw value was used in non-parametric analyses when necessary.
Because there were only 8 participants within study 1, non-parametric tests were used throughout where necessary.

To assess for differences between 2 tests in the healthy cohort Wilcoxon rank tests were used, and between both pre-intervention tests in the patient cohort, paired t-tests were employed for parametric data, and Wilcoxon rank tests for non-parametric data. In the healthy cohort the steep and shallow ramps were compared against the first standard ramp (to minimise any potential confounding from a learning effect).

Data were examined using Bland Altman plots (Bland et al 1986) and the mean difference and standard deviation of the difference (SDD) between pairs of tests was calculated. A within-subject coefficient of variation (CoV) was calculated as the ratio of the SDD and the mean value of the test variable. This gives a convenient measure of the absolute reliability. CoV measurements performed this way are typically applicable for most data, even non-parametric data, as it is the distribution of the residuals that concerns us, not the distribution of the original data. Residuals are typically normally distributed. However I also present for the main study CoV values calculated using logarithmic transformation of the data (and identify in which variables this method has been used) with an exponential correction back at the end for any variable that displays residuals that diverge as the magnitude of the mean value increases, i.e. the residual is proportional to the mean. These logarithmic CoV are calculated via the following equation (where \(x_1\) and \(x_2\) are the values for each variable on test 1 and test 2 respectively):

\[
\text{exp}\left(-\frac{\left(\log_e x_1 - \log_e x_2\right)^2}{2}\right)
\]

Bootstrapped intraclass correlation coefficients (ICC) were also calculated for the pooled group and each individual disease category, to give a measure of relative reliability. In the simplest terms the ICC is a ratio of the within-patient variance to the whole group variance, and as a ratio of 2 variances is a unit-less value. This allows, even if 2 variables have different units of measurement, a comparison of 2 ICC values. 95% confidence intervals of each ICC were calculated by bootstrapping and a z-score of Fisher transformation was calculated with p values for direct comparison of some of the more important variables’ ICCs within the patient cohort (the numbers were too small in the healthy cohort to render this reliable). Excellent reliability was defined as ICC >0.90, good reliability 0.7-0.89, and moderate reliability 0.4-0.69 (as standard). Reliability was also calculated with regards to comparison to peak \(\dot{V}_{O_2}\), i.e. better, worse or similar reliability to peak \(\dot{V}_{O_2}\). ICC calculations
always used raw data, not logarithmically transformed data, even when previously stated the variable was non-normally distributed.

I analysed for 6 potential confounders to reproducibility that differed between patients; age (separated into above and below median), gender, weight status (BMI above or below 25), aetiology of disease, whether the 2 tests were performed on the same day (inter-test interval) and finally, whether the ramp protocol was steeper, shallower or the same as the original protocol. I constructed repeated measures mixed linear models using an interaction term between test and potential confounder. A significant interaction term (p<0.05) identified that the difference in a variable between tests 1 and 2 was influenced by that confounder. The presence of a significant interaction term however does not identify the direction of the affect on reproducibility (improving or reducing it). Therefore oneway analysis of variance was also used, with the differences between the tests as the dependent variable. Because disease aetiology is an important potential confounder, ICC and SDD/mean values were produced for each variable within each disease category.

Within the healthy cohort only ramp protocol was assessed for its potential role as a confounder.

A p value <0.05 was considered statistically significant throughout.
5.4 Results

5.4.1 Study 1 - Participant and test characteristics

Eight healthy participants undertook four CPX tests each. The healthy adults included two females and six males, median age 25 (range 22-43). The characteristics of the participants are shown in Table 5.1. None of the participants were smokers. Spirometry was normal with a percentage of predicted FEV$_1$ 105.7% (IQR 87.6 – 111.8%) and percentage of predicted FVC 114.4% (IQR 97.4 – 136.1%).

The steep ramp protocol was designed to elicit an exercise time of approximately 6 minutes. For one participant the protocol was more challenging than expected (exercise time 04:49 minutes) and another not challenging enough (exercise time 08:25 minutes) but all others lay between 05:39-07:28 minutes. On the shallow ramp protocol an exercise time of approximately 15 minutes was the target. 2 participants were significantly short of this target at 12:03 and 12:53 minutes. On the standard protocol, aiming for an exercise time of 8-12 minutes, one participant stopped earlier than this (06:58 minutes) and one participant beyond this exercise time (12:23 minutes). These were the same participants who were significantly shorter and longer respectively on the steep ramp.

5.4.2 Study 1 - Identification of intensity onset

From these 32 tests, on 30 occasions the participant remembered to indicate when they detected an increase in work rate. The work rate first detected ranged from 9-35 W with a mean value of $22.4 \pm 7.2$ W (Figure 5.1). The intensity of the protocol did not significantly affect when the participant detected the increasing work rate ($p=0.41$ by ANOVA for the 3 groups of test intensity, $p=0.61$ by linear regression of ramp used in W/min).
<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Gender</th>
<th>Age (years)</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
<th>Considered Athletic?</th>
<th>Steep ramp (W/min) and ET (min:sec)</th>
<th>Shallow ramp (W/min) and ET (min:sec)</th>
<th>Standard ramp (W/min) and ET (min:sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>28</td>
<td>183</td>
<td>77</td>
<td>22.99</td>
<td>No</td>
<td>40W/ 06:29</td>
<td>10W/ 18:48</td>
<td>25W/ 09:34</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>25</td>
<td>175</td>
<td>75</td>
<td>24.49</td>
<td>Yes</td>
<td>50W/ 07:28</td>
<td>15W/ 20:44</td>
<td>30W/ 11:34</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>24</td>
<td>178</td>
<td>68</td>
<td>21.46</td>
<td>Yes</td>
<td>50W/ 06:44</td>
<td>15W/ 16:44</td>
<td>25W/ 10:34</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>25</td>
<td>162</td>
<td>45</td>
<td>17.15</td>
<td>No</td>
<td>25W/ 04:49</td>
<td>8W/ 12:03</td>
<td>15W/ 06:58</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>22</td>
<td>178</td>
<td>70</td>
<td>22.09</td>
<td>No</td>
<td>40W/ 05:39</td>
<td>15W/ 12:53</td>
<td>25W/ 08:39</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>28</td>
<td>168</td>
<td>56</td>
<td>19.84</td>
<td>Yes</td>
<td>30W/ 08:25</td>
<td>12W/ 18:25</td>
<td>20W/ 12:23</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>43</td>
<td>185</td>
<td>69</td>
<td>20.16</td>
<td>No</td>
<td>50W/ 07:08</td>
<td>16W/ 18:59</td>
<td>30W/ 10:44</td>
</tr>
</tbody>
</table>

**Table 5.1: Baseline characteristics for eight healthy adults undertaking four CPX tests.** The final 3 columns show the watt per minute increase during the incremental portion of the exercise test for each of the 3 levels of difficulty, as well as the exercise time (ET) achieved in each test. The exercise time for the standard protocol shown was the first time it was performed in each participant. Considered athletic? was the participant’s opinion of their own level of fitness, being athletic meant they considered themselves of above average fitness.
Figure 5.1: Distribution of work rate in an incremental test when participant perception of increasing intensity is first noticed by the participant.
5.4.3 Study 1 – Repeatability of 2 identical CPX tests

When comparing the difference between the 2 standard tests only peak minute ventilation and \( P_{ET}CO_2 \) at the AT showed a borderline significant difference between these 2 tests. Although there was no significant difference in peak \( \dot{V}O_2 \) between these 2 tests, if we just consider the 3/8 participants who undertook the standard protocol first there was a mean increase of 3.1 mL/min/kg between this first and final test. In comparison for the 5/8 participants where the first standard protocol was performed on test 2 or 3 the mean increase was 0.14 mL/min/kg. Although the numbers are too small to draw firm conclusions this could be evidence of a learning effect within this group. In conflict with this hypothesis are the results shown in Figure 5.2. This shows the mean peak \( \dot{V}O_2 \) by the order of the tests performed. Test 1 displayed lower mean values when compared to tests 2 and 4 but it was test 3 which had the lowest mean peak \( \dot{V}O_2 \). There were no significant differences in peak \( \dot{V}O_2 \) between any pairs of tests in test order (p>0.05 Wilcoxon rank test for all 6 pairs, i.e. test 1 vs test 2, test 1 vs test 3 etc).

Table 5.2 shows the main measures of test-retest reliability between the 2 standard tests. Peak \( \dot{V}O_2 \), in mL/min, mL/min/kg and as a percentage of predicted, showed good repeatability as evidenced by ICC values ≥ 0.93 and CoV < 0.07. Most other gas exchange variables showed an intermediate degree of repeatability (0.80≤ ICC <0.91; 0.10≤ CoV <0.20). Some variables such as the \( \dot{V}E/\dot{V}CO_2 \) slope showed a good reliability using CoV, but poor ICC values, whilst the OUES/kg showed a good ICC but slightly weaker CoV. Using the whole data (rather than ending the slope at the VCP) did not improve test-retest reliability measures of the \( \dot{V}E/\dot{V}CO_2 \) slope; both ICC and CoV were worse, although this difference was not statistically significant.
Figure 5.2: Median and interquartile range for peak $V_{O_2}$ (mL/min/kg) by test order. Tests 1 and 3 show the lowest median values, which is not suggestive of a learning effect. There was no significant difference between any pairs of these values. The whiskers represent the lowest data point within 1.5 IQR of the 25th and 75th percentiles.
<table>
<thead>
<tr>
<th>Overall mean (Test 1 and 2)</th>
<th>Mean difference (and 95% LOAs)</th>
<th>Standard deviation of difference (SDD)</th>
<th>CoV</th>
<th>ICC (and 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $\dot{V}_{O_2}$ (mL/min)</td>
<td>2858</td>
<td>88 (-292, 468)</td>
<td>193.8</td>
<td>0.068</td>
</tr>
<tr>
<td>Peak $\dot{V}_{O_2}$ (mL/min/kg)</td>
<td>42.9</td>
<td>1.23 (-4.49, 6.96)</td>
<td>2.92</td>
<td>0.068</td>
</tr>
<tr>
<td>% predicted Peak $\dot{V}_{O_2}$</td>
<td>107</td>
<td>2.61 (-10.95, 16.18)</td>
<td>6.92</td>
<td>0.065</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>1758</td>
<td>-39 (-528, 450)</td>
<td>249.6</td>
<td>0.142</td>
</tr>
<tr>
<td>AT (% of pred peak $\dot{V}_{O_2}$)</td>
<td>66.0</td>
<td>-2.69 (-23.9, 18.52)</td>
<td>10.82</td>
<td>0.164</td>
</tr>
<tr>
<td>OUES</td>
<td>2.91</td>
<td>-0.05 (-0.73, 0.63)</td>
<td>0.35</td>
<td>0.120</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>44.2</td>
<td>-1.09 (-11.7, 9.6)</td>
<td>5.44</td>
<td>0.123</td>
</tr>
<tr>
<td>OUES 25-75</td>
<td>2.92</td>
<td>0.13 (-0.64, 0.9)</td>
<td>0.39</td>
<td>0.135</td>
</tr>
<tr>
<td>OUES 50</td>
<td>2.81</td>
<td>-0.31 (-1.21, 0.58)</td>
<td>0.46</td>
<td>0.162</td>
</tr>
<tr>
<td>OUES 70</td>
<td>2.94</td>
<td>-0.18 (-1.02, 0.66)</td>
<td>0.43</td>
<td>0.146</td>
</tr>
<tr>
<td>OUES 90</td>
<td>2.95</td>
<td>-0.07 (-0.77, 0.62)</td>
<td>0.35</td>
<td>0.120</td>
</tr>
<tr>
<td>OUEP</td>
<td>44.8</td>
<td>-1.4 (-7.86, 5.06)</td>
<td>3.30</td>
<td>0.074</td>
</tr>
<tr>
<td>$O_2$ Pulse</td>
<td>16.8</td>
<td>0.12 (-3.59, 3.82)</td>
<td>1.89</td>
<td>0.113</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 1</td>
<td>24.3</td>
<td>-0.42 (-3.26, 2.42)</td>
<td>1.45</td>
<td>0.060</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 2</td>
<td>27.7</td>
<td>1.3 (-3.81, 6.41)</td>
<td>2.61</td>
<td>0.094</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio nadir</td>
<td>24.5</td>
<td>0.45 (-2.28, 3.17)</td>
<td>1.39</td>
<td>0.057</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio at AT</td>
<td>26.1</td>
<td>0.73 (-1.43, 2.88)</td>
<td>1.10</td>
<td>0.042</td>
</tr>
<tr>
<td>RER at peak</td>
<td>1.24</td>
<td>0.02 (-0.16, 0.19)</td>
<td>0.09</td>
<td>0.073</td>
</tr>
<tr>
<td>$P_{ET}CO_2$ at AT *</td>
<td>43.5</td>
<td>-1 (-3.57, 1.57)</td>
<td>1.31</td>
<td>0.030</td>
</tr>
<tr>
<td>HR at peak</td>
<td>175</td>
<td>2.88 (-16.28, 22.03)</td>
<td>9.78</td>
<td>0.056</td>
</tr>
<tr>
<td>$R_t$</td>
<td>37.7</td>
<td>1.39 (-4.72, 7.51)</td>
<td>3.12</td>
<td>0.083</td>
</tr>
<tr>
<td>Peak $\dot{V}_E$ *</td>
<td>108</td>
<td>9.83 (-12.96, 32.62)</td>
<td>11.63</td>
<td>0.108</td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ Slope</td>
<td>9.72</td>
<td>0.04 (-0.93, 1.01)</td>
<td>0.49</td>
<td>0.051</td>
</tr>
<tr>
<td>Maximum Wattage</td>
<td>254</td>
<td>8.38 (-12.97, 29.72)</td>
<td>10.89</td>
<td>0.043</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Slope</td>
<td>0.04</td>
<td>0 (-0.01, 0.02)</td>
<td>0.01</td>
<td>0.178</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Intercept</td>
<td>56.7</td>
<td>-6.8 (-38.65, 25.05)</td>
<td>16.25</td>
<td>0.286</td>
</tr>
<tr>
<td>Duration (s)</td>
<td>613</td>
<td>17 (-36, 70)</td>
<td>27</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Table 5.2: The mean value, mean difference, standard deviation, standard deviation of the difference and the bootstrapped intraclass correlation coefficient with confidence interval when comparing the two standard tests undertaken by 8 healthy volunteers. Slope 1 refers to the slope calculated up until the VCP. Slope 2 refers to the slope calculated with all exercise data. CoV is the coefficient of variation and equals the SDD/mean. ICC is the intraclass correlation coefficient, calculated using bootstrapping measures. LOA = limits of agreement. CI = confidence intervals. * p<0.05 for difference between tests.
5.4.4 Study 1 – Repeatability of CPX tests using different ramp protocols

Table 5.3 shows the mean difference between the first standard protocol and either the steep or shallow ramp protocols for each of the CPX variables. When comparing the standard and steep protocols only 3 variables were affected by the change in protocol, peak RER (p=0.025), peak work rate (p=0.01) and the $\dot{V}_{O_2} - WR$ slope (p=0.01). When the shallow protocol was compared to the standard protocol the respiratory frequency at peak exercise (p=0.049), peak work rate (p=0.01) and the $\dot{V}_{O_2} - WR$ slope (p=0.02) were affected by a change in protocol.

Table 5.3 also shows how the 2 measures of test-retest reliability, ICC and CoV performed. On the steep protocol heart rate data was poorly recorded in one participant, the 4 variables utilising heart rate are therefore calculated from 7, not 8 participants. Peak $\dot{V}_{O_2}$, OUES, OUES/kg and $O_2$ pulse showed good test-retest reliability on both measurements. The AT and OUEP performed slightly less well. Peak heart rate, the $HR - \dot{V}_{O_2}$ intercept, and the $\dot{V}_{O_2} - WR$ slope performed poorly when measuring the ICC. The $\dot{V}_{O_2} - WR$ slope had, however, a good CoV, reflecting the small amounts of variance relative to the magnitude of the variable. The $\dot{V}_{O_2} - WR$ slope was significantly less reliable when comparing the steep/standard protocols with the 2 standard protocols (ICC 0.02 vs 0.78, p=0.02).
<table>
<thead>
<tr>
<th></th>
<th>Mean difference</th>
<th>CoV for</th>
<th>ICC (and 95% CI)</th>
<th>Mean difference</th>
<th>CoV for</th>
<th>ICC (and 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>steep vs standard (95% LOAs)</td>
<td></td>
<td></td>
<td>shallow vs standard (95% LOAs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak $\dot{V}_{\text{O}_2}$ (mL/min)</td>
<td>-66.2 (-555, 383)</td>
<td>0.086</td>
<td>0.92 (0.53, 0.99)</td>
<td>-42.4 (-583, 498)</td>
<td>0.099</td>
<td>0.91 (0.42, 0.99)</td>
</tr>
<tr>
<td>Peak $\dot{V}_{\text{O}_2}$ (mL/min/kg)</td>
<td>-1.29 (-8.38, 5.8)</td>
<td>0.087</td>
<td>0.88 (0.64, 0.96)</td>
<td>-0.61 (-8.47, 7.25)</td>
<td>0.096</td>
<td>0.93 (0.5, 0.99)</td>
</tr>
<tr>
<td>% predicted Peak $\dot{V}_{\text{O}_2}$</td>
<td>-3.69 (-21.52, 14.14)</td>
<td>0.088</td>
<td>0.91 (0.76, 0.97)</td>
<td>-1.76 (-20.86, 17.34)</td>
<td>0.093</td>
<td>0.91 (0.67, 0.98)</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>-85 (-583, 413)</td>
<td>0.146</td>
<td>0.87 (0.55, 0.97)</td>
<td>10 (-648, 669)</td>
<td>0.189</td>
<td>0.82 (0.43, 0.95)</td>
</tr>
<tr>
<td>AT (% of pred max)</td>
<td>-3.59 (-24.38, 17.19)</td>
<td>0.162</td>
<td>0.85 (0.65, 0.94)</td>
<td>-0.82 (-29.08, 27.44)</td>
<td>0.215</td>
<td>0.75 (0.43, 0.9)</td>
</tr>
<tr>
<td>OUES</td>
<td>0.07 (-0.48, 0.62)</td>
<td>0.094</td>
<td>0.88 (0.31, 0.99)</td>
<td>-0.1 (-0.53, 0.34)</td>
<td>0.077</td>
<td>0.91 (0.55, 0.99)</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>1.02 (-7.08, 9.13)</td>
<td>0.091</td>
<td>0.87 (0.34, 0.98)</td>
<td>-1.68 (-8.49, 5.13)</td>
<td>0.078</td>
<td>0.87 (0.55, 0.97)</td>
</tr>
<tr>
<td>OUES 25-75</td>
<td>0.15 (-0.82, 1.13)</td>
<td>0.170</td>
<td>0.69 (-0.03, 0.94)</td>
<td>0.1 (-0.68, 0.89)</td>
<td>0.138</td>
<td>0.81 (0.33, 0.96)</td>
</tr>
<tr>
<td>OUES 50</td>
<td>0.2 (-0.99, 1.39)</td>
<td>0.198</td>
<td>0.65 (0.41, 0.8)</td>
<td>-0.36 (-1.61, 0.89)</td>
<td>0.229</td>
<td>0.64 (-0.02, 0.91)</td>
</tr>
<tr>
<td>OUES 70</td>
<td>0.13 (-0.91, 1.17)</td>
<td>0.172</td>
<td>0.71 (0.26, 0.91)</td>
<td>-0.24 (-1.23, 0.76)</td>
<td>0.174</td>
<td>0.73 (0.3, 0.91)</td>
</tr>
<tr>
<td>OUES 90</td>
<td>0.14 (-0.51, 0.8)</td>
<td>0.109</td>
<td>0.84 (0.18, 0.98)</td>
<td>-0.08 (-0.63, 0.46)</td>
<td>0.094</td>
<td>0.88 (0.48, 0.98)</td>
</tr>
<tr>
<td>OUEP</td>
<td>-0.97 (-8.45, 6.52)</td>
<td>0.085</td>
<td>0.81 (0.33, 0.96)</td>
<td>-2.48 (-10.50, 5.03)</td>
<td>0.087</td>
<td>0.77 (0.05, 0.96)</td>
</tr>
<tr>
<td>$O_2$ Pulse</td>
<td>0.23 (-2.01, 2.48)</td>
<td>0.068</td>
<td>0.97 (0.83, 1)</td>
<td>-0.66 (-3.78, 2.46)</td>
<td>0.097</td>
<td>0.94 (0.21, 1)</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{\text{CO}_2}$ slope 1</td>
<td>-0.83 (-4.28, 2.61)</td>
<td>0.073</td>
<td>0.56 (-0.26, 0.91)</td>
<td>0.13 (-2.77, 3.03)</td>
<td>0.060</td>
<td>0.81 (0.44, 0.94)</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{\text{CO}_2}$ slope 2</td>
<td>-1.11 (-4.33, 2.11)</td>
<td>0.062</td>
<td>0.78 (0.26, 0.95)</td>
<td>1.76 (-2.45, 5.91)</td>
<td>0.076</td>
<td>0.61 (-0.2, 0.92)</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{\text{CO}_2}$ ratio nadir</td>
<td>-0.15 (-3.47, 3.17)</td>
<td>0.070</td>
<td>0.73 (0.09, 0.94)</td>
<td>0.13 (-2.21, 2.48)</td>
<td>0.049</td>
<td>0.9 (0.52, 0.98)</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{\text{CO}_2}$ ratio at AT</td>
<td>-0.52 (-3.3, 2.3)</td>
<td>0.056</td>
<td>0.79 (-0.21, 0.98)</td>
<td>0.51 (-4.50, 2.02)</td>
<td>0.088</td>
<td>0.65 (-0.01, 0.91)</td>
</tr>
<tr>
<td>RER at peak *</td>
<td>0.05 (-0.03, 0.12)</td>
<td>0.029</td>
<td>0.86 (0.54, 0.96)</td>
<td>-0.06 (-0.23, 0.11)</td>
<td>0.071</td>
<td>0.55 (0.03, 0.84)</td>
</tr>
<tr>
<td>$P_{ETCO_2}$ at AT</td>
<td>0.5 (-3.56, 4.56)</td>
<td>0.047</td>
<td>0.77 (0.2, 0.95)</td>
<td>0.38 (-5.64, 6.39)</td>
<td>0.069</td>
<td>0.62 (-0.05, 0.91)</td>
</tr>
<tr>
<td>HR at peak</td>
<td>-3.6 (-19.9, 12.7)</td>
<td>0.048</td>
<td>0.38 (-0.52, 0.88)</td>
<td>2.6 (-12.7, 18)</td>
<td>0.045</td>
<td>0.74 (0.44, 0.9)</td>
</tr>
<tr>
<td>$R_t$ †</td>
<td>-0.41 (-5.88, 5.06)</td>
<td>0.076</td>
<td>0.78 (0.48, 0.91)</td>
<td>3.61 (-4.05, 11.28)</td>
<td>0.101</td>
<td>0.48 (0.15, 0.71)</td>
</tr>
<tr>
<td>Peak $\dot{V}_E$</td>
<td>-2.68 (-31.25, 25.84)</td>
<td>0.143</td>
<td>0.89 (0.62, 0.97)</td>
<td>4.24 (-27.78, 36.25)</td>
<td>0.155</td>
<td>0.85 (0.58, 0.91)</td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ Slope ††</td>
<td>-1.05 (-1.77, -0.32)</td>
<td>0.040</td>
<td>0.02 (-0.55, 0.58)</td>
<td>0.69 (-0.58, 1.96)</td>
<td>0.064</td>
<td>0.47 (0.12, 0.71)</td>
</tr>
<tr>
<td>Maximum Wattage ††</td>
<td>28.5 (-14.9, 71.9)</td>
<td>0.084</td>
<td>0.87 (0.43, 0.98)</td>
<td>-27.38 (-58.1, 3.4)</td>
<td>0.066</td>
<td>0.87 (0.46, 0.97)</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Slope</td>
<td>-0.001 (-0.01, 0.01)</td>
<td>0.160</td>
<td>0.86 (0.59, 0.95)</td>
<td>0.004 (-0.01, 0.02)</td>
<td>0.215</td>
<td>0.78 (0.3, 0.94)</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ Intercept</td>
<td>1.84 (-37.6, 41.27)</td>
<td>0.327</td>
<td>0.06 (-0.35, 0.44)</td>
<td>-3.27 (-42.91, 36.36)</td>
<td>0.346</td>
<td>0.07 (-0.47, 0.58)</td>
</tr>
</tbody>
</table>

Table 5.3: The mean difference, standard deviation of the difference: mean ratio and the bootstrapped intraclass correlation coefficient with confidence interval when comparing the first standard test to either the steep or shallow tests undertaken by 8 healthy volunteers (a negative difference indicates that the steep or shallow protocol had lower values than the standard test). Slope 1 refers to the slope calculated up until the VCP. Slope 2 refers to the slope calculated with all exercise data. LOA = limits of agreement. CI = confidence intervals. * p<0.05 for steep vs standard mean value, † p<0.05 for shallow vs standard mean value.
5.4.5 Study 2 – Patient and test characteristics

100 patients were recruited into the Observational study to identify an ideal variable on CPX to distinguish between cardiac and respiratory limitation. These patients all underwent 2 baseline CPX tests. The details of these 100 patients, and exclusion criteria of 4 patients will be discussed in Chapter 6. From the 96 patients deemed appropriate for the study, 3 did not have acceptable data from 2 tests although they had completed the study protocol. The reproducibility statistics are based on the final 93.

Of these 93 patients (70 male); 24 had a primary diagnosis of COPD, 43 a primary diagnosis of heart failure and the remaining 26 mitral valve disease requiring surgery. Some patients had a secondary problem (for example COPD is not uncommon in my heart failure cohort and some of the mitral regurgitation patients had impaired left ventricular systolic function). The reproducibility analysis will be based on their primary diagnosis only.

The patient characteristics of the total group and by disease category are shown in Table 5.4. 18 from 26 patients with mitral regurgitation were symptomatic.

On analysis of CPX variables, OUES/kg, both measures of the $\dot{V}_E/\dot{V}_{CO_2}$ slope, all measures of the $\dot{V}_E/\dot{V}_{CO_2}$ ratio, resting heart rate, arterial oxygen saturations, breathing reserve at the AT and peak, and the HR – $\dot{V}_{O_2}$ slope all showed a significantly non-normal distribution (p<0.0026 by Shapiro-Wilk test). All measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship were made marginally more normally distributed by logarithmic transformation, but resting heart rate was the only variable which became significantly more normally distributed following logarithmic transformation.

On analysis of first test exercise times there was a non-significant trend to shorter times within the COPD group and longer times within the mitral group. The protocol chosen for test 2 reflected this difference in test 1 time, with a non-significant trend to the highest average protocol chosen in the MV group and lowest in the COPD group. Figure 5.3 shows the proportion of ramp protocols chosen for test 2. 84 from 93 patients had an identifiable anaerobic threshold. Only 28 patients reached and exceeded a VCP.
<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>COPD (n=24)</td>
<td>CHF (n=43)</td>
<td>MV (n=26)</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>65.2 ± 10.9</td>
<td>66.6 ± 9.5</td>
<td>66.6 ± 11.1</td>
<td>61.7 ± 11.3</td>
</tr>
<tr>
<td>Male (n, %)</td>
<td>70, 75.3%</td>
<td>16, 66.7%</td>
<td>38, 88.4%</td>
<td>16, 61.5%</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>170.1 ± 10.0</td>
<td>166.8 ± 7.8</td>
<td>171.3 ± 9.2</td>
<td>171.2 ± 12.3</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>77.6 ± 15.8</td>
<td>72.0 ± 16.1</td>
<td>82.7 ± 15.5</td>
<td>74.4 ± 13.8</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.7 ± 4.4</td>
<td>25.8 ± 4.6</td>
<td>28.1 ± 4.1</td>
<td>25.3 ± 4.0</td>
</tr>
<tr>
<td>Diabetics (n, %)</td>
<td>10, 10.7%</td>
<td>1, 4.2%</td>
<td>8, 18.6%</td>
<td>1, 3.8%</td>
</tr>
<tr>
<td>Current smokers (n, %)</td>
<td>12, 12.9%</td>
<td>5, 20.8%</td>
<td>7, 16.3%</td>
<td>0, 0%</td>
</tr>
<tr>
<td>Hypertensive (n, %)</td>
<td>40, 43.0%</td>
<td>10, 41.7%</td>
<td>19, 44.2%</td>
<td>11, 42.3%</td>
</tr>
<tr>
<td>BNP (pg/mL)</td>
<td>86 (37, 203)</td>
<td>36 (14, 52)</td>
<td>118 (85, 286)</td>
<td>85 (43, 222)</td>
</tr>
<tr>
<td>FVC (L)</td>
<td>3.29 (2.59, 3.94)</td>
<td>2.75 (2.34, 3.18)</td>
<td>3.59 (3.01, 4.31)</td>
<td>3.29 (2.75, 3.88)</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>97.2 ± 23.0</td>
<td>88.3 ± 26.4</td>
<td>100.1 ± 18.0</td>
<td>100.5 ± 25.8</td>
</tr>
<tr>
<td>FEV₁ (L)</td>
<td>2.21 ± 0.89</td>
<td>1.32 ± 0.50</td>
<td>2.52 ± 0.73</td>
<td>2.50 ± 0.87</td>
</tr>
<tr>
<td>FEV₁ (% predicted)</td>
<td>79.6 ± 25.7</td>
<td>52.8 ± 19.8</td>
<td>89.7 ± 20.2</td>
<td>87.7 ± 21.4</td>
</tr>
<tr>
<td>FEV₁:FVC ratio</td>
<td>67.2 (55.9, 74.0)</td>
<td>46.2 (38.2, 56.2)</td>
<td>71.1 (62.7, 77.0)</td>
<td>68.6 (65.1, 75.1)</td>
</tr>
<tr>
<td>DLCO (mL/min/mmHg)</td>
<td>5.97 ± 2.11</td>
<td>4.89 ± 1.92</td>
<td>6.19 ± 1.94</td>
<td>6.61 ± 2.25</td>
</tr>
<tr>
<td>Ramp protocol (W/min)</td>
<td>11.4 ± 3.3</td>
<td>10.5 ± 2.8</td>
<td>11.3 ± 2.9</td>
<td>12.3 ± 4.0</td>
</tr>
<tr>
<td>Resting heart rate (bpm)</td>
<td>78.1 ± 14.3</td>
<td>85.5 ± 16.8</td>
<td>72.0 ± 11.2</td>
<td>81.3 ± 12.8</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (mL/min/kg)</td>
<td>17.2 ± 4.2</td>
<td>17.4 ± 4.2</td>
<td>16.3 ± 3.6</td>
<td>18.6 ± 5.0</td>
</tr>
<tr>
<td>% Predicted Peak $V_{O_2}$</td>
<td>73.9 ± 17.0</td>
<td>75.6 ± 18.6</td>
<td>71.0 ± 14.4</td>
<td>77.2 ± 19.1</td>
</tr>
<tr>
<td>AT (% Predicted $V_{O_2}$)</td>
<td>53.7 ± 13.3</td>
<td>56.3 ± 12.5</td>
<td>51.7 ± 12.9</td>
<td>55.2 ± 14.5</td>
</tr>
</tbody>
</table>

Table 5.4: Patient characteristics overall and for each of the 3 disease categories, chronic obstructive pulmonary disease (COPD), chronic heart failure (CHF) and mitral valve disease (MV). Data are displayed as mean ± SD for normally distributed variables and median (IQR) for non-normally distributed variables.
Figure 5.3: Proportion of each ramp protocol chosen for test 2 based on the results of test 1.
5.4.6 Study 2 – Difference between test 1 and 2

Table 5.5 shows the mean values for test 1 and test 2 for 35 CPX variables. The mean within-patient difference was calculated and shown along with the 95% limits of agreement of the difference. This allows clinicians, when comparing two exercise tests in the same patient, to see if the magnitude of change lies outside of this range, thereby indicating the likely presence of a significant deterioration or improvement (rather than inherent variation).

The Bland-Altman plots for 9 of the most commonly used variables are shown in Figure 5.4. Bland-Altman plots are still appropriate when performed on non-parametric data so no log transformations were done for the basic plots. Peak $\dot{V}_{O_2}$ and $\dot{V}_{O_2}$ at the AT showed a significant relationship between the residuals and magnitude of the mean values ($p<0.05$ by Bradley-Blackwood F statistic). For peak $\dot{V}_{O_2}$, there was a significant decrease in the difference between tests (test 2 minus test 1) as peak $\dot{V}_{O_2}$ increased, so that for the patients with the greatest peak $\dot{V}_{O_2}$ there was on average a slight decline from test 1 to test 2. For $\dot{V}_{O_2}$ at the AT this relationship is in the opposite direction. This is reflected in the Bland-Altman plots for these 2 variables with non-horizontal limits of agreement.

All measures of peak $\dot{V}_{O_2}$ showed a significant mean increase between tests 1 and 2 (+24 mL/min, $p=0.046$; +0.35 mL/min/kg, $p=0.024$; +1.6%, $p=0.017$). The $\dot{V}_{O_2}$ at the AT also showed a significant mean increase between tests 1 and 2 (+47 mL/min, $p<0.001$; +2.8%, $p<0.001$). $P_{ETCO_2}$ at the AT (-0.63 mmHg, $p=0.02$), breathing reserve at the AT (-1.5%, $p=0.02$ by Wilcoxon rank-test), breathing reserve at peak (-1.9%, $p=0.02$ by Wilcoxon rank-test), peak minute ventilation (+2 L/min, $p=0.02$), peak respiratory frequency (+1.7 breaths per minute, $p=0.001$), peak heart rate (+3.5 bpm, $p<0.01$) and double product (+ 671 mmHg bpm, $p=0.03$) also showed significant differences between test 1 and test 2 respectively. All measures of the OUES, the OUEP, $O_2$ pulse, measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship, the $\dot{V}_{O_2} – WR$ slope and the measures of the $HR – \dot{V}_{O_2}$ relationship did not display a systematic difference from test 1 to test 2.
Table 5.5: Test-retest reliability of the full cohort. Coefficients of variation calculated via logarithmic transformation of the raw data for variables (starred) where residuals are non-normally distributed. 

† p<0.05 for significant difference between tests 1 and 2. For abbreviations see text.
Figure 5.4: Bland–Altman plots of 9 important CPX variables. To maximise viewing of the data points axes are deliberately minimally labelled. The x-axis of each plot is the mean values of tests 1 and 2, with the y-axis the difference between the 2 tests. The 9 variables are peak $\dot{V}_{O_2}$ (mL/min), $\dot{V}_{O_2}$ at the AT (mL/min), $\dot{V}_E$/ $\dot{V}_{CO_2}$ slope, OUES (L/min/10-fold increase in $\dot{V}_E$), OUEP, $\dot{V}_E$/ $\dot{V}_{CO_2}$ ratio at nadir, $O_2$ pulse (mL/beat), breathing reserve at the AT (%) and $\dot{V}_{O_2} - WR$ slope (mL/min/W).
**5.4.7 Study 2 – Test-retest reliability of various CPX variables**

In a similar manner to the healthy cohort, the ICC and SDD/mean (as a simple measure of the coefficient of variation) were calculated for the patient cohort. Separately the coefficients of variation were calculated using the logarithmic transformation of the raw values for each variable that showed dispersion of the residuals with increasing magnitude, as noted on the Bland-Altman plots of O$_2$ pulse, $\dot{V}_E$/\dot{V}$_{CO2}$ slopes and ratios and minute ventilation. For these variables the logarithmic coefficients of variation are shown, rather than the SDD/mean coefficient of variation, but these are equivalent. These measures are all shown in Table 5.5. The confidence interval for each variable’s ICC allows a visual comparison of two ICC values to assess for a significantly stronger test-retest reliability in one variable over another.

It can be seen for gas exchange variables that peak $\dot{V}_{O2}$, OUES measured using the full data and weight adjusted, OUEP, the O$_2$ pulse, peak circulatory power, and the $\dot{V}_E$/\dot{V}$_{CO2}$ ratio at the VCP and nadir, and breathing reserve at the AT and peak all had excellent test-retest reliability as defined by an ICC ≥0.90. Despite small differences in the ICC values and noticeable overlap in the CIs, only circulatory power, O$_2$ pulse and the breathing reserve at AT were not significantly weaker than peak $\dot{V}_{O2}$. Double product, peak watts and heart rate also showed high test-retest reliability. Gas exchange variables with only fair or moderate test-retest reliability (0.2 ≤ICC ≤0.69) included RER at rest and the AT and resting arterial O$_2$ saturations. All other measures including the foreshortened OUES data, $\dot{V}_{O2}$ at the AT, the $\dot{V}_E$/\dot{V}$_{CO2}$ slope and $\dot{V}_E$/\dot{V}$_{CO2}$ ratio at the AT, the $\dot{V}_{O2} - WR$ slope, and the $HR - \dot{V}_{O2}$ relationship showed good test-retest reliability (0.7 ≤ICC <0.89).

The $\dot{V}_E$/\dot{V}$_{CO2}$ ratio measured at nadir was significantly more reproducible than both the ratio measured at the anaerobic threshold (p=0.002) and the P$_{ET}CO_2$ (p=0.001) at the anaerobic threshold. The slopes of the $\dot{V}_E$/\dot{V}$_{CO2}$ relationship using full data or only data up to the VCP showed similar test-retest reliability, and these slopes did not have significantly lower ICC values than the $\dot{V}_E$/\dot{V}$_{CO2}$ ratio measured at nadir.

OUES had significantly better test-retest reliability than the OUES measured using foreshortened data from the 25-75$^{th}$ percentiles, 0-50$^{th}$ percentiles, 0-70$^{th}$ percentiles and 0-90$^{th}$ percentiles. OUES was similar to OUEP (p=0.45).

When assessing reproducibility using coefficients of variation peak $\dot{V}_{O2}$, OUEP, O$_2$ pulse, the $\dot{V}_E$/\dot{V}$_{CO2}$ ratio at the VCP, AT and nadir, both measures of the $\dot{V}_E$/\dot{V}$_{CO2}$ slope, and the P$_{ET}CO_2$ at the AT all had excellent test-retest reliability (CoV <0.10). OUES, OUES/kg, DP and peak circulatory power had slightly worse values for
The breathing reserve at peak, the $HR - \dot{V}_{O_2}$ relationship and foreshortened OUES measures had poor coefficients of variation (>0.20). The biggest discrepancies between CoV and ICC values can be seen with the arterial oxygen saturations, which at both rest and peak showed excellent CoV values (0.02) but poor ICC values (0.60 and 0.80 respectively). The BR also displayed some disparities; at anaerobic threshold it had excellent CoV and ICC, whilst at peak, despite a similar ICC, the CoV was vastly different. This probably can be explained by the nature of CoV which cannot manage negative numbers (10 patients had a negative BR one either test 1, 2 or both).

**5.4.8 Study 2 - Determinants of test-retest reliability**

6 factors (age, gender, BMI, disease category, change in ramp protocol from test 1 to 2, and tests on the same or different days) were assessed for their potential to influence the test-retest reliability of CPX variables in 2 ways: firstly through a mixed regression model looking for an interaction between the factor and condition (a significant interaction implies that the potential determinant influences the difference between 2 tests, i.e. its repeatability); and secondly by comparing the confidence limits of ICCs calculated for each group. Only the principal CPX variables were assessed in this way to avoid achieving statistical significance because of multiple measures. Table 5.6 shows the influence of disease category on mean test-retest difference, CoV and ICC. Determinants could affect in one of 2 ways, by affecting the mean difference between 2 tests, or by affecting the between-test variance.

Gender affected reproducibility minimally. The $\dot{V}_E/\dot{V}_{CO_2}$ slope had a significantly lower ICC in females (0.71 vs 0.93 compared to males $p=0.002$) and females exhibited a greater mean difference in heart rate between tests at both rest and peak exercise, with non-significantly worse ICCs.

Age (divided into 2 groups by median age) only affected the repeatability of the $\dot{V}_{O_2} - WR$ slope, with a significantly lower ICC (0.46) in older patients compared with younger (0.86) ($p = 0.0001$).

A BMI above 25 led to significantly poorer repeatability of the OUES (ICC 0.90 vs 0.96 against BMI $\leq 25$, $p=0.02$), $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir and AT (ICC 0.88 vs 0.96, $p=0.005$; 0.88 vs 0.96, $p=0.005$ respectively), and breathing reserve at the AT (ICC 0.77 vs 0.91, $p=0.005$). Double product and peak heart rate were more reproducible in heavier individuals. Comparing mean differences, the $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir, $\dot{V}_{O_2} - WR$ slope and resting heart rate all fell from test 1 to 2 in patients of normal weight, but rose in those where BMI >25.
When comparing reliability between disease categories, the $\dot{V}_O_2$ at AT had a non-significant but slightly weaker reliability in the CHF group than the COPD group and it can be seen that there was a much greater difference between tests in the mitral group than the COPD group. There was a significantly improved reliability of the OUES in patients with mitral valve disease ($p=0.01$) over the other 2 conditions, however ICC was 0.88 or above in all 3 groups. Category of disease influenced the reliability of the $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir and at AT with better ICC values and smaller mean differences in the COPD and mitral groups compared to the CHF group. Patients with mitral valve disease had a larger mean difference in resting and peak heart rate between tests, but the ICC was not different between groups, whilst the $HR - \dot{V}_O_2$ slope was significantly less reliable in the COPD group.

There was a significant interaction between test and inter-test interval for $\dot{V}_O_2$ at the AT, with similar ICC values but a much greater mean difference in patients performing the tests on the same day (+69 mL/min vs -70 mL/min). There was almost no mean difference in O$_2$ pulse when tests were performed on the same day (-0.1 mL/beat), with a large increase on different days (+0.5 mL/beat), and significantly worse ICC values when measured on different days compared to the same day (ICC 0.85 vs 0.96, $p=0.03$). Patients undergoing the tests on the same day had a greater mean difference in heart rate, both at rest and peak, but ICCs were not significantly different. The $\dot{V}_O_2 - WR$ slope was significantly less reliable when measured on different days compared to the same day (ICC 0.37 vs 0.76, $p=0.03$).

Ramp protocol affected a number of variables’ reliability. O$_2$ pulse was significantly less reliable in patients remaining on the 10W protocol compared to moving to either a shallower ($p=0.0007$) or steeper ($p=0.03$) protocol; the $\dot{V}_E/\dot{V}_{CO_2}$ slope was less reliable in the patients moving to a shallower protocol, with the best ICC in the group moving to a steeper protocol. There were changes to the between test mean difference of double product, circulatory power, maximum watts and the $\dot{V}_O_2 - WR$ slope; the latter showing similar changes to those seen in Study 1 in healthy individuals (mean 0.13 mL/min/W decrease in slope in those moving to steeper protocols, and a mean 0.63 mL/min/W increase in slope in shallower protocols), with similar ICC values in patients on similar or harder protocols, but a significantly worse ICC in those requiring easier protocols. Peak heart rate and the $HR - \dot{V}_O_2$ slope displayed the strongest ICCs in patients moving to the steeper protocols and the worst in those remaining on the 10W protocol.

Overall the test-retest reliability of most variables was largely resistant to these 6 potential determinants. Age and gender affected reproducibility minimally; whilst weight affected the ICCs for a few variables, generally
with gas exchange variables worsening, albeit to a small degree, in the heavier individuals. Inter-test interval did not affect the majority of variable’s ICCs. The two determinants that made the most difference are disease category and ramp protocol. However within disease category there was no single disease that consistently outperformed the others; whereas for ramp protocol it was largely the patients moving to the steepest ramp that displayed the strongest test-retest reliability.

Peak $\hat{V}_{O_2}$, OUEP and to a lesser extent the $O_2$ pulse and OUES were largely unaffected. Conversely the $\hat{V}_{O_2} - WR$ slope was affected by 4 of the 6 determinants.

5.4.9 Changes in slopes throughout testing

I have plotted sparkline plots of the OUES, $\dot{V}_E/\dot{V}_{CO_2}$ slope and $\dot{V}_{O_2} - WR$ slope for 8 healthy individuals, 8 randomly chosen patients with heart failure and 8 randomly chosen patients with COPD, to examine the statement that slopes are reproducible because they do not change as the test evolves. This is a qualitative view only; it may be of interest but does not lend itself to statistical techniques. Sparkline plots do not have labelled axes. The x-axis in all plots represents time, but scales will differ between plots. The y-axes are the three slopes calculated using data only until that time point, therefore the value at the right-most point of the plot represents the final value and everything before that ever-decreasing foreshortened data (Figure 5.5). The $\dot{V}_E/\dot{V}_{CO_2}$ slope plots utilises data throughout exercise and so we would expect the slope to rise towards the end of exercise if the ventilatory compensation point is reached and passed.

We can see that OUES was largely planar for a significant proportion of latter exercise. In a couple of patients with COPD it appeared to be slightly rising at termination of exercise, this pattern was not seen in the CHF patients.

For the $\dot{V}_{O_2} - WR$ slope it was interesting to see that no slopes appear to decline towards the end of the test, generally they were planar with some of the COPD patients experiencing a rise.
## Table 5.6: Test-retest reliability measurements between disease categories for selected variables.

<table>
<thead>
<tr>
<th>COPD (n=24)</th>
<th>CHF (n=43)</th>
<th>MV (n=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td><strong>Mean diff</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td>(+LOA)</td>
<td>(+LOA)</td>
<td>(+LOA)</td>
</tr>
<tr>
<td><strong>SDD/mean</strong></td>
<td><strong>SDD/mean</strong></td>
<td><strong>SDD/mean</strong></td>
</tr>
<tr>
<td><strong>ICC</strong></td>
<td><strong>ICC</strong></td>
<td><strong>ICC</strong></td>
</tr>
</tbody>
</table>

- **Peak $V_{O2}$ (mL/min)**: 1220, Mean diff 23.9 (-230, 278), Mean 0.11, SDD/mean 0.91 (0.85, 0.95), ICC 0.92 (0.84, 0.96)
- **Peak $V_{O2}$ (mL/min/kg)**: 17.3, Mean diff 0.33 (-3.46, 4.11), Mean 0.11, SDD/mean 0.88 (0.79, 0.94), ICC 0.95 (0.91, 0.97)
- **$V_{O2}$ at AT (mL/min)**: 892, Mean diff 15 (-174, 204), Mean 0.11, SDD/mean 0.88 (0.74, 0.95), ICC 0.95 (0.91, 0.97)
- **OUES (L/min/10-fold $V_{O2}$)**: 1.93, Mean diff 0 (-0.56, 0.56), Mean 0.15, SDD/mean 0.88 (0.79, 0.93), ICC 0.95 (0.91, 0.97)
- **OUEP**: 31.7, Mean diff 0.74 (-3.25, 4.72), Mean 0.06, SDD/mean 0.93 (0.87, 0.96), ICC 0.94 (0.83, 0.95)
- **O2 Pulse (mL/beat)**: 9.98, Mean diff 0.03 (-1.56, 1.63), Mean 0.08, SDD/mean 0.92 (0.82, 0.96), ICC 0.94 (0.83, 0.95)
- **$V_{E}/V_{CO2}$ slope**: 33.2, Mean diff -0.61 (-9.12, 7.91), Mean 0.13, SDD/mean 0.86 (0.75, 0.92), ICC 0.94 (0.83, 0.95)
- **$V_{E}/V_{CO2}$ ratio nadir**: 34.5, Mean diff -0.43 (-4.66, 3.79), Mean 0.06, SDD/mean 0.95 (0.91, 0.98), ICC 0.94 (0.83, 0.95)
- **$V_{E}/V_{CO2}$ ratio at AT**: 36.7, Mean diff -0.98 (-7.91, 5.96), Mean 0.10, SDD/mean 0.88 (0.7, 0.96), ICC 0.95 (0.83, 0.95)
- **RER at peak**: 1.01, Mean diff -0.01 (-0.13, 0.11), Mean 0.06, SDD/mean 0.84 (0.72, 0.91), ICC 0.95 (0.83, 0.95)
- **HR at rest (bpm)**: 84.7, Mean diff 1.54 (-13.8, 16.9), Mean 0.09, SDD/mean 0.85 (0.67, 0.94), ICC 0.95 (0.83, 0.95)
- **HR at peak (bpm)**: 125, Mean diff 1.67 (-19, 22.4), Mean 0.08, SDD/mean 0.83 (0.58, 0.93), ICC 0.95 (0.83, 0.95)
- **DP (mmHg bpm)**: 22442, Mean diff 951 (-4963, 6864), Mean 0.13, SDD/mean 0.86 (0.67, 0.94), ICC 0.95 (0.83, 0.95)
- **Circulatory Power (mmHg mL/min)**: 218478, Mean diff 9786 (-49792, 69365), Mean 0.14, SDD/mean 0.9 (0.8, 0.95), ICC 0.95 (0.83, 0.95)
- **BR at AT**: 49.5, Mean diff 0.66 (-9.43, 10.74), Mean 0.10, SDD/mean 0.91 (0.79, 0.96), ICC 0.95 (0.83, 0.95)
- **Breathing Reserve (%)**: 10.0, Mean diff 0.1 (-20.8, 21), Mean 1.07, SDD/mean 0.86 (0.68, 0.94), ICC 0.95 (0.83, 0.95)
- **$V_{O2} - WR$ slope (mL/min/W)**: 9.35, Mean diff 0.24 (-2.76, 3.23), Mean 0.16, SDD/mean 0.61 (0.15, 0.85), ICC 0.95 (0.83, 0.95)
- **Maximum Wattage (W)**: 80, Mean diff 2 (-18.6, 22.6), Mean 0.13, SDD/mean 0.92 (0.84, 0.96), ICC 0.95 (0.83, 0.95)
- **HR – $V_{O2}$ slope**: 0.05, Mean diff 0 (-0.03, 0.03), Mean 0.30, SDD/mean 0.64 (0.14, 0.88), ICC 0.95 (0.83, 0.95)

**Note:** LOA = limits of agreement. For other abbreviations see text.
Figure 5.5: Sparkline plots for the change in OUES, $\dot{V}_E/\dot{V}_{CO_2}$ slope, and $\dot{V}_{O_2} - WR$ slope in 24 participants over time. Data from 8 healthy participants, 8 patients with CHF and 8 patients with COPD were analysed. The x-axis on all is time, although this will differ between each test they are all scaled differently. Y-axes are the 3 slopes as they evolve over time, using data measured from the start of the test to each instantaneous point.


5.5 Discussion

Cardiopulmonary exercise testing is regarded to have an excellent degree of reproducibility, as assessed in multiple studies (Elborn et al 1990, Lehmann et al 1996, Marburger et al 1998, Skinner et al 1999, Hansen et al 2004, Bensimhon et al 2008, Keteyian et al 2010). Previous studies have largely focussed on CPX as a whole, rather than comparing variables to identify differences in repeatability between them, and also if any other factors can influence the repeatability. Another problem with quotations for reproducibility statistics in medicine is that studies often perform under the most optimal conditions. They take optimal patients and have experts in the field performing the tests. This is not the picture seen in day-to-day clinical practice.

Here I show multiple measures of the test-retest reliability of multiple CPX variables performed under typical clinical conditions, and how they compare to one another in a population with 3 forms of cardiorespiratory disease, and in a separate healthy population. I also show how age, gender, weight, disease category and test-retest interval influence these results in the patient population, and importantly how changing the ramp protocol influences reliability in both the patients and healthy adults.

5.5.1 Test-retest reliability and the measures of reproducibility

The large patient group (93 patients) showed excellent test-retest reliability for peak $\dot{V}_O_2$, OUES measured using the full data and weight adjusted, OUEP, the $O_2$ pulse, peak circulatory power, the $\dot{V}_E/\dot{V}_CO_2$ ratio at the VCP and nadir, and breathing reserve at the AT and peak. However only the ICCs for peak circulatory power, $O_2$ pulse and the breathing reserve at the AT, were not significantly worse than peak $\dot{V}_O_2$. Within the healthy adult study, when comparing each participant’s 2 standard tests, the ICCs are generally weaker than for the patient population, with the exception of peak $\dot{V}_O_2$, $\dot{V}_O_2$at the AT and the $\dot{V}_O_2 - WR$ slope. This may just reflect a much smaller cohort (n=8). It has been suggested that slopes will have a greater test-retest reliability when compared to peak values as they are resistant to differences in effort. My findings would largely dispute this concept. Peak $\dot{V}_O_2$ and $O_2$ pulse are the most reliable of gas analysis variables. Whilst OUES had similar ICC and CoV values, the $\dot{V}_E/\dot{V}_CO_2$ slope and the $\dot{V}_O_2 - WR$ slope had significantly worse values, than peak $\dot{V}_O_2$. This is most noticeable for the $\dot{V}_O_2 - WR$ slope with an ICC of only 0.70 despite the components, peak $\dot{V}_O_2$ and maximum work rate, both being individually very reliable. For the main variables these measures of reproducibility are very similar to those seen in the previously described trials, performed under optimal research conditions.
The concept of reproducibility is a difficult one and no method is without limitations. Let us consider the ICC. It is a function of the within-patient variance and the total group variance. There are a couple of significant problems with this. Firstly its magnitude is dependent on the variance of the whole group, but that is not what we as clinicians find relevant; it is only the within-patient variance that we are interested in. The magnitude of an ICC can be increased by increasing the variance of the population studied, i.e. by choosing a more heterogeneous group. This is quite relevant in CPX testing as variables such as peak $\dot{V}_{O_2}$ will have a wide range of values (for example the range of peak $\dot{V}_{O_2}$ within our patients was 546.5 mL/min – 2248.9 mL/min, a more than 4-fold increase). It is not that this incorrectly inflates the ICC in variables like peak $\dot{V}_{O_2}$, but it must be considered when comparing variables whose magnitudes may not behave similarly to one another. An example of this is $P_{ET}CO_2$ at the AT, which within the same population had a range of 22-50 mmHg, much smaller than that of peak $\dot{V}_{O_2}$. A potential way to overcome this limitation is to use a measure that does not involve the whole group variance. Coefficients of variation, which I display here alongside the ICC for each variable, are a function of the ratio of within-subject standard deviation to the mean of the whole group. This standardises the between patient variance but is influenced heavily by the mean value chosen. A commonly described example of a problem with CoV is when a temperature measuring device is tested for reliability. Temperature measured in Kelvin or Celsius will show identical variance but because they exist on very different scales the CoV measurements will be vastly different. Hence CoV measurements are generally only applicable to variables that range from zero upwards in a linear fashion without negative values. Breathing reserve in our analysis will be highly susceptible to anomalous CoV measurements as the mean is unrepresentative because it is influenced by a number of negative values; this will artificially worsen CoV. In contrast to this, variables that display a clinical range an order of magnitude below their values will display artificially good CoV measurements. An example of this is arterial oxygen saturations, which rarely fall below 90% before a test is stopped, hence there exists a possible clinical range of only 0-10 readings, but where mean values will be approximately 10 times this. Therefore both ICC and CoV must be viewed together, and the limitations of both considered for each variable separately.

The second problem with ICC, and also a problem for CoV, is that the within-patient variance can be relatively unaffected by systematic bias between two tests, i.e. a persistent effect where a second test is higher or lower than the first. ICCs were developed to compare 2 or more tests where the values are interchangeable, i.e. test 1 could be test 2 and vice versa. This may not be the case with clinical exercise testing which may display an ordering effect; this may be familiarisation, which I will discuss later, or fatigue. The presence of a systematic
bias may not influence the ICC but will certainly influence a clinician’s interpretation of 2 sequential tests, therefore knowledge of the mean difference between two tests (and the limits of agreement) is vital to determine if a patient’s clinical situation has altered. What is helpful about ICC measurements is that they are non-scalar, i.e. they are comparable between any unit of measurement. This allows a comparison of two ICC values, which I have performed in this analysis in the form of a z-score of Fisher transformation.

5.5.2 The difference between 2 tests – A familiarisation effect?
It was proposed many years ago that CPX results can be improved upon on subsequent tests (Elborn et al 1990). A substantial difference on treadmill testing existed between tests 1 and 2, with a much smaller difference on subsequent tests. This has been termed a learning effect, but has largely not been reproduced to the same degree in any study since. Russell et al suggested that the increase in exercise time found by Elborn et al was largely due to improved efficiency of walking, as they also found evidence of an increase between tests 1 and 2 (but not 2 and 3) in exercise time, but unlike Elborn et al, they found no significant change in peak $\dot{V}_O_2$ (Russell et al 1998). Similarly Bensimhon et al showed a mean increase in exercise time between tests 1 and 2, without a concomitant increase in peak $\dot{V}_O_2$ (Bensimhon et al 2008). I show that there was a significant increase in peak $\dot{V}_O_2$ between tests 1 and 2, however this had a mean value of approximately 24 mL/min, only about 2.5% of the peak values, suggesting that whilst it is statistically significant, it doesn’t appear to be clinically relevant. $\dot{V}_O_2$ at the AT also shows a significant bias with statistically significantly higher values on test 2, whilst measures of the breathing reserve show the opposite, with significantly lower values on test 2; the magnitude of this systematic bias remains small. So these results largely disagree with the original findings by Elborn et al which showed a very large learning effect. A small, clinically non-relevant learning effect is apparent for peak $\dot{V}_O_2$, AT and breathing reserve.

5.5.3 The effect of potential determinants on test-retest reliability
Almost all previous studies of reproducibility have failed to assess factors that may influence the reproducibility. Bensimhon et al attempted to answer this question using multivariate analysis. A low peak $\dot{V}_O_2$ on test 1 predicted a larger change on test 2; however this could just represent regression to the mean. Other
factors analysed did not affect the reproducibility (as defined as a change in peak \( \dot{V}_{O_2} \)) and included site experience, gender, symptom status, ethnicity and cause of heart failure.

I assessed within the patient study, 6 potential determinants for their influence on test-retest reliability. One potential problem with this analysis is the problem with multiple testing. For almost 20 variables, for 6 determinants, p values were calculated. There was no a priori hypothesis for which variables and determinants should be affected so a correction factor should be employed. This would lead to a p value for significance of 0.0026. None of the ICC measurements were that significantly different between groups, therefore all the significant differences described in the results section must be interpreted with caution.

Almost all variables’ repeatabilities appeared unaffected by gender and age. Heavier patients had reduced repeatability of a handful of variables but the absolute difference in ICC was not large. Almost all variables’ repeatabilities were unaffected if the 2 tests were performed on different days compared to the same day suggesting fatigue does not play a role in repeatability. Disease categorisation affected test-retest measures of a few variables including the \( \dot{V}_{O_2} \) at the AT, the OUES, \( \dot{V}_{E}/\dot{V}_{CO_2} \) ratio at nadir and AT, resting and peak heart rate and the \( HR - \dot{V}_{O_2} \) slope. However one disease category did not consistently show an improved or reduced reliability over the other two, suggesting that these differences may largely be down to random variation.

5.5.4 The effect of a change in ramp protocol on test-retest reliability

The final determinant, a change in ramp protocol, is especially relevant and therefore warrants specific discussion. The other 5 determinants (age, gender, weight, type of disease, and inter-test interval) cannot be altered by the physician performing serial tests; their influence on test-retest reliability is unavoidable. However chosen protocol may have the potential to significantly alter results.

Separate to the patient study, a healthy cohort of adults was specifically tested to assess for the influence of a change in ramp protocol on test-retest reliability. Maximum work load and \( \dot{V}_{O_2} \) – \( WR \) slope were significantly affected by both a shallower and steeper ramp compared to the standard ramp. The \( \dot{V}_{O_2} \) – \( WR \) slope had much lower ICC measurements in the standard vs steeper group (ICC=0.02) and standard vs shallower group (ICC=0.47) than the comparison of two standard tests (ICC=0.78), although only when comparing the former with 2 standard tests is this difference in ICC significant. The magnitude of the slope decreased in steeper protocols in
all participants by an average of 1.05, and rose in shallower protocols in 7 of the 8 participants by an average of 0.69. This pattern was also seen in the patient study.

A study (Agostoni et al 2005) has investigated the effect of work rate protocol on exercise test variables and tested heart failure patients using bicycle ergometry at 3 levels of difficulty (individualised to the patient following an initial test). When testing patients to maximum within 5 minutes, peak $\dot{V}_{O_2}$ and the $\dot{V}_{O_2} - WR$ slope were lower than at 10 and 15 minutes. The shallowest ramp protocol (cessation at 15 minutes) had a higher $\dot{V}_{O_2} - WR$ slope than 10 minutes with a similar peak $\dot{V}_{O_2}$. It was postulated that these results may represent the delay in oxygen kinetics displayed by heart failure patients. $O_2$ consumption at the muscles and the corresponding oxygen uptake at the lungs occur later than in healthy controls; this lag means that the measured $\dot{V}_{O_2}$ at any given work rate on an incremental protocol represents the oxygen requirement of a preceding level of the ramp. The faster the ramp, the greater the difference that exists between the $\dot{V}_{O_2}$ we measure and the $\dot{V}_{O_2}$ we expect for the work rate at that instant. I do not believe this has been comprehensively studied in healthy controls, and so my results may be the first to help confirm or refute this hypothesis. If the explanation is a delay in oxygen kinetics, as explained above, we would expect for healthy adults, with no delay to their oxygen kinetics, that these values should be unaffected or at least less affected by ramp intensity. This is not the case in my study, with a greater reduction in the slope in the healthy adults than the heart failure patients. Another explanation proposed for the higher values of the slope in longer, shallower protocols, is that alternative processes that add to oxygen consumption have time to become prominent. These processes include the use of muscle groups alongside the legs (such as arms); less efficient muscle fibre use as fatigue sets in, increased use of oxygen by the respiratory muscles; and the reconversion of lactate to glycogen within the liver which in the faster, steeper protocols “waits” until exercise finishes (Wasserman et al 2005). I know of no studies to prove these hypotheses, but they lead to two important assumptions which can be explored with the results of my study. Firstly it is difficult to explain how these effects would differ between patients with heart disease and healthy adults, i.e. the influence of protocol intensity on $\dot{V}_{O_2} - WR$ slope is equal for all, and that seems probable given my results with the change in slope being of identical direction when altering ramp difficulty between the 2 studies. This is not the same as saying the slope is the same between patients with heart disease and healthy adults, only that the slope behaves similarly in both groups to a change in protocol. Secondly these hypotheses (and one explanation for the difference in slope between normal adults and those with CHF) is that the slope falls after the AT, most noticeably in patients with CHF as a greater proportion of work is performed
by anaerobic sources (and therefore does not require oxygen). As can be seen in Figure 5.5 when the slope is measured adding progressively more of the test in from the beginning of incremental exercise we do not see a clear deterioration in the slope in the latter stages (which would correspond to post-AT) in either CHF or healthy adults. They typically plateau with a rise in the slope in many COPD patients. Therefore the argument that the effect of protocol intensity on the $\dot{V}_{O_2} - WR$ slope only appears after the AT appears incorrect based on my data.

Choice of protocol is relevant because within research they are normally standardised so that all participants undergo the same ramp. In the reproducibility studies described earlier 10 used a standard protocol for all patients with no deviation. Only a single study (Skinner et al 1999) had an easier protocol for more limited patients, without deviation in protocol between test 1 and test 2. It is established practice (although this seems to be largely experiential rather than evidence-based) that to obtain optimal results for your patient an individualised protocol, aiming for maximal exercise within 8-12 minutes, should be employed (Balady et al 2010), which only the Skinner et al study could claim to have done. The only study that appears to have definitively looked into the effect of test duration, did so on only 5 healthy adult males (Buchfuhrer et al 1983). My results suggest that should a policy exist for standardised protocols within a clinical institution or a research trial, then $\dot{V}_{O_2} - WR$ slope may be affected but other variables should be largely unaffected. A different protocol when comparing two tests in the same patient should be interpreted with caution but aside from the $\dot{V}_{O_2} - WR$ slope should be comparable.

**5.5.5 Limitations**

Patients within this study were not blinded to the role of the 2 exercise tests. They were aware that test 1 was a familiarisation test and as such may have been inclined to perform to less maximal effort on this test. However the difference in peak $\dot{V}_{O_2}$ between the two tests is only 2.5%, a much smaller difference than that seen between test 1 and 2 in the study suggesting the need for a familiarisation test (Elborn et al 1990). Peak RER was, on average, identical between test 1 and 2, suggesting equal effort. Some patients had undergone previous CPX testing (largely treadmill) which may also have affected repeatability, especially if they were more common in one group (for example heart failure). However data was not collected as to patient familiarisation with the technique.
The decision to perform both tests on the same day or different days was not randomised. This may affect the validity of the statement that performing 2 tests on the same day does not change the test-retest reliability.

Seven of the eight healthy volunteers were aged under 30; this small range may have affected reproducibility and ideally these results should be confirmed in a cohort of healthy adults with a broader age range.

A criticism of this reproducibility is that all tests were performed only on a bicycle ergometer. As altering protocol had an effect, so too may changing ergometer affect reproducibility. It has previously been shown that cycle exercise is associated with a lower peak $\dot{V}_{O_2}$, $\dot{V}_{E}/\dot{V}_{CO_2}$ slope and $HR - \dot{V}_{O_2}$ slope in patients with CHF compared with treadmill (Witte et al 2005). Another study comparing 3 treadmill protocols and a cycle protocol showed that the $\dot{V}_{O_2}$ at the AT was very reproducible between these four tests, but peak $\dot{V}_{O_2}$ was not (Piña et al 1990). However a direct comparison of the reproducibility of variables between a treadmill and a bicycle CPX test is still lacking, but was just not possible within this study’s protocol.

5.5.6 Conclusions

Many variables obtained during CPX show excellent test-retest reliability. This study shows, for the first time, multiple values of reproducibility for many variables, and allows direct comparison between variables. Peak $\dot{V}_{O_2}$, OUES, $O_2$ pulse, peak circulatory power and the $\dot{V}_{E}/\dot{V}_{CO_2}$ ratio at nadir showed excellent test-retest reliability. Measures of the ventilatory threshold and the $\dot{V}_{E}/\dot{V}_{CO_2}$ slope showed good test-retest reliability. The $\dot{V}_{E}/\dot{V}_{CO_2}$ ratio at nadir was superior to the ratio at the AT, whilst OUES using foreshortened data was inferior to the full OUES.

The difference in variables between two tests was largely unaffected, or only affected to a small degree by age, gender, weight, aetiology of disease, and inter-test interval. The protocol intensity chosen can affect certain variables, the $\dot{V}_{O_2} - WR$ slope was the variable most affected by changes in ramp protocol.

Peak $\dot{V}_{O_2}$ and $\dot{V}_{O_2}$ at the AT showed a significant familiarisation effect, however this was of small magnitude, and unlikely to be of clinical relevance. Most other gas analysis variables did not display this familiarisation effect. Overall there appears to be little necessity for a preliminary test in the majority of patients undergoing bicycle ergometry CPX testing.
6.0 The Observational Study
6.1 Abstract

Cardiopulmonary exercise testing (CPX) has established its role in identifying a limiting pathophysiological state in patients with a variety of cardiorespiratory diseases. However it remains unclear if any variable can show specificity and sensitivity for cardiac over respiratory dysfunction. Certain variables such as the anaerobic threshold (AT) and oxygen uptake to work rate relationship (\(\dot{V}_{O2} - WR\) slope) have been proposed to be specific for cardiac dysfunction, but the relative strengths of various variables for this purpose has not been tested.

96 patients with chronic obstructive airways disease (COPD), chronic heart failure (CHF), severe mitral valve disease (MV) or combined COPD and CHF, were prospectively enrolled in an observational study and underwent CPX testing on a bicycle ergometer. Following a familiarisation test, each patient underwent a personalised second test aiming for maximal exercise after approximately 10 minutes. Multiple variables were measured from this second test. Although patients weren’t matched for exercise capacity, peak \(\dot{V}_O2\) was similar between the groups (mean 17.3 mL/min/kg COPD, 16.4 mL/min/kg CHF and 18.6 mL/min/kg MV).

Breathing reserve (area under curve 0.912) and percent of predicted oxygen uptake efficiency slope (OUES) (AUC 0.877) had the greatest ability to discriminate between COPD and CHF on ROC curve analysis. \(\dot{V}_O2 - WR\) slope performed moderately (AUC 0.676), however the \(\dot{V}_O2\) at the AT was not significantly lower in the CHF patients and did not show any ability to discriminate (AUC 0.569 for AT as percent of predicted peak \(\dot{V}_O2\)).

Breathing reserve was confirmed as an excellent variable discriminating cardiac from respiratory disease. OUES had similar ability to discriminate and importantly abnormal values were found in the CHF and MV groups.

Despite it being considered an important determinant of cardiac dysfunction, the AT could not discriminate in this typical clinical population. The \(\dot{V}_O2 - WR\) slope showed only moderate discriminant ability.
6.2 Introduction

Cardiopulmonary exercise testing (CPX) is recommended for the identification of the key limiting organ in a patient presenting with exercise intolerance or dyspnoea (Balady et al 2010).

Algorithms have been developed which allow the physician to determine limiting physiology using a handful of CPX variables (Neuberg et al 1988, Eschenbacher et al 1990, Milani et al 2004, Wasserman et al 4th Edition 2005). An example of one of these is shown below (Figure 6.1).

Most authors use an algorithm similar to the one displayed here, with peak $\dot{V}_O_2$ used to determine if patients are truly limited and the combination of anaerobic threshold and breathing reserve to thereafter determine cause. Eschenbacher et al added the ventilatory equivalent for CO$_2$, a change in arterial oxygen saturations, ischaemic changes/symptoms and heart rate reserve (which is similar to the $HR - \dot{V}_O_2$ slope I have measured in all) for a more comprehensive analysis. A breathing reserve cut-off of 30% and an anaerobic threshold cut-off of 40% of predicted peak $\dot{V}_O_2$ have typically been used to discriminate. It is difficult to validate these algorithms; even if you identify resting abnormalities in cardiac and/or respiratory function is it possible to be confident that these abnormalities alone limit exercise? Breathing reserve, measured at peak (Nery et al 1983) or at the anaerobic threshold (Medoff et al 1998) appears to have good ability to discriminate patients with known respiratory disease from healthy adults and those with heart disease. The former study also showed that the $\dot{V}_O_2$ at the anaerobic threshold was significantly lower in the cardiac (mitral valve disease) group when compared to a COPD and control group. However this was a small study and I have not seen the results replicated elsewhere.

Whilst an algorithm (if validated) is an excellent approach to identify the limiting physiology in a patient with breathlessness or exercise intolerance, it is not useful in serial measurements of a single patient. Whilst anaerobic threshold may theoretically fall linearly with worsening disease severity the same cannot be said about the breathing reserve. Peak $\dot{V}_O_2$ may be the best way to track changes in a patient with a single disease to establish if they are deteriorating or perhaps improving after the initiation of therapy. Increasingly however, patients are presenting with multiple pathologies. In one study of heart failure clinic patients 40% had spirometry suggestive of COPD (Mascarenhas et al 2008). Therefore should we use peak $\dot{V}_O_2$ to detect changes over time? How can we be sure that a reduction in peak $\dot{V}_O_2$ reflects a worsening of cardiac and not respiratory function?
I aim within this chapter to show, in a prospective cohort of patients with COPD, heart failure, mitral valve disease and a mixture of the former 2 conditions, which variables show the best ability to discriminate broadly between respiratory and cardiac dysfunction. I propose that peak $\dot{V}_{O_2}$ will not show an ability, and will be similarly reduced in all groups; breathing reserve will show good ability to discriminate respiratory patients; and based on the work from the pilot data and other authors’ work, the OUES, the OUEP, the $O_2$ pulse, $\dot{V}_{O_2} - WR$ slope and the $\dot{V}_{O_2}$ at the AT will show good ability to discriminate patients with cardiac disease.

---

6.3 Methods

6.3.1 Patient recruitment

This chapter is the central analysis of the principal Observational cohort recruited for the study. Therefore the recruitment details are explained in detail within the Methods Chapter alongside inclusion and exclusion criteria. Briefly patients from a respiratory clinic with COPD, patients with CHF either from a heart failure clinic or awaiting CRT implantation and patients with severe mitral valve disease awaiting surgical correction were recruited into this exercise study. Patients identified to have both COPD and CHF were not excluded but for much of the analyses performed within this chapter were analysed separately to those with only one diagnosis. Each patient underwent lung function tests with gas transfer measurements, total lung volumes and spirometry; an echocardiogram; and venous blood sampling including renal and liver biochemistry, full blood count and B-type natriuretic peptide measurement (BNP).

6.3.2 CPX analysis

Each patient recruited undertook 2 bicycle ergometer CPX tests performed on a COSMED Quark CPET System (COSMED S.r.l. Rome, Italy). The first test, which all patients undertook on an identical 10W/minute ramp protocol, was for familiarisation purposes and was ignored in this primary analysis. The second test was personalised for each patient based on the length of time cycled on test 1 to achieve an incremental cycle time of around 10 minutes. Analysis was conducted on this second test. The identification and measurement of CPX variables was performed as per directions from within the Methods Chapter and will not be discussed again here. However following on from the identification of a new set of contemporary reference equations (Gläser et al 2010, Chapter 4) percent of predicted values for peak $\dot{V}_{O_2}$, OUES and $O_2$ pulse will be calculated using these equations shown below (Table 6.1). The $\dot{V}_{O_2}$ at the AT will be presented as a percent of the predicted peak $\dot{V}_{O_2}$ using the SHIP reference equation for peak $\dot{V}_{O_2}$. Circulatory power was calculated as the product of peak systolic blood pressure and peak $\dot{V}_{O_2}$. Whilst this data has also generated reference ranges for the $\dot{V}_E/\dot{V}_{CO_2}$ relationship, $\dot{V}_{O_2} - WR$ slope and the $HR - \dot{V}_{O_2}$ slope these variables do not behave in a typically linear manner, the magnitude of a percent of predicted value is not easily interpretable, and therefore these variables will just be displayed giving their raw values. $\dot{V}_E/\dot{V}_{CO_2}$ slope is displayed in 2 forms: slope 1 using data until the VCP; and slope 2 using data throughout exercise, in keeping with previous chapters. It was noticed that the $\dot{V}_E/\dot{V}_{CO_2}$ slope and ratio behaved differently between the groups, so I therefore hypothesised that the difference
between these 2 values may have some discriminatory power. This was calculated as $\dot{V}_E/\dot{V}_{CO_2}$ slope (data to VCP only) minus the $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir.

### Reference Equations for CPX variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sex</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $\dot{V}_{O_2}$</td>
<td>M</td>
<td>$-69 + 1.48\times\text{Age} + 14.02\times\text{Height} + 7.44\times\text{Weight} - 233.72\times\text{cs} - 0.2256\times\text{Age}^2$</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>$-588 - 11.33\times\text{Age} + 9.13\times\text{Height} + 26.88\times\text{Weight} - 0.12\times\text{Weight}^2$</td>
</tr>
<tr>
<td>$O_2$ pulse</td>
<td>M</td>
<td>$-0.7 - 0.044\times\text{Age} + 0.064\times\text{Height} + 0.086\times\text{Weight} - 0.62\times\text{cs} + 1.73\times\text{bb}$</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>$-3.7 - 0.004\times\text{Age} + 0.056\times\text{Height} + 0.075\times\text{Weight} + 0.42\times\text{bb}$</td>
</tr>
<tr>
<td>OUES</td>
<td>M</td>
<td>$+907.7 - 11.51\times\text{Age} + 5.67\times\text{Height} + 8.62\times\text{Weight} - 49.99\times\text{bb} - 214.53\times\text{cs} + 172.97\times f$</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>$-182.4 - 8.89\times\text{Age} + 10.12\times\text{Height} + 10.51\times\text{Weight} - 117.65\times\text{bb} - 21.45\times\text{cs} + 40.31\times f$</td>
</tr>
</tbody>
</table>

Table 6.1: Reference ranges for peak $\dot{V}_{O_2}$ (mL/min), $O_2$ pulse (mL/beat) and OUES (mL/min/10-fold increase in $\dot{V}_E$) for males (M) and females (F) based on the SHIP data. Age is measured in years; height in cm; weight in kg; cs (current smoker) = 0 for no, 1 for yes; bb (beta-blocker) = 0 for no, 1 for yes; f = FEV$_1$ (L).

### 6.3.3 Hypotheses

I hypothesised that peak $\dot{V}_{O_2}$ would show no ability to discriminate between patients with cardiac or respiratory disease. Markers of lung function and cardiac function would influence peak $\dot{V}_{O_2}$. Conversely variables including the OUES, the OUEP, the $O_2$ pulse, $\dot{V}_{O_2} - WR$ slope and the $\dot{V}_{O_2}$ at the AT, would all successfully discriminate these 2 disease categories, with high sensitivity and specificity for cardiac disease, and would therefore not be affected by markers of lung function.

### 6.3.4 Statistical analysis

Statistical analysis was performed using Stata version 11.1 for Windows (StataCorp LP, College Station, Texas). All patient characteristics and CPX variables were assessed for normal distribution using Shapiro-Wilk test. Because of multiple variables, and no a priori hypotheses regarding which variables may behave non-normally, a Sidak correction was employed. For those variables showing a non-normal distribution only if logarithmic transformation significantly normalised distribution was this transformation used. For all other non-
normal distributions the raw value was used in non-parametric analyses when possible. All values within tables are displayed as mean ± standard deviation for normally distributed data, and median (25th, 75th percentiles-the interquartile range) for non-normally distributed data.

To analyse for differences between groups, oneway analysis of variance (ANOVA) was used, with addition of the co-variates age, gender and weight. For between group differences in CPX variables three other co-variates were added to the model, beta-blocker use, eGFR and current smoking status, for which ex-smokers were classified with non-smokers. This distinction was made because cigarette smoke (principally carbon monoxide) can limit oxygen transport in haemoglobin; this is the effect I wish to analyse for as a confounder, and would therefore be inappropriate in ex-smokers. Although patients were encouraged to refrain from smoking prior to testing, carbon monoxide can remain in the blood for 48 hours, and we could not mandate smoking cessation for this length of time to all patients. For categorical variables the χ² test was performed to assess for between group differences. For non-normally distributed variables the Kruskall-Wallis test was used in place of ANOVA.

When comparing 2 continuous variables linear regression was performed. To assess for the influence of respiratory function on CPX variables a multivariate model including age, gender and weight was constructed.

The primary analysis of this study is the ability of each variable to discriminate between COPD and cardiac disease. This was assessed using Receiver Operator Characteristic (ROC) curve analysis with area under curve (AUC) measurements. Comparison of the AUC values from key variables were analysed for significant differences to identify optimal variables. The groups tested against each other were: 1) all cardiac patients against COPD patients (excluding mixed disease); 2) CHF patients against COPD patients (excluding mixed disease); 3) CHF patients against COPD with patients identified by their primary diagnosis. The ROC model was constructed so that the AUC was always >0.50, however this does not identify direction of the discrimination i.e. whether higher or lower values predict cardiac disease or COPD. Therefore for some variables the AUC reflects ability to discriminate cardiac disease, and for others the ability to discriminate respiratory disease. However because the purpose of the study is not to identify a specific threshold for detection, merely to identify a variable which would appear to be relatively sensitive and specific for cardiac disease, this distinction is not relevant.

A p value of <0.05 was considered significant throughout (with the exception of the Shapiro-Wilk test as described above).
6.4 Results

6.4.1 Patient recruitment and characteristics

257 patients were approached following identification as potentially appropriate for the study by members of their clinical team. 135 patients were either clearly inappropriate, declined, could not be further contacted, or met one or more of the exclusion criteria after detailed discussion. 122 patients agreed to be part of the study and appeared appropriate. Of these 122 a further 9 did not turn up to the original and repeated appointments; 7 were either too unwell, had/were experiencing a recent decompensation or could not cycle; 1 patient had (unknown to him) being experiencing significant ventricular arrhythmias confirmed on CRT interrogation; 1 had an excessive BP response soon after testing began requiring immediate cessation of testing; 1 had completely normal spirometry (from the COPD group) so was not pursued, and a final patient could not tolerate the CPX mask.

This left 102 patients that underwent the study protocol. 2 patients recruited near the beginning of the study had moderate to severe, asymptomatic, aortic stenosis, which was then dropped as an inclusion criteria. 2 patients with COPD were excluded because their spirometry failed to show respiratory limitation at rest significant enough for inclusion (although it was still abnormal). 1 patient with CRT was excluded due to recovery in left ventricular function following CRT (aetiology of left ventricular systolic dysfunction was alcoholic cardiomyopathy and the patient was now abstinent) and 1 patient with mitral valve disease was excluded because his MR was later deemed not of sufficient severity to require reparative surgery.

3 patients did not have data from both a familiarisation and main test, but these 3 were still included in the final analysis (two patients underwent the full study protocol but did not have reliable data in the familiarisation test, and one patient did not wish to perform a second test or echocardiogram – his familiarisation test is used for the main analysis).

Of the final 96 patients (72 male), 43 were patients with CHF (14 undergoing CRT and 13 having undergone CRT), 26 with mitral valve disease and 27 with lung disease (25 showed obstructive physiology and 2 restrictive physiology but had a formal physician-made diagnosis of COPD). The patient characteristics of the total group and by disease category condition are shown in Table 6.2. 18 from 26 patients with mitral valve disease were symptomatic. 7 patients fulfilled criteria for inclusion in the mixed group. 4 of these patients had COPD as their primary diagnosis but 3 were known to have ischaemic heart disease and on echocardiography had abnormalities of left ventricular function, and a fourth had an elevated BNP and right atrial and ventricular
dilatation. 3 heart failure patients were known to have COPD. 5 further heart failure patients and 5 mitral valve patients had obstructive spirometry which met GOLD criteria category 2 but not higher. 2 were known to have long-standing asthma (not COPD) but none of the remaining patients had a formal secondary diagnosis or symptoms suggestive of COPD, and only 1 was a current smoker. Of these 10 patients 4 were reclassified with non-obstructive or mildly obstructive spirometry (GOLD category 1) on repeat spirometry, 2 had FEV\textsubscript{1} values just below the 80% cut-off to distinguish between category 1 and 2, and 4 lay well within category 2. They were categorised within their principal disease category. Within the COPD category, 2 patients met the criteria for GOLD category 1 (mild), 12 patients met criteria for GOLD category 2 (moderate), 10 for category 3 (severe) and 3 for category 4 (very severe).

Patient variables were assessed for a normal distribution. FVC and the FEV\textsubscript{1}:FVC ratio were the only major respiratory variables non-normally distributed. Echocardiographic variables were largely non-normally distributed, however they showed a bivariate distribution (which would be expected) and within each disease category the distribution was normal by Shapiro-Wilk testing; the FVC and FEV\textsubscript{1}:FVC ratio largely showed this pattern as well. LA volume was non-normally distributed throughout each group. BNP was non-normally distributed throughout the whole group and the individual groups. Therefore where necessary non-parametric tests were used but BNP and LA volume were the only variables logarithmically transformed.

There were significantly more males in the CHF group, which also had a significantly greater BMI than the other 2 groups. Haemoglobin (Hb) and estimated glomerular filtration rate (eGFR) were significantly lower in the CHF group, whilst BNP was significantly higher in the 2 cardiac groups compared to the COPD group. FVC was significantly lower in the COPD group compared to the others, but only significantly lower than the CHF group as a percent of predicted. FEV\textsubscript{1} and percent of predicted FEV\textsubscript{1}, along with the FEV\textsubscript{1}:FVC ratio were significantly lower in the COPD group compared to both cardiac groups. All cardiac indices displayed in Table 6.2 were significantly different between the CHF and COPD groups, and the CHF and mitral groups. LV diastolic volume (LVEDV) and LA volume were significantly larger in the mitral compared to the COPD group, but systolic volume (LVESV), ejection fraction (LVEF), fractional shortening (LVFS), RV S’ wave and TAPSE were similar. A significantly greater proportion of patients within the CHF group (34/40 patients), and a significantly smaller proportion within the COPD group (1/23 patients) regularly took beta-blockers. The use of beta-blockers did not significantly affect the results when added to any model of baseline characteristics.
<table>
<thead>
<tr>
<th></th>
<th>p between groups</th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
<th>Mean ± SD/ Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All</td>
<td>COPD</td>
<td>CHF</td>
<td>MV</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>0.32</td>
<td>65.0 ± 10.9</td>
<td>65.8 ± 9.8</td>
<td>66.6 ± 11.1</td>
<td>61.7 ± 11.3</td>
</tr>
<tr>
<td>Male (n, %)</td>
<td>0.02</td>
<td>72, 75%</td>
<td>18, 66.7%</td>
<td>38, 88.4%</td>
<td>16, 61.5%</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>0.08</td>
<td>169.9 ± 9.9</td>
<td>166.5 ± 7.5</td>
<td>171.3 ± 9.2</td>
<td>171.2 ± 12.3</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>0.03</td>
<td>77.3 ± 16.0</td>
<td>71.3 ± 16.4</td>
<td>82.7 ± 15.5</td>
<td>74.4 ± 13.8</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>0.01</td>
<td>26.6 ± 4.5</td>
<td>25.6 ± 4.9</td>
<td>28.1 ± 4.1</td>
<td>25.3 ± 4.0</td>
</tr>
<tr>
<td>Diabetics (n, %)</td>
<td>0.06</td>
<td>10, 10.4%</td>
<td>1, 3.7%</td>
<td>8, 18.6%</td>
<td>1, 3.8%</td>
</tr>
<tr>
<td>Current smokers (n, %)</td>
<td>0.09</td>
<td>12, 12.5%</td>
<td>5, 18.5%</td>
<td>7, 16.3%</td>
<td>0, 0%</td>
</tr>
<tr>
<td>Hypertensive (n, %)</td>
<td>0.17</td>
<td>42, 43.8%</td>
<td>12, 44.4%</td>
<td>19, 44.2%</td>
<td>11, 42.3%</td>
</tr>
<tr>
<td>Hb (g/dL)</td>
<td>0.02</td>
<td>14.0 ± 1.24</td>
<td>14.4 ± 1.17</td>
<td>13.8 ± 1.37</td>
<td>14.0 ± 1.03</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73m²)</td>
<td>&lt;0.01</td>
<td>77.8 ± 19.1</td>
<td>84.7 ± 19.6</td>
<td>70.6 ± 18.2</td>
<td>82.8 ± 16.4</td>
</tr>
<tr>
<td>BNP (pg/mL)</td>
<td>&lt;0.001</td>
<td>85 (35, 195)</td>
<td>32.5 (13, 50)</td>
<td>118 (85, 286)</td>
<td>85 (43, 222)</td>
</tr>
<tr>
<td>FVC (L)</td>
<td>&lt;0.01</td>
<td>3.26 (2.57, 3.91)</td>
<td>2.75 (2.29, 3.06)</td>
<td>3.59 (3.01, 4.31)</td>
<td>3.29 (2.75, 3.88)</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>0.03</td>
<td>96.6 ± 22.9</td>
<td>87.3 ± 25.1</td>
<td>100.1 ± 18.0</td>
<td>100.5 ± 25.8</td>
</tr>
<tr>
<td>FEV₁ (L)</td>
<td>&lt;0.001</td>
<td>2.17 ± 0.90</td>
<td>1.28 ± 0.49</td>
<td>2.52 ± 0.73</td>
<td>2.50 ± 0.87</td>
</tr>
<tr>
<td>FEV₁ (% predicted)</td>
<td>&lt;0.001</td>
<td>78.2 ± 26.5</td>
<td>50.8 ± 19.6</td>
<td>89.7 ± 20.2</td>
<td>87.7 ± 21.4</td>
</tr>
<tr>
<td>FEV₁:FVC ratio (%)</td>
<td>&lt;0.001</td>
<td>66.9 (52.8, 74)</td>
<td>43.3 (35.8, 55.9)</td>
<td>71.1 (62.7, 77)</td>
<td>68.6 (65.1, 75.1)</td>
</tr>
<tr>
<td>KCO (min⁻¹)</td>
<td>0.002</td>
<td>1.16 ± 0.35</td>
<td>0.96 ± 0.40</td>
<td>1.20 ± 0.32</td>
<td>1.30 ± 0.29</td>
</tr>
<tr>
<td>Ramp protocol (W/min)</td>
<td>0.05</td>
<td>11.3 ± 3.3</td>
<td>10.3 ± 2.8</td>
<td>11.3 ± 2.9</td>
<td>12.3 ± 4.0</td>
</tr>
<tr>
<td>LVEDV (cm³)</td>
<td>&lt;0.001</td>
<td>1245 (89, 165)</td>
<td>72 (55, 92)</td>
<td>149 (112, 193)</td>
<td>125 (106, 147)</td>
</tr>
<tr>
<td>LVESV (cm³)</td>
<td>&lt;0.001</td>
<td>56.5 (37.2, 99.8)</td>
<td>28.9 (19.7, 37.2)</td>
<td>99.9 (75.9, 145.3)</td>
<td>49 (38.5, 54.2)</td>
</tr>
<tr>
<td>LVEF (%)</td>
<td>&lt;0.001</td>
<td>51.3 (33.8, 59.5)</td>
<td>59.5 (53.0, 64.5)</td>
<td>33.5 (30.8, 39.1)</td>
<td>58.4 (54.9, 65.7)</td>
</tr>
<tr>
<td>LVFS (%)</td>
<td>&lt;0.001</td>
<td>23.2 (16.0, 29.8)</td>
<td>27.5 (23.3, 31.9)</td>
<td>15.2 (10.9, 18.8)</td>
<td>30.4 (25.6, 34.6)</td>
</tr>
<tr>
<td>LA volume (cm³)</td>
<td>&lt;0.001</td>
<td>74.3 (52.2, 116.5)</td>
<td>43.0 (33.3, 63.8)</td>
<td>74.2 (56.4, 103.3)</td>
<td>135 (101, 170)</td>
</tr>
<tr>
<td>TAPSE (mm)</td>
<td>0.001</td>
<td>21 ± 5</td>
<td>22 ± 4</td>
<td>19 ± 4</td>
<td>23 ± 5</td>
</tr>
<tr>
<td>RV S’ wave (cm/s)</td>
<td>&lt;0.001</td>
<td>11.4 (9.4, 14)</td>
<td>12.5 (10.7, 15.6)</td>
<td>9.6 (8.7, 11.8)</td>
<td>12.8 (10.6, 16.6)</td>
</tr>
</tbody>
</table>

Table 6.2: Patient characteristics. p values between groups by ANOVA with age and gender as co-variates for continuous variables and χ² analysis for categorical variables. Normally distributed data is presented as mean ± SD, with median (IQR) for non-normally distributed data. For abbreviations see text.
6.4.2 CPX results and influencing factors on exercise capacity

Table 6.3 gives the mean values for the CPX variables overall and within each group. 8 patients in the COPD group, 1 in the CHF group and 1 in the mixed group did not reach anaerobic threshold, the variables recorded at the AT reflect only the patients reliably achieving the AT. On analysis of CPX variables, OUES/kg both measures of the \( \dot{V_E}/\dot{V_{CO2}} \) slope, all measures of the \( \dot{V_E}/\dot{V_{CO2}} \) ratio, resting heart rate, arterial oxygen saturations, breathing reserve at the AT and peak, and the \( HR - \dot{V_O2} \) slope all showed a significantly non-normal distribution (p<0.0026 by Shapiro-Wilk test). All measures of the \( \dot{V_E}/\dot{V_{CO2}} \) relationship were made marginally more normally distributed by logarithmic transformation, but resting heart rate was the only variable which became significantly more normally distributed following logarithmic transformation and therefore transformed in this way for this analysis.

Peak \( \dot{V_{O2}} \) (as the considered overall marker of exercise capacity) was assessed for the influence of certain patient factors. Creatinine and eGFR was related to peak \( \dot{V_{O2}}/kg \) (\( R^2=0.05, p=0.03 \) and \( R^2=0.08, p<0.01 \)). Haemoglobin, sodium and bicarbonate did not relate to peak \( \dot{V_{O2}} \) as an absolute or weight adjusted value. BNP was related to both measures of peak \( \dot{V_{O2}} \) (\( R^2=0.12, p=0.002 \) and \( R^2=0.13, p=0.001 \) respectively). Atrial fibrillation compared with sinus rhythm did not significantly reduce peak \( \dot{V_{O2}} \) (p=0.90). No other rhythm abnormalities were found in the Observational study.

6.4.3 The influence of the lungs on CPX variables

As shown in the pilot data in Chapter 3 in a retrospective analysis of patients with heart failure, abnormal spirometry related to the peak \( \dot{V_{O2}} \), loosely to the AT, and not to the \( \dot{V_E}/\dot{V_{CO2}} \) slope or OUES. The influence of FEV\textsubscript{1} and percent of predicted FEV\textsubscript{1} as continuous variables on CPX variables were therefore assessed using a regression model including age, gender, weight and eGFR (as these variables were significantly different between groups and affected peak \( \dot{V_{O2}} \)) as co-variates on this prospective dataset. I also assessed for a relation between these CPX variables and the diffusional capacity of the lungs as measured by the transfer coefficient corrected for Hb (K\textsubscript{CO}(Hb)). Because this is a multivariate model I have not shown the \( R^2 \) values, as these would pertain to the whole model rather than just the influence of FEV\textsubscript{1} on the CPX variable.

FEV\textsubscript{1} and percent of predicted FEV\textsubscript{1} related to peak \( \dot{V_{O2}} \) (p<0.02 for both), peak \( \dot{V_{O2}}/kg \) (p<0.005 for both), percent of predicted peak \( \dot{V_{O2}} \) (p<0.02), OUES (p<0.01), OUES/kg (p<0.005), percent predicted OUES
(p<0.001), OUEP (p<0.01), $\dot{V}_E/\dot{V}_{CO_2}$ slope measured throughout exercise (but not when measured until the VCP) (p<0.05) breathing reserve both at the AT (p<0.001) and at peak (p<0.001). The difference between the $\dot{V}_E/\dot{V}_{CO_2}$ slope and $\dot{V}_E/\dot{V}_{CO_2}$ ratio related to the percent of predicted FEV$_1$ only (p=0.02). It was interesting to see a relation between both $\dot{V}_E/\dot{V}_{CO_2}$ slope and OUES and FEV$_1$; however these significant relationships were in the opposite direction to that expected, i.e. a rise in FEV$_1$ related to lower OUES and higher $\dot{V}_E/\dot{V}_{CO_2}$ slope measures. For all measures of peak $\dot{V}_{O_2}$, the OUEP and breathing reserve a fall in FEV$_1$ led to a fall in these values.

Peak $\dot{V}_{O_2}$ related to the K$_{CO}$ (Hb) (p<0.001), as did peak $\dot{V}_{O_2}/kg$ (p<0.001), percent predicted peak $\dot{V}_{O_2}$ (p=0.001), $\dot{V}_{O_2}$ at the AT (p=0.03) (but not as a percent of predicted peak $\dot{V}_{O_2}$), OUEP (p<0.001), O$_2$ pulse (p=0.01) and percent predicted O$_2$ pulse (p=0.01), $\dot{V}_E/\dot{V}_{CO_2}$ slope until the VCP, (p=0.01) (but not measured throughout exercise), $\dot{V}_E/\dot{V}_{CO_2}$ ratio at the nadir (p<0.001) and AT (p<0.01), but not VCP, $P_{ET}CO_2$ at the AT (p=0.003), circulatory power (p=0.006), and breathing reserve at peak (p=0.01) (but not at AT). For all of these variables, including the measures of $\dot{V}_E/\dot{V}_{CO_2}$, a reduction in K$_{CO}$ related to a deterioration in the CPX variable (so therefore a rise in the measures of $\dot{V}_E/\dot{V}_{CO_2}$).

### 6.4.4 Differences in variables between groups

ANOVA (with age, gender and weight as co-variates) was performed to look for between group differences ignoring the 7 patients included in the mixed group. Figure 6.2 shows how each patient within each group and the within group means for the 3 diagnostic categories. $\dot{V}_E/\dot{V}_{CO_2}$ ratio at the VCP was not investigated because patients with COPD rarely reached this threshold.

No measure of the peak $\dot{V}_{O_2}$ or the $\dot{V}_{O_2}$ at the AT was significantly different between the 3 groups (p>0.17 in all). OUES as an absolute value (p<0.001), adjusted for weight (p<0.001) and as percent predicted (p<0.001) was significantly different between groups with significantly lower values in the 2 cardiac groups compared to the COPD groups (but not different between the 2 cardiac groups). Foreshortened OUES measures were inferior to full data OUES (not shown in tables). OUEP and O$_2$ pulse showed a borderline significant difference between the groups (p=0.02 and p=0.03 respectively). O$_2$ pulse as a percent of predicted was not different between the groups. No measure of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship was different between groups, except for the difference between 202
the slope and ratio (p=0.003) which was significantly lower in the cardiac groups. Heart rates at rest and peak exercise were significantly different between groups (p<0.001 both) with lower resting heart rates in the 2 cardiac groups, significantly lower heart rate at peak in the CHF group compared to both, and significantly higher peak heart rate in the MV group. Both double product and peak circulatory power were significantly lower in the CHF group (p<0.001 both between groups) than the MV or COPD groups. Breathing reserve at the AT and at peak exercise, and the peak minute ventilation were also all different between groups (p<0.005 for all), with similar results for the 2 cardiac groups, and significantly lower values in the COPD group. The \( \dot{V}_{\text{O}_2} - WR \) slope was significantly lower in the CHF and MV groups (p=0.02), the \( HR - \dot{V}_{\text{O}_2} \) slope was significantly higher in the MV group (p=0.01) and the \( HR - \dot{V}_{\text{O}_2} \) intercept was significantly lower in both cardiac groups (0.002). Because measures of the \( \dot{V}_{\text{E}}/\dot{V}_{\text{CO}_2} \) slope, breathing reserve, peak \( \dot{V}_{\text{E}}, \dot{V}_{\text{O}_2} - WR \) slope and \( HR - \dot{V}_{\text{O}_2} \) slope were non-normally distributed they were also assessed using the non-parametric Kruskall-Wallis test (no co-variates), with results the same as for ANOVA. The addition of the use of beta-blockers, eGFR and current smoking status to the model largely didn’t change the results with the exception of the O\textsubscript{2} pulse which was no longer significantly different between groups after adjustment for beta-blocker use, and DP and circulatory power which were not significantly different between groups after adjustment for beta-blockers and smoking.Whilst beta-blocker use was related to peak heart rate, it did not affect the between categories difference.

The absence of a reliably measurable AT was found in 7 patients with COPD (30.4%) and 1 patient with CHF (2.5%), and this may have affected the ability of the ANOVA to differentiate between groups; therefore for these 8 patients the peak \( \dot{V}_{\text{O}_2} \) was also recorded as the \( \dot{V}_{\text{O}_2} \) at the AT for a secondary analysis. This did not alter the results seen when analysing only patients who did achieve AT.
<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>COPD (n=23)</th>
<th>CHF (n=40)</th>
<th>MV (n=26)</th>
<th>Mixed (n=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $V_{O_2}$ (mL/min)</td>
<td>1311.6 ± 376.8</td>
<td>1190.4 ± 316.5</td>
<td>1348.9 ± 348.0</td>
<td>1390.3 ± 460.6</td>
<td>1205.4 ± 322.6</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (mL/min/kg)</td>
<td>17.1 ± 4.3</td>
<td>17.3 ± 4.7</td>
<td>16.4 ± 3.6</td>
<td>18.6 ± 5.0</td>
<td>15.3 ± 2.8</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (% predicted)</td>
<td>71.5 ± 16.2</td>
<td>71.2 ± 18.2</td>
<td>70.2 ± 14.9</td>
<td>73.7 ± 16.7</td>
<td>68.8 ± 9.0</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>964.2 ± 248.3</td>
<td>885.3 ± 182.1</td>
<td>975.9 ± 212.5</td>
<td>979.5 ± 308.0</td>
<td>1019.0 ± 336.1</td>
</tr>
<tr>
<td>AT (% of pred peak $V_{O_2}$)</td>
<td>52.2 ± 12.5</td>
<td>53.0 ± 11.0</td>
<td>51.0 ± 13.3</td>
<td>52.6 ± 12.5</td>
<td>57.0 ± 11.1</td>
</tr>
<tr>
<td>OUES *</td>
<td>1.67 ± 0.58</td>
<td>1.91 ± 0.56</td>
<td>1.57 ± 0.47</td>
<td>1.55 ± 0.65</td>
<td>1.87 ± 0.78</td>
</tr>
<tr>
<td>OUES/kg *</td>
<td>20.9 (18.0, 25.3)</td>
<td>25.6 (21.4, 30.7)</td>
<td>19.4 (16.2, 22.7)</td>
<td>20.2 (17.2, 24.7)</td>
<td>21.2 (19.3, 32.1)</td>
</tr>
<tr>
<td>OUEP *</td>
<td>33.1 ± 5.2</td>
<td>30.8 ± 5.5</td>
<td>33.3 ± 5.1</td>
<td>35.0 ± 4.2</td>
<td>32.7 ± 6.4</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat) *</td>
<td>11.1 ± 3.1</td>
<td>9.4 ± 2.0</td>
<td>12.5 ± 3.0</td>
<td>10.1 ± 3.2</td>
<td>13.4 ± 2.6</td>
</tr>
<tr>
<td>$\dot{V}_E/\dot{V}_CO_2$ slope 1</td>
<td>34.3 (29.5, 38.1)</td>
<td>32.4 (27.2, 40.1)</td>
<td>34.8 (30.8, 38.1)</td>
<td>33.8 (29.1, 36.4)</td>
<td>34.7 (25.6, 39.8)</td>
</tr>
<tr>
<td>$\dot{V}_E/\dot{V}_CO_2$ slope 2</td>
<td>35.4 (31.2, 40.0)</td>
<td>32.4 (29.2, 40.1)</td>
<td>36.4 (33.2, 39.0)</td>
<td>36.2 (33.0, 41.6)</td>
<td>34.7 (25.6, 39.8)</td>
</tr>
<tr>
<td>$\dot{V}_E/\dot{V}_CO_2$ ratio nadir</td>
<td>32.5 (29.9, 38.1)</td>
<td>33.0 (29.7, 39.3)</td>
<td>32.8 (29.8, 36.6)</td>
<td>31.7 (30.2, 35.6)</td>
<td>35.9 (29.5, 41.7)</td>
</tr>
<tr>
<td>$\dot{V}_E/\dot{V}_CO_2$ ratio at AT</td>
<td>34.5 (31.5, 38.4)</td>
<td>35.8 (30.9, 40.1)</td>
<td>34.5 (31.5, 39.0)</td>
<td>32.8 (31.7, 37.1)</td>
<td>35.8 (30.1, 37.4)</td>
</tr>
<tr>
<td>RER at peak</td>
<td>1.08 ± 0.12</td>
<td>1.00 ± 0.13</td>
<td>1.11 ± 0.10</td>
<td>1.14 ± 0.11</td>
<td>1.00 ± 0.05</td>
</tr>
<tr>
<td>$P_{2}CO_2$ at AT (mmHg)</td>
<td>35.4 ± 4.3</td>
<td>35.7 ± 4.5</td>
<td>35.2 ± 4.2</td>
<td>35.5 ± 3.9</td>
<td>36.3 ± 6.8</td>
</tr>
<tr>
<td>HR at peak (bpm) *</td>
<td>124 ± 26</td>
<td>128 ± 18</td>
<td>114 ± 22</td>
<td>143 ± 26</td>
<td>98 ± 18</td>
</tr>
<tr>
<td>DP (mmHg bpm) *</td>
<td>20356 ± 6552</td>
<td>23740 ± 5764</td>
<td>17566 ± 5029</td>
<td>23538 ± 6795</td>
<td>13839 ± 3485</td>
</tr>
<tr>
<td>Circ power (mmHg mL/min) *</td>
<td>216355 ± 84191</td>
<td>221155 ± 77057</td>
<td>208376 ± 68638</td>
<td>236728 ± 111159</td>
<td>171200 ± 58918</td>
</tr>
<tr>
<td>Peak oxygen saturations (%)</td>
<td>94 (95, 98)</td>
<td>94 (91, 98)</td>
<td>98 (97, 99)</td>
<td>97 (96, 98)</td>
<td>94 (91, 96)</td>
</tr>
<tr>
<td>BR at AT (%) *</td>
<td>68.8 (54.3, 76.3)</td>
<td>48.5 (43.6, 66.0)</td>
<td>72.3 (63.3, 77.4)</td>
<td>71.1 (63.3, 78.4)</td>
<td>38.0 (24.8, 59.3)</td>
</tr>
<tr>
<td>BR (%) *</td>
<td>34.8 (16.8, 45.2)</td>
<td>11.1 (0.7, 28.0)</td>
<td>44.1 (34.7, 52.7)</td>
<td>39.8 (26.2, 45.3)</td>
<td>12.4 (-7.6, 38.1)</td>
</tr>
<tr>
<td>$V_{O_2} - WR$ Slope *</td>
<td>0.044 (0.03, 0.06)</td>
<td>0.046 (0.04, 0.06)</td>
<td>0.038 (0.03, 0.05)</td>
<td>0.057 (0.04, 0.08)</td>
<td>0.036 (0.03, 0.04)</td>
</tr>
<tr>
<td>$HR - V_{O_2}$Slope *</td>
<td>11.03 ± 1:54</td>
<td>10:11 ± 2:43</td>
<td>11:22 ± 1:27</td>
<td>11:26 ± 1:38</td>
<td>10:36 ± 1:07</td>
</tr>
</tbody>
</table>

Table 6.3: CPX results for a number of principal variables for the whole cohort and then divided into disease categories; chronic obstructive pulmonary disease (COPD), chronic heart failure (CHF), mitral valve disease (MV) and mixed cardiac and respiratory abnormalities. Results are displayed as mean ± SD for normally distributed variables and median (25th, 75th percentiles) for non-normally distributed variables. * p<0.05 for difference between first 3 groups (see text for differences between which groups)
Figure 6.2: Dotplots for 9 CPX variables in patients with COPD, CHF, mitral valve disease (MV) and mixed disease. Each dot represents a single patient and mean values are represented by black lines.
6.4.5 ROC curve analysis

Received Operator Characteristics (ROC) curve analysis was conducted on all the major variables to identify which had the ability to discriminate between a cardiac and respiratory diagnosis. The area under curve (AUC) for each variable was calculated for all cardiac patients against COPD patients (excluding mixed disease), CHF patients against COPD patients (excluding mixed disease) and CHF patients against COPD with patients identified by their primary diagnosis (Table 6.4). The AUC for each of the divisions of OUES (25-75, 0-50, 0-70 and 0-90) were also calculated but, because they were all inferior to OUES using the full data, are not included in the table. The ROC model was constructed so that the AUC was always >0.50, however this does not identify direction of the discrimination i.e. whether higher or lower values predict cardiac disease or COPD. There were not sufficient measurements of the $\bar{V}_{E}/\bar{V}_{CO2}$ slope at the VCP in the COPD group to construct an ROC curve.

The close nature of the AUC for percent predicted peak $\bar{V}_{O2}$ to 0.500 confirms that these groups are largely well matched in terms of exercise capacity.

Variables with good AUC values to discriminate between COPD and cardiac disease included breathing reserve (AUC 0.899), breathing reserve at the AT (AUC 0.861), OUES/kg (AUC 0.822) and percent of predicted OUES (AUC 0.864). None of these values were significantly different from one another. For these variables the ability to discriminate improves marginally if only CHF patients were compared with COPD, and drops marginally if patients with mixed disease were also included under their primary diagnosis. The $\bar{V}_{O2} - WR$ slope had an intermediate AUC, but this was significantly worse than percent of predicted OUES. Measures of the $\bar{V}_{E}/\bar{V}_{CO2}$ relationship did not reliably discriminate except for the difference between the slope and ratio at nadir. This value had an AUC of 0.761 between the 2 groups with positive measures (corresponding to ratios greater than slopes) predicting COPD, and negative values (corresponding to ratios less than slopes) predicting cardiac disease. The $\bar{V}_{O2}$ at the AT as an absolute value and as a percentage of predicted peak $\bar{V}_{O2}$ were not strong discriminators.

The $O_2$ pulse showed a reasonable ability to discriminate when comparing only CHF with COPD; however the highest values were seen in the CHF group and the differences disappear when displayed as a percentage of predicted. Double product was useful to discriminate between CHF and COPD but when mitral valve patients are introduced it significantly worsened. Reason for test cessation compared between fatigue/leg fatigue (n=50)
and breathlessness (n=22), ignoring patients stopping for both or alternative reasons (n=24), was not a strong discriminator (AUC 0.6691).

A secondary analysis was performed re-classifying the 3 CHF and 1 MV patient with obstructive spirometry (but not known COPD) into the mixed group. All AUC values were within 2% (0.02) for the CHF versus COPD and cardiac versus COPD analyses. ROC curves for some of the key variables are shown in Figure 6.3.

Accepting OUES and breathing reserve as the 2 best discriminators these variables were examined more closely. OUES had a good ability to discriminate between CHF and COPD (AUC 0.8772; 95% confidence intervals 0.791, 0.963) when measured as a percentage of predicted. As a weight adjusted model the AUC was similar, but as an absolute value it was significantly worse, which may reflect the significant demographic differences between the groups. Lower values were predictive of CHF over COPD. Although the hypothesis of this study is not to generate a threshold, a percentage of predicted OUES of 89.3% gave a sensitivity of 78.3% with a specificity of 85% to identify CHF and correctly categorised 82.5% of patients (Figure 6.4).

Breathing reserve had a slightly better, albeit not significant, ability to discriminate (AUC 0.9120; 95% confidence intervals 0.844, 0.980). In this case the lowest values were suggestive of COPD not CHF. A BR threshold of 33.6% gave a sensitivity of 80% with a specificity of 100% to identify COPD and correctly categorised 87.3% of patients (Figure 6.4).

I looked to see if the combination of RER and BR at peak improved discrimination. The product of these 2 variables showed a non-significant improvement in AUC for discriminating CHF from COPD (AUC=0.9283; 95% confidence intervals 0.867, 0.989) and cardiac disease from COPD (AUC = 0.9177; 95% confidence intervals 0.862, 0.972).

When patients were categorised based on the recommended algorithm (Figure 6.1) the results were as follows: 15 from 23 with COPD were categorised as respiratory limitation, none were categorised as circulatory limitation; only 6 from 40 patients with CHF, and 2 from 26 with mitral valve disease, were categorised as circulatory limitation, the majority being categorised as normal, or poor effort/deconditioning; and none of the 7 patients with significant cardiac and respiratory disease were categorised as mixed lesions.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Cardiac vs COPD (excluding mixed)</th>
<th>CHF vs COPD (excluding mixed)</th>
<th>Cardiac vs COPD (by primary diagnosis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $V_O_2$ (mL/min)</td>
<td>0.6364</td>
<td>0.6370</td>
<td>0.6141</td>
</tr>
<tr>
<td>Peak $V_O_2$ (mL/min/kg)</td>
<td>0.5053</td>
<td>0.5554</td>
<td>0.5105</td>
</tr>
<tr>
<td>Peak $V_O_2$ (% predicted)</td>
<td>0.5204</td>
<td>0.5087</td>
<td>0.5030</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>0.6021</td>
<td>0.6239</td>
<td>0.5605</td>
</tr>
<tr>
<td>AT (% of pred peak $V_O_2$)</td>
<td>0.5405</td>
<td>0.5692</td>
<td>0.5784</td>
</tr>
<tr>
<td>OUES</td>
<td>0.6703</td>
<td>0.6625</td>
<td>0.6637</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>0.8221</td>
<td>0.8478</td>
<td>0.8084</td>
</tr>
<tr>
<td>OUES (% predicted)</td>
<td>0.8656</td>
<td>0.8772</td>
<td>0.8642</td>
</tr>
<tr>
<td>OUEP</td>
<td>0.6686</td>
<td>0.6272</td>
<td>0.6253</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat)</td>
<td>0.6902</td>
<td>0.7941</td>
<td>0.6678</td>
</tr>
<tr>
<td>$O_2$ Pulse (% predicted)</td>
<td>0.5007</td>
<td>0.5446</td>
<td>0.5070</td>
</tr>
<tr>
<td>$V_E$/V $CO_2$ slope 1</td>
<td>0.5903</td>
<td>0.6120</td>
<td>0.6001</td>
</tr>
<tr>
<td>$V_E$/V $CO_2$ slope 2</td>
<td>0.6644</td>
<td>0.6750</td>
<td>0.6702</td>
</tr>
<tr>
<td>$V_E$/V $CO_2$ ratio nadir</td>
<td>0.5191</td>
<td>0.5043</td>
<td>0.5056</td>
</tr>
<tr>
<td>$V_E$/V $CO_2$ ratio AT</td>
<td>0.5579</td>
<td>0.5419</td>
<td>0.5033</td>
</tr>
<tr>
<td>$V_E$/V $CO_2$ slope - ratio</td>
<td>0.7609</td>
<td>0.7750</td>
<td>0.7751</td>
</tr>
<tr>
<td>RER at peak</td>
<td>0.7470</td>
<td>0.7337</td>
<td>0.7665</td>
</tr>
<tr>
<td>$P_{ET}CO_2$ at AT (mmHg)</td>
<td>0.5490</td>
<td>0.5526</td>
<td>0.6070</td>
</tr>
<tr>
<td>HR at peak (bpm)</td>
<td>0.5224</td>
<td>0.6913</td>
<td>0.5172</td>
</tr>
<tr>
<td>DP (mmHg bpm)</td>
<td>0.6784</td>
<td>0.7943</td>
<td>0.6388</td>
</tr>
<tr>
<td>Circ power (mmHg mL/min)</td>
<td>0.5213</td>
<td>0.5534</td>
<td>0.5167</td>
</tr>
<tr>
<td>Peak oxygen saturations (%)</td>
<td>0.7748</td>
<td>0.7962</td>
<td>0.7526</td>
</tr>
<tr>
<td>Breathing reserve at AT (%)</td>
<td>0.8605</td>
<td>0.8752</td>
<td>0.8668</td>
</tr>
<tr>
<td>Breathing reserve (%)</td>
<td>0.8986</td>
<td>0.9120</td>
<td>0.8991</td>
</tr>
<tr>
<td>$V_O_2$ – $WR$ Slope</td>
<td>0.7039</td>
<td>0.6761</td>
<td>0.6918</td>
</tr>
<tr>
<td>$HR$ – $V_O_2$ Slope</td>
<td>0.5321</td>
<td>0.6315</td>
<td>0.5169</td>
</tr>
</tbody>
</table>

Table 6.4: Area under curve (AUC) for a number of CPX variables following 3 ROC curve analyses: patients with only heart failure or mitral valve disease versus COPD (n=89); patients with only heart failure versus COPD (n=63); and patients with cardiac disease versus COPD by primary diagnosis (mixed diseases allowed) (n=96). The AUC does not indicate the direction of the discrimination.
Figure 6.3: Receiver Operator Characteristic (ROC) curves to differentiate between COPD and CHF. Variables shown include peak $\dot{V}O_2$ and OUES (as percent of predicted), $\dot{V}O_2$ at the AT and the $\dot{V}O_2 - WR$ slope, $\dot{V}E/\dot{V}CO_2$ slope and ratio at nadir, and the breathing reserve at the AT and peak exercise. In the top 2 graphs lower values of each variable discriminate CHF, whilst in the bottom 2 graphs lower values discriminate COPD. At multiple cut-offs the sensitivity and specificity of each cut-off to predict one disease over the other is calculated. These are then plotted against each other as 1-specificity on the x-axis, and sensitivity on the y-axis. The area under the curve defined by these points is calculated (AUC). A highly sensitive and specific variable will have an AUC close to 1. An AUC of 0.5 (the line of identity on each plot) will not discriminate.
**Figure 6.4: Diagnostic capabilities of OUES and BR cut-offs.** Using the optimal cut-offs identified for OUES (as a percent of predicted) and breathing reserve (%) from ROC curve analysis we can see how many patients would have their diagnosis of either COPD of CHF correctly identified by each variable. Numbers in red represent an incorrect diagnosis.

<table>
<thead>
<tr>
<th>OUES (% predicted)</th>
<th>COPD</th>
<th>CHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤89.3%</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>&gt;89.3%</td>
<td>18</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breathing reserve (%)</th>
<th>COPD</th>
<th>CHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤33.6%</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>&gt;33.6%</td>
<td>0</td>
<td>31</td>
</tr>
</tbody>
</table>
6.5 Discussion

Cardiopulmonary exercise test variables are generally closely related to one another, but there exist specific differences between them which may mean certain variables are more useful in certain circumstances than others. Peak $\dot{V}_{O_2}$ is the most widely used and discussed exercise variable for patients, both with respiratory and cardiac diseases. I hypothesised that this would not have any ability to discriminate between the 2 disease categories and we can see with an AUC of approximately 0.500 that the null hypothesis is proved. This is an important proof of concept to have in place for further analysis. Peak $\dot{V}_{O_2}$ is known to be affected by both categories of disease, and the lack of a significant difference between groups in ANOVA and by ROC curve analysis means that the groups are potentially well matched in exercise capacity. This allows me to analyse further variables without concern that observed differences merely relate to the underlying group differences rather than pathophysiological differences. There existed a difference between groups in gender and weight; these can be corrected for in 2 ways, reference ranges and separate analyses by groups. I chose to correct variables that are affected by these variables (for example OUES, $O_2$ pulse, peak $\dot{V}_{O_2}$, AT) by the use of reference equations. This has been performed in the primary analyses.

I find, to partially confirm my hypotheses, that OUES (when corrected as a percentage of predicted) had excellent discriminatory ability for CHF (and CHF plus MV disease) versus COPD. The other variables hypothesised to have strong discriminatory power performed less well. The $\dot{V}_{O_2} - WR$ slope had reasonable capacity, and was in keeping with the expected direction of difference, i.e. lower values in the cardiac groups. The $O_2$ pulse had highest values in the cardiac group which was in the opposite direction to the expected trend. This difference between groups was lost when $O_2$ pulse was calculated as a percentage of predicted. It appears that the high use of beta-blockers is the principal contributing factor to this variable; slower heart rates at equivalent points of exercise necessitate a higher stroke volume or oxygen extraction at the muscle, reflected in a higher $O_2$ pulse. The $\dot{V}_{O_2}$ at the AT (as a percent of predicted peak $\dot{V}_{O_2}$) and OUEP cannot discriminate, with lowest values in the latter associated with respiratory, not cardiac disease. Conceptually (and etymologically) the OUEP and OUES appear closely related, however they seem to behave very differently within this analysis. The breathing reserve, both at peak and at AT (the former non-significantly stronger than the latter) are very strong discriminators. Interestingly the long-standing proposed algorithm to differentiate patients with breathlessness performed badly. It could be argued that the presence of COPD or heart failure does not define
the lungs or heart as the limiting organ, but it does seem curious that so few patients (certainly within the CHF group) were classified in agreement with their objectively proven diagnoses.

Because there are a number of variables that have performed differently than expected, some of them will now be discussed separately.

### 6.5.1 Oxygen uptake efficiency slope

The OUES was first described in 1996 by Baba et al. This group noticed that the relationship between oxygen uptake and minute ventilation did not conform to a linear relationship unless the minute ventilation was logarithmically transformed. Defined as the slope of this relationship, a unit increase in OUES represents a unit increase in $\dot{V}_{O_2}$ for a ten-fold increase in minute ventilation. It is considered effort independent and so is more reliable than the peak $\dot{V}_{O_2}$ in patients where non-cardiorespiratory disease curtails exercise prematurely, or where patients are not maximally exercised for safety reasons. From my analyses and previous work OUES appears to be relatively insusceptible to respiratory abnormalities. I am unaware of any previous work involving OUES in patients with primary respiratory disease. Hollenberg et al identified the first reference ranges for OUES in a group of older healthy adults and found that FEV$_1$ and smoking status affected the OUES (Hollenberg et al 2000). However in that multivariate model used to identify an individual’s predicted OUES the contribution from FEV$_1$ and smoking was small. For example in an average-sized male adult the reduction in predicted OUES caused by a reduction in FEV$_1$ of 1L (a large change) is only 2.5%. Smoking contributed about 10%. The FEV$_1$ was a slightly more important contributor in the female reference ranges. Within the SHIP cohort I have shown that FEV$_1$ was strongly related to OUES, but as a percent of predicted FEV$_1$ the relationship was only weakly significant, suggesting that most of the influence comes from the impact of patient characteristics on both FEV$_1$ and OUES (i.e. age, gender, height, weight). The relationship with smoking was weak and in opposite directions in males and females (Chapter 4). Within a population of patients with CHF I have shown a strong relationship between percentage of predicted FEV$_1$ and peak $\dot{V}_{O_2}$, but not OUES (Barron et al 2010, Chapter 3). Similar results were found in another study where the addition of respiratory disease to heart failure minimally, and non-significantly, reduced OUES (Marinov et al 2008).

Within this observational study FEV$_1$ relates to OUES but with an inverse relationship, i.e. lower FEV$_1$ values related to higher OUES values. This likely represents an example of Simpson’s Paradox: a relationship noted individually in 2 or more groups can appear to reverse in direction when these groups are analysed together if
their group means are widely discrepant. In other words there may be a small positive relationship within each group (reduction in FEV\textsubscript{1} leads to a small reduction in OUES) but because the mean OUES in the cardiac group is significantly lower than in the COPD group, and the FEV\textsubscript{1} differences between groups lie in the opposite direction, when grouped together it looks like higher FEV\textsubscript{1} values lead to lower OUES. Whilst I do not believe that the inverse relationship is real, this result does add further weight to the argument that OUES is largely resistant to abnormalities in lung function. OUES also did not relate to $K_{\text{CO}}$ (Hb), a marker of diffusion capacity of the lungs that is corrected for lung volume and haemoglobin concentration.

However most important are the findings that OUES is significantly lower in the CHF and mitral valve groups despite similar exercise capacity (we would actually expect the mean OUES to be higher in the CHF group given the increased average weight and higher proportion of males) and on ROC curve analysis OUES, when corrected for weight or as a percentage of predicted value, was the strongest of the variables at identifying cardiac from respiratory disease where lower values predict cardiac abnormalities. Unlike breathing reserve, the only other variable with similar AUC values, its magnitude probably relates to exercise capacity (this hypothesis will be developed within Chapter 7), i.e. a reduction in OUES will mean a worsening of disease pathophysiology. Conceptually the same cannot be said for BR at peak, although the magnitude of BR at AT may be relevant if AT is unaffected by deteriorating respiratory disease. Why might OUES be largely resistant to abnormal lung function? Patients with heart failure typically have an abnormal peak $V_{\text{O}_2}$ but may still ventilate to high levels. They spend a considerable portion of exercise where the $V_E$ and $V_{\text{O}_2}$ relationships are decoupled because of anaerobic metabolism, with greater ventilatory inefficiency. Hence the linear relationship is “flattened”. In contrast COPD patients behave more like healthy adults that just do not progress to maximal exercise. They spend less time in anaerobic zones and so the relationship between $V_{\text{O}_2}$ and $V_E$, whilst still possibly abnormal, is more even throughout, leading to a more “normalised” OUES. This is only conjecture and is not examined specifically with the data I present here.

### 6.5.2 Breathing reserve

Multiple authors have suggested the use of breathing reserve as a discriminator of respiratory limitation (Eschenbacher et al 1990, Messner-Pellenc et al 1994, Medoff et al 1998, Milani et al 2004, Wasserman et al 4th Edition 2005). My analysis identified that a cut-off of 33.6% for the BR allowed for the greatest correct
classification of patients with a specificity of 100%; very close to the cut-off of 30% that has been suggested in previous studies.

One of these previous authors suggested using the BR at the AT, as this was proposed to reduce the unwanted impact on diagnostic power if patients voluntarily cease exercise before their respiratory system is maximally stressed (Medoff et al 1998). I found good discriminatory power of the BR at the AT, although the BR at peak appeared to perform better. The optimal cut-off in my data for BR at the AT was 48.7% (results not shown), not dissimilar to their cut-off of 42%. However 40% of the patients within the Medoff study and 35% of my COPD patients did not achieve the anaerobic threshold, limiting the applicability of this variable within a general population of patients with cardiorespiratory disease. In contrast BR at peak and the OUES are measurable in all.

Because RER was lower in the COPD group I investigated the role of the product of RER and BR. This showed a small, non-significant, improvement on BR at peak alone. Given the added complexity needed to record this variable this small improvement does not appear warranted.

My results suggest that using the BR at peak is perfectly acceptable if you wish a single value to identify principal limiting physiology. The magnitude of the value appears to be largely meaningless when measuring severity of disease/exercise intolerance (this will be tested further in Chapter 7), so the use of the BR in serial studies may be limited.

### 6.5.3 Anaerobic threshold

It has long been believed that a reduced anaerobic threshold on CPX is pathognomonic for heart failure. This forms a central part of an established algorithm for the identification of the principal organ system limiting exercise (Milani et al 2004, Wasserman et al 4th Edition 2005). Yet the evidence for this fact is scarce. In a study by Weber et al, a reduced $\dot{V}_{O_2}$ at the AT was found in patients with heart failure and related to functional limitation (Weber et al 1982). AT was found to be more sensitive than peak $\dot{V}_{O_2}$ at predicting prognosis in CHF (Gitt et al 2002). In a study by Nery et al, the $\dot{V}_{O_2}$ at the AT in a group of patients with mitral valve disease was lower than patients with COPD and healthy controls (Nery et al 1983); however the numbers were small with significant differences in gender and average age between groups. These studies alone appear to have laid the foundation of what has become a firmly held belief; that AT reflects cardiac function. A low AT has become
synonymous with “inadequate oxygen delivery” in the literature, but is this the case? There has been plenty of criticism of what the AT reflects and how it is measured (Hopker et al 2011), but if it really is the point when oxygen supply and demand diverge then does not the muscle have the central role in determining the AT? And muscular changes occur in COPD, CHF, sedentary behaviour and many other chronic diseases. Very few CPX studies performed on patients with COPD report the AT but in the study on the BR at the AT described above, Medoff et al showed that the $\dot{V}_{O_2}$ at the AT was the same in COPD and CHF patients with similar exercise capacities. Only the early study by Nery et al showed values in COPD similar to healthy controls. Therefore I suggest that the anaerobic threshold is critically determined by the muscle and any chronic process that impairs muscular function will reduce AT. Within this study the $\dot{V}_{O_2}$ at the AT (as a percentage of predicted peak $\dot{V}_{O_2}$) did not differ between groups and the AUC for its discriminant ability was very poor at 0.569. The use of absolute $\dot{V}_{O_2}$ at the AT was a better discriminator, but in the wrong direction, i.e. higher values in the CHF group.

It may be that with early heart failure a reduction in cardiac output will lead to a reduction in $\dot{V}_{O_2}$ at the AT prior to the muscular maladaptation that occurs later, but given the very wide normal limits of AT, this would probably be difficult to detect. A potential limitation to my criticism of AT as discriminant is the range of values seen within my population. Only 14 patients (2 with COPD, 7 with CHF, 4 with MV disease and 1 with mixed disease) had a reduced AT as defined by the cut-off of 40% predicted peak $\dot{V}_{O_2}$, and 28 with a value below 45%.

Categorising by an AT above or below 40% did not help its ability to discriminate (AUC 0.456, results not shown).

6.5.4 Other variables

Similarly to the AT, the $\dot{V}_{O_2} - WR$ slope has been regarded as a variable reflecting oxygen delivery, and therefore is believed to be reduced in patients with heart failure but not COPD (Hansen et al 1987). Hypotheses for the cause of the reduced slope include slow oxygen kinetics and increased utilisation of anaerobic sources of metabolism after the AT. In this seminal study by Hansen et al, (where they measured $\Delta \dot{V}_{O_2} / \Delta WR$ instead of the regression line of the slope), the value for patients with circulatory disease was 8.29, compared with 10.29 for the healthy controls. However, although they attempted to adjust the protocol for exercise capacity to ensure participants exercised for similar lengths of time, they do not show the average duration of exercise within each group, and 14/51 patients didn’t achieve 6 minutes; they did not reveal if similar findings were seen in the control group. As I have shown in Chapter 5, the $\dot{V}_{O_2} - WR$ slope is heavily dependent on protocol with longer
protocols elevating the slope in healthy adults and patients with CHF and COPD. Therefore our study protocol
is ideally designed to answer the hypothesis regarding the slope being a marker of impaired circulatory function,
in that test 2 in all patients was specifically picked based on their own results from test 1 to lead to a
standardised length of time undertaking incremental exercise (around 10 minutes). With this attempt at
standardisation of time, patients with mitral valve disease and CHF had significantly lower $\dot{V}_{O_2} - WR$ slopes
(8.7) compared to COPD patients (9.6) although on average there was a trend to longer exercise times in the
cardiac groups. The AUC was approximately 0.70 in the 3 analyses, so the slope appears to have moderate
ability to discriminate. One of the key findings from Chapter 5 is that the $\dot{V}_{O_2} - WR$ slope does not have
excellent test-retest reliability and it would be unexpected for a variable that does not have good reproducibility
to show a good ability to discriminate. Perhaps if methodology for the measurement of the slope alters to
improve reproducibility this analysis could be rerun to see if this improves the ability of the slope to
discriminate.

The oxygen uptake efficiency plateau (OUEP) is a newly described variable that has shown similar prognostic
power to OUES. However, unlike OUES, it didn’t appear to significantly discriminate between these groups
(AUC 0.63-0.67) with mean OUEP lowest in the COPD group.

The O$_2$ pulse showed reasonable ability to discriminate, however the lowest values were paradoxically found in
the COPD group, and once corrected using reference ranges any differences largely disappeared. It is believed
that the O$_2$ pulse reflects stroke volume and therefore in heart failure, a condition of reduced contractile ability
of the heart (amongst other abnormalities), it should be lowest. This may have been appropriate in the pre-beta
blockade era, but now most CHF patients are treated with beta blockers and the reduced heart rate leads to
increased compensatory stroke volume (through increased ventricular filling time) or oxygen extraction, which
will be reflected by an O$_2$ pulse that normalises or may even become supranormal.

It appears that neither the OUEP nor O$_2$ pulse are ideally suited to discriminate.

Measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship have been shown to be abnormal in heart failure and COPD, and are strong
prognostic markers, but their association with both conditions unsurprisingly means that within this study
population they did not significantly differ between groups or have the ability to discriminate. There was an
interesting observation that on average the slope was higher in the cardiac patients than the ratio at nadir, whilst
in COPD patients the difference was in the opposite direction. When I created a variable of slope minus ratio at
nadir (using the slope measuring data pre-VCP as is convention) there was a significant difference between
groups with a good ability to discriminate (AUC 0.76-0.78). Whilst this new variable may show some promise,
the magnitude does not clearly relate to a specific pathophysiological abnormality and so may not add anything
further to the variables already commonly used. Further work is necessary.

The double product was significantly lower in the CHF group and this is reflected in its reasonable
discriminatory power (AUC 0.79 for CHF vs COPD), however this was significantly reduced by the addition of
mitral valve disease, suggesting it may not so much reflect the abnormal circulatory function as the increased
numbers of anti-hypertensive agents used within the CHF group.

Unsurprisingly RER at peak, peak heart rate and oxygen saturations were different between the groups, but did
not add much over the other variables in terms of discriminatory ability. The $HR - V_{O_2}$ slope was lowest in the
CHF group and highest in the mitral group. This U shaped relationship is to be expected. Patients with COPD
behave reasonably typically, like healthy adults, but patients with circulatory abnormalities can behave in one of
two ways; firstly there can be an excessive heart rate response to compensate for the reduced forward stroke
volume; and secondly chronotropic incompetence (often due to beta blockers). Valve patients, who are less
frequently prescribed beta blockers, would be expected to behave in the former manner as we largely see here,
with CHF patients the latter. This bivariate effect limits the likelihood of this variable being useful; when mitral
and CHF patients are considered together it had no discriminant ability, although for CHF alone it is better
(AUC 0.63).

Reason for stopping, when categorised into either breathlessness or fatigue, was not particularly discriminant.
Given that this also hasn’t been shown to be prognostically powerful (Witte et al 2008) it argues against a need
for the common approach of recording reason for test cessation.

6.5.5 Clinical relevance

My results suggest that whilst breathing reserve should retain its role at the heart of a decision-making tree into
principal limiting physiology, the oxygen uptake at the AT is not sufficiently discriminant to add further
information and should not be used in this capacity. Few typical heart failure patients fulfil the recommended
AT cut-off of 40%. The OUES shows real promise; it is highly reproducible with a likely linear relationship to
disease severity, hence I would propose that should a physician be attempting to track a patient’s “cardiac”
contribution to exercise intolerance this would be the best choice for a variable to follow. In a patient with coexistent respiratory and cardiac disease it could help differentiate between a decline in exercise capacity due to worsening cardiac versus respiratory pathology.

### 6.5.6 Limitations

The main limitation is the numbers of patients recruited. I had performed sample study calculations suggesting a requirement of 70 cardiac and 35 COPD patients; I was just short of both of those targets, however, perhaps because OUES had a greater AUC than anticipated, the results are highly significant. It will be interesting to see if this AUC is repeatable, and with 2 equally sized populations. This would ideally be performed by an independent study group.

Secondly some of the patients within the study population, although proven to have either COPD or CHF, were not as symptomatic as expected, with a reasonable proportion achieving a reasonably normal peak $\dot{V}_{O_2}$ (17 patients achieved 85% predicted peak $\dot{V}_{O_2}$ (18%), with 3 achieving 100% predicted). It would be interesting to see if these results could be replicated in a population of more severely affected patients, for example those awaiting heart or lung transplant.

### 6.5.7 Conclusions

OUES and breathing reserve are the best variables routinely measured on CPX at discriminating heart failure from COPD. These variables also have results largely unaltered by the addition of mitral valve pathology to the CHF group, and including patients with both pathologies under their principal diagnosis, unlike other variables.

The anaerobic threshold does not discriminate CHF patients. The $\dot{V}_{O_2} - WR$ slope shows an intermediate ability to discriminate, no other variable shows much promise. In a patient with heart failure and COPD, breathing reserve and OUES could be used to identify principal limiting physiology and for serial assessment of progression track the changes caused by each pathology separately. Further assessment of the role of these variables in patient with mixed disease is necessary.
7.0 Correlation of CPX variables to disease severity
7.1 Abstract

Different cardiopulmonary exercise test (CPX) variables give us different information on the impact of cardiorespiratory disease on a patient. Some of these variables can be useful to distinguish limiting pathology, but it is unclear which variables most closely reflect disease severity.

96 patients with COPD, CHF or mitral valve disease, recruited into the study had two CPX tests, an echocardiogram, full lung function tests and BNP measurement. Symptom status prior to exercise testing was established using the NYHA classification system. NYHA class, echocardiographic measures including ejection fraction, fractional shortening, s’ velocity, E:e’ ratio, left atrial volumes, TAPSE and right ventricular S’, and BNP were analysed against 24 CPX variables using ANOVA, interaction regression models and Spearman’s rank correlation.

NYHA class related to a number of variables amongst the full patient population and to OUES, anaerobic threshold and $\dot{V}_{O_2} - WR$ slope in cardiac patients alone. OUES, double product and the $\dot{V}_{O_2} - WR$ slope correlated most closely to the echocardiographic variables. Most variables correlated to BNP within the full population and the cardiac patients when analysed alone, but not in the COPD patients; anaerobic threshold did not correlate to BNP. When the correlation of other variables against peak $\dot{V}_{O_2}$ (as the overall measure of exercise capacity) was tested almost none correlated in the COPD patients, and almost all correlated within the cardiac patients. Breathing reserve did not correlate to peak $\dot{V}_{O_2}$ within either group confirming its inability to reflect disease severity in either COPD or heart failure.

OUES and the $\dot{V}_{O_2} - WR$ slope appear to relate closely to disease and symptom severity in patients with heart failure and mitral valve disease, but not in patients with COPD. They may be ideally suited to the follow-up of patients with cardiac disease to specifically reflect disease severity and progression.
7.2 Introduction

One of the key roles of CPX is to identify the principal limiting organ in a patient complaining of reduced exercise tolerance. I have shown in previous chapters how certain CPX variables are more discriminant of cardiac over respiratory disease (or vice versa). However the ability of a single exercise test to discriminate primary pathology limiting exercise is not the main concern. An algorithm approach (although not the recommended one) may well be sensitive and specific enough to identify principal limiting physiology, but where it fails to help is when a patient returns for a subsequent test to assess disease progression. For this physicians have largely used peak $\dot{V}_{O_2}$, considered the endpoint of all disease processes on exercise capacity. Peak $\dot{V}_{O_2}$ is often used synonymously with exercise capacity.

For many patients peak $\dot{V}_{O_2}$ will perform admirably when serial tests are required, but if, for whatever reason, peak $\dot{V}_{O_2}$ is not felt to be reflective of a patient’s disease progression it is unknown if other variables can perform in this role. An alternative variable would need to reflect both the disease severity and influence of the disease on exercise capacity.

Within this chapter I explore the relationships between a number of the principal CPX variables and subjective and objective measures of disease severity, to see which variables would be appropriate to use in the serial testing of patients with a cardiac cause of exercise limitation.
7.3 Methods

7.3.1 Patient recruitment
This chapter involves only patients analysed within the Observational study (Chapter 6) and methodology for patient recruitment is described in detail within the Methods Chapter and Chapter 6. Each patient, prior to exercise test 1, was assessed for symptoms, and a symptom score based on NYHA classes 1-4 was assigned. Initially this was only performed for cardiac patients (as NYHA class has only been validated within cardiac patients) but was introduced for COPD patients after the first 7 patients had already completed the protocol. A symptom score validated for respiratory patients is the MRC breathlessness scale; this is a 1-5 scale and so was not adopted for this study because it could not be easily compared with a 1-4 scale for cardiac patients. Prior to exercise testing each patient also had venous bloods taken including B-Natriuretic Peptide (BNP). Based on previous analyses I have presented, it is clear that within my population BNP fits a skewed distribution that appears normally distributed following logarithmic transformation. Therefore this was performed for all patients BNP values.

7.3.2 CPX analysis
This is described in detail elsewhere. Briefly patients undertook 2 CPX bicycle ergometer tests. The first was a familiarisation test, with 3 minutes of rest, 3 minutes of unloaded cycling and then a 10W/min incremental protocol until symptoms limited further exercise and the test was stopped. Based on the results of this first test each patient had a second test with the same rest and unloaded sections, but with an incremental protocol designed to reach symptom limited maxima at approximately 10 minutes. All analysis was performed on the patients’ second tests. The following CPX variables were analysed within these analyses; peak $\dot{V}_{O_2}$ (mL/min), peak $\dot{V}_{O_2}$ (mL/min/kg), peak $\dot{V}_{O_2}$ (percent of predicted), $\dot{V}_{O_2}$ at the AT (mL/min), $\dot{V}_{O_2}$ at the AT (percent of predicted peak $\dot{V}_{O_2}$), OUES (L/min/10-fold increase in $\dot{V}_E$), OUES/kg, OUES (percent of predicted), OUEP, $O_2$ pulse (mL/beat), $O_2$ pulse (percent of predicted), $\dot{V}_E/\dot{V}_{CO_2}$ slope until the VCP, $\dot{V}_E/\dot{V}_{CO_2}$ slope throughout full exercise, $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir, $\dot{V}_E/\dot{V}_{CO_2}$ ratio at AT, $\dot{V}_E/\dot{V}_{CO_2}$ slope minus the $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir, end-tidal partial pressure of CO$_2$ (P$_{ET CO_2}$), double product (DP), peak circulatory power, breathing reserve (BR) at the AT and peak exercise, peak minute ventilation, $\dot{V}_{O_2} - WR$ slope and $HR - \dot{V}_{O_2}$ slope.

The percentages of predicted for peak $\dot{V}_{O_2}$ (and by extension $\dot{V}_{O_2}$ at the AT), OUES and $O_2$ pulse use the reference equations described previously (Gläser et al 2010, Chapter 4).
7.3.3 Echocardiographic analysis

All patients underwent an echocardiogram on the same day as exercise testing. I undertook each echo using the IE33 imaging system (Philips, Amsterdam, The Netherlands). Standard echocardiographic views were attempted in every patient and are explained in detail in the Methods Chapter. For this analysis only 7 measures were analysed. The first 3 were measures of systolic function and included: ejection fraction as measured by Simpson’s BiPlane method (occasionally when views in only one plane were possible these single plane volumes alone were used); fractional shortening, using the left ventricular internal dimension at diastole and systole as measured in the parasternal long axis view at the mitral leaflet tips; and mean s’ velocity, this is the average of the mitral and lateral annular s wave velocities on tissue Doppler echocardiography (these represent the forward, systolic, movements). The next 2 were variables of diastolic function and included: left atrial volume (LAV) which was calculated using the area/length method; and the E:e’ ratio, where the ratio of the transmitial filling E wave to the mean of the mitral and lateral annular e wave velocities on tissue Doppler echocardiography (these represent the first backwards, or passive diastolic, movements). Finally I also measured two right ventricular variables: the tricuspid annular plane systolic excursion (TAPSE) and right ventricular systolic (S’) wave on tissue Doppler imaging. Full techniques are described in the Methods Chapter.

7.3.4 Hypotheses

I hypothesised that an ideal variable of cardiac function would correlate to symptom severity in cardiac, but not COPD, patients; and correlate to objective measures of cardiac function (for example echocardiographic measurements of systolic and diastolic function) within cardiac, but not COPD, patients. I also hypothesised that, as peak $\dot{V}_O_2$ is the accepted measure of overall exercise capacity, an ideal variable will correlate with peak $\dot{V}_O_2$. This correlation should be strong ($r>0.50$), but not close to 1, as this will show that almost no extra information is gained from adding this second variable to peak $\dot{V}_O_2$.

7.3.5 Statistical analysis

Tests for normally distributed data have been performed on all CPX variables and BNP in previous chapters. Briefly BNP is skewed and appears normally distributed when logarithmically transformed; hence log transformed BNP will be used throughout. Similarly left atrial volume (LAV) is skewed and is normally distributed after logarithmic transformation and this will be used throughout. Whilst some of the CPX variables
are non-normally distributed ($\dot{V}_E/\dot{V}_{CO_2}$ slopes), they do not become normally distributed on logarithmic
transformation so were analysed in their raw format and non-parametric tests used wherever necessary.

NYHA class was analysed for its relation to CPX variables using one-way analysis of variance (ANOVA) within
the whole patient population and by respiratory and cardiac groups. Spearman’s rank correlation (because some
of the CPX variables are non-normally distributed) was used to analyse CPX variables for a correlation with
echocardiographic measures of systolic function (ejection fraction, fractional shortening and mean $s$’),
echocardiographic measures of diastolic function (LA volume and E:e’ ratio), echocardiographic measures of
right ventricular function (TAPSE and RV S’) and BNP.

Peak $\dot{V}_{O_2}$, as an overall variable of exercise capacity, was correlated to the remaining CPX variables; also using
Spearman’s rank correlation test.

To analyse if the relationship between a CPX variable and the outcome variable (i.e. NYHA class,
echocardiographic measures etc) is different in COPD compared with cardiac disease, a multivariate model was
designed with the CPX variable and diagnosis as an interaction term. A statistically significant interaction term
indicates that the CPX variable is significantly differently related to the outcome variable in cardiac versus
respiratory patients.

A p value of <0.05 is considered significant throughout.
7.4 Results

7.4.1 Patient characteristics
Details of recruitment numbers and dropout rates are reported in Chapter 6. This chapter reports on the same 96 patients undertaking the Observational study. Briefly from 96 patients (72 male), 43 were patients with CHF (14 undergoing CRT and 13 having undergone CRT), 26 with mitral valve disease requiring surgical correction and 27 with lung disease (25 showed obstructive physiology and 2 restrictive physiology but had a formal physician-made diagnosis of COPD). 18 from 26 patients with mitral valve disease were self-described as symptomatic. 7 from 96 patients fulfilled criteria for inclusion in the mixed group. 4 of these patients had COPD as their primary diagnosis but 3 were known to have ischaemic heart disease and on echocardiography had abnormalities of left ventricular function, and a fourth had right atrial and ventricular dilatation with an elevated BNP. 3 heart failure patients were known to have COPD. 5 further heart failure patients and 5 mitral valve patients had obstructive spirometry which met GOLD criteria category 2 but not higher. 2 were known to have long-standing asthma (not COPD) but none of the remaining patients had a formal secondary diagnosis or symptoms suggestive of COPD, and only 1 was a current smoker. Of these 10 patients 4 were reclassified with non-obstructive or mildly obstructive spirometry (GOLD category 1) on repeat spirometry, 2 had FEV₁ values just below the 80% cut-off to distinguish between category 1 and 2, and 4 lay well within category 2. They were categorised within their principal disease category. Within the COPD category, 2 patients met the criteria for GOLD category 1 (mild), 12 patients met criteria for GOLD category 2 (moderate), 10 for category 3 (severe) and 3 for category 4 (very severe).

There were significantly more males in the CHF group; this group was also heavier, with a lower haemoglobin and estimated glomerular filtration rate (eGFR).

7.4.2 Relation to symptom class
The simplest way to describe a patient’s functional status is a symptom class. Within this study 65/66 patients with cardiac disease had an NYHA class recorded prior to exercise testing. For COPD patients the equivalent scoring system was performed on 16/23 patients for reasons described above.

Symptom class was significantly related by ANOVA to peak $\dot{V}O_2$ (F= 10.35, p<0.001), peak $\dot{V}O_2$/kg (F= 14.3, p<0.001) and percent predicted peak $\dot{V}O_2$ (F= 9.67, p<0.001). $\dot{V}O_2$ at the AT, OUEP, O₂ pulse, variables of the $\dot{V}E/\dot{V}CO_2$ relationship, circulatory power, and peak $\dot{V}E$ were also all highly significantly correlated to symptom
class (Table 7.1). OUES and the $\dot{V}_{O_2} - WR$ slope were also significantly correlated, but with borderline p values. Measures of the AT and OUES (Figure 7.1), the $\dot{V}_{O_2} - WR$ slope, and most variables of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship did not differ between groups of symptom class in patients with COPD, but they did in patients with CHF and MV disease. Breathing reserve at the AT, but not at peak, was related to symptom class in patients with COPD but not cardiac disease.

**7.4.3 Correlation with markers of cardiac dysfunction**

For a variable to be sensitive and specific for cardiac dysfunction it follows that it will correlate to objective markers of cardiac dysfunction; at least within the CHF and MV populations. CPX variables were analysed against echocardiographic variables and BNP.

For markers of systolic function, ejection fraction (EF - measured by Simpson’s biplane method), fractional shortening (FS) and the mean of the mitral septal and lateral annular s’ velocities, correlations were performed against some of the principal CPX variables. Double product was highly correlated to all 3 measures ($r=0.49$, $r=0.53$, $r=0.52$ respectively, $p<0.0001$ for all). OUES as an absolute value, weight adjusted and percent of predicted was highly correlated to mean s’ velocity ($r=0.40$, $r=0.41$, $r=0.34$ respectively, $p<0.0001$ for all) and when adjusted for weight to fractional shortening ($r=0.31$, $p=0.003$). Peak $\dot{V}_{O_2}$ was highly correlated to mean s’ velocity only when weight adjusted ($r=0.36$, $p<0.001$) but not as a percent of predicted.

As markers of diastolic function left atrial volume (LAV) and the ratio of transmitral E wave to mean mitral septal and lateral annular e’ velocities (E:e’ ratio) were calculated. Only OUES/kg ($r=-0.39$, $p<0.001$), percent of predicted OUES ($r=-0.38$, $p<0.001$) and the $\dot{V}_{O_2} - WR$ slope ($r=-0.36$, $p<0.001$) correlated strongly with LAV, whereas OUES (as an absolute, weight adjusted and percent of predicted value) was highly correlated to the E:e’ ratio ($r=-0.43$, $p<0.001$; $r=-0.41$, $p<0.001$; and $r=-0.40$, $p<0.001$ respectively) along with the $\dot{V}_{O_2} - WR$ slope ($r=-0.42$, $p<0.001$). The relationship between E:e’ and s’ on OUES and peak $\dot{V}_{O_2}$ is shown in Figure 7.2.

Measures of right ventricular function were also tested for correlation against some of the principal CPX variables. Peak $\dot{V}_{O_2}/kg$ ($r=0.29$, $p=0.005$), $\dot{V}_{O_2}$ at the AT ($r=0.30$, $p=0.007$), OUES ($r=0.30$, $p=0.004$), $\dot{V}_E/\dot{V}_{CO_2}$ slope ($r=-0.33$, $p=0.002$) and ratio ($r=-0.31$, $p=0.003$), double product ($r=0.41$, $p<0.001$) and circulatory power ($r=0.41$, $p<0.001$) were all correlated to TAPSE.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Full group</th>
<th>COPD group</th>
<th>Cardiac group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $V_{O_2}$ (mL/min)</td>
<td>10.35</td>
<td>4.65</td>
<td>9.29</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (mL/min/kg)</td>
<td>14.30</td>
<td>2.18</td>
<td>19.53</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (% predicted)</td>
<td>9.67</td>
<td>3.54</td>
<td>9.86</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>8.80</td>
<td>0.38</td>
<td>7.51</td>
</tr>
<tr>
<td>AT (% of pred peak $V_{O_2}$)</td>
<td>4.01</td>
<td>0.08</td>
<td>3.91</td>
</tr>
<tr>
<td>OUES</td>
<td>2.91</td>
<td>0.29</td>
<td>6.80</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>3.01</td>
<td>0.23</td>
<td>10.99</td>
</tr>
<tr>
<td>OUES (% predicted)</td>
<td>0.77</td>
<td>0.38</td>
<td>4.49</td>
</tr>
<tr>
<td>OUEP</td>
<td>13.38</td>
<td>4.69</td>
<td>10.32</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat)</td>
<td>4.22</td>
<td>8.32</td>
<td>2.08</td>
</tr>
<tr>
<td>$O_2$ Pulse (% predicted)</td>
<td>4.92</td>
<td>4.83</td>
<td>4.33</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 1</td>
<td>8.34</td>
<td>0.75</td>
<td>11.58</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 2</td>
<td>3.53</td>
<td>0.25</td>
<td>4.80</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio nadir</td>
<td>9.62</td>
<td>2.52</td>
<td>8.47</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio AT</td>
<td>14.01</td>
<td>5.81</td>
<td>7.31</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope - ratio</td>
<td>3.59</td>
<td>0.94</td>
<td>7.88</td>
</tr>
<tr>
<td>RER at peak</td>
<td>2.90</td>
<td>6.20</td>
<td>1.20</td>
</tr>
<tr>
<td>$P_{ET}CO_2$ at AT (mmHg)</td>
<td>13.1</td>
<td>6.85</td>
<td>6.85</td>
</tr>
<tr>
<td>DP (mmHg bpm)</td>
<td>2.12</td>
<td>0.64</td>
<td>3.29</td>
</tr>
<tr>
<td>Circ power (mmHg mL/min)</td>
<td>7.59</td>
<td>3.93</td>
<td>7.45</td>
</tr>
<tr>
<td>BR at AT (%)</td>
<td>2.68</td>
<td>9.02</td>
<td>1.96</td>
</tr>
<tr>
<td>BR (%)</td>
<td>1.49</td>
<td>1.65</td>
<td>0.08</td>
</tr>
<tr>
<td>Peak $\dot{V}_E$ (L/min)</td>
<td>4.41</td>
<td>3.33</td>
<td>3.18</td>
</tr>
<tr>
<td>$V_{O_2} - WR$ slope</td>
<td>3.14</td>
<td>0.46</td>
<td>7.88</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ slope</td>
<td>0.26</td>
<td>0.31</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 7.1: Association (by ANOVA) between symptom class (I-IV) and CPX variables in the full group, patients with COPD, and patients with mitral valve disease and CHF.
Figure 7.1: Change in 4 CPX variables with symptom class. Left hand plot shows values for the full population, the right hand plot shows pairs within each variable, COPD on the left and cardiac disease on the right. All variables are expressed as a percent of predicted. It can be seen that for peak $\dot{V}_{O_2}$, O$_2$ pulse and the $\dot{V}_{O_2}$ at the AT the trends in both disease categories are similar, whilst for OUES the directions are opposite between the two disease categories.
Figure 7.2: Correlation between tissue Doppler imaging, OUES and peak $V_{O_2}$. E:e’ ratio as a marker of diastolic function and mean s’ velocity as a marker of systolic function were correlated against 2 CPX variables, peak $V_{O_2}$ and OUES (both as weight adjusted variables). It can be seen that the correlations with OUES are stronger than for peak $V_{O_2}$. 
Peak $\dot{V}_{O_2}\text{/kg (r=0.29, p=0.007)}, \dot{V}_{O_2}$ at the AT (r=0.26, p=0.02), OUES/kg (r=0.38, p<0.001), double product (r=0.51, p<0.001) and circulatory power (r=0.44, p<0.001), but not the measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship were correlated to the right ventricular S’ wave.

On the interaction model double product was the only CPX variable influenced by disease category in its relation to the mean LV s’ velocity. There were no other significant interactions when assessing any of the principal CPX variables on E:e’ ratio, ejection fraction, fractional shortening, LAV, TAPSE or RV S’ wave.

BNP was strongly correlated to a number of CPX variables within the whole group. Figure 7.3 shows the relationship between BNP and both OUES and peak $\dot{V}_{O_2}$. Peak $\dot{V}_{O_2}$ as an absolute value (r=−0.32, p=0.003) and weight adjusted (r=−0.37, p<0.001), OUES as an absolute, weight adjusted and percent of predicted (r=−0.53, r=−0.57, r=−0.56 respectively, p<0.001 for all), $\dot{V}_E/\dot{V}_{CO_2}$ slope (r=0.48, p<0.001), $\dot{V}_E/\dot{V}_{CO_2}$ ratio at nadir (r=0.41, p<0.001) and at AT (r=0.42, p<0.001), double product (r=−0.56, p<0.001) and circulatory power (r=−0.49, p<0.001), breathing reserve at peak (r=0.33, p=0.002) and the $\dot{V}_{O_2} - WR$ slope (r=−0.44, p<0.001) were all strongly correlated with BNP. Almost all variables had a significant interaction term; on further analysis the presence of COPD significantly reduced the correlation between CPX variables and BNP so that none were significantly correlated, whilst the same correlation in the cardiac patients was largely improved or remained the same, with the exception of breathing reserve at peak which was not correlated to BNP in either group independently.

### 7.4.4 Correlation with peak $\dot{V}_{O_2}$

Almost all principal CPX variables correlated to peak $\dot{V}_{O_2}$ (as weight adjusted and percent predicted) (Table 7.2). Within the full population of patients only the $\dot{V}_E/\dot{V}_{CO_2}$ slope-ratio difference, breathing reserve and the $HR - V_{O_2}$ slope did not correlate strongly with peak $\dot{V}_{O_2}$/kg or percent predicted peak $\dot{V}_{O_2}$ (p<0.002 for all).

On the interaction model all measures of OUES and peak $\dot{V}_E$ were the only variables that showed a significant interaction with disease category on their correlation with peak $\dot{V}_{O_2}$. The correlation between OUES and peak $\dot{V}_{O_2}$ was significantly stronger in the cardiac patients (OUES/kg vs peak $\dot{V}_{O_2}$/kg r=0.84, p<0.0001; percent predicted OUES vs percent predicted peak $\dot{V}_{O_2}$ r=0.71, p<0.0001) than the COPD patients (OUES/kg vs peak $\dot{V}_{O_2}$/kg r=0.50, p=0.02; percent predicted OUES vs percent predicted peak $\dot{V}_{O_2}$ r=0.20, p=0.36) (Figure 7.4).
Figure 7.3: Correlation between log BNP and both OUES and peak $\dot{V}_{O_2}$ (both as weight adjusted variables). It can be seen that the correlation between log BNP and OUES is stronger than for peak $\dot{V}_{O_2}$. 
Figure 7.4: Scatter plot of percent of predicted OUES against percent of predicted peak $V_{O_2}$. Plot 1 shows all COPD (red) and cardiac patients (blue). Plot 2 shows only COPD patients with a weaker correlation coefficient. Plot 3 shows only cardiac patients with a stronger correlation coefficient.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Peak $\dot{V}_{O_2}$ (mL/min/kg)</th>
<th>Peak $\dot{V}_{O_2}$ (% predicted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$p$</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>0.596</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AT (% of pred peak $\dot{V}_{O_2}$)</td>
<td>0.670</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OUES</td>
<td>0.511</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>0.666</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OUES (% predicted)</td>
<td>0.392</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OUEP</td>
<td>0.550</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat)</td>
<td>0.400</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$O_2$ Pulse (% predicted)</td>
<td>0.554</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 1</td>
<td>-0.500</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 2</td>
<td>-0.399</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio nadir</td>
<td>-0.548</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio AT</td>
<td>-0.469</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope - ratio</td>
<td>0.241</td>
<td>0.018</td>
</tr>
<tr>
<td>$P_{ETCO_2}$ at AT (mmHg)</td>
<td>0.467</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DP (mmHg bpm)</td>
<td>0.461</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Circulatory power</td>
<td>0.681</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BR at AT (%)</td>
<td>0.031</td>
<td>0.780</td>
</tr>
<tr>
<td>Breathing reserve (%)</td>
<td>-0.100</td>
<td>0.335</td>
</tr>
<tr>
<td>Peak $\dot{V}_E$ (L/min)</td>
<td>0.536</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\dot{V}_O_2$ – WR slope</td>
<td>0.458</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HR – $\dot{V}_O_2$ slope</td>
<td>0.026</td>
<td>0.805</td>
</tr>
</tbody>
</table>

**Table 7.2: Correlation between peak $\dot{V}_{O_2}$ and other CPX variables.** Correlation coefficients and $p$ values were calculated for peak $\dot{V}_{O_2}$ (mL/min/kg) and percent of predicted peak $\dot{V}_{O_2}$. 
7.5 Discussion

The purpose of this thesis is to establish which CPX variable should be used, if your patient has both cardiac and respiratory disease, to monitor the cardiac impact on exercise capacity. Previous chapters have explored the ability of certain variables to discriminate better than others, and future chapters will investigate the change in variables following cardiac interventions.

But an important criterion for a variable to be truly reflective of cardiac dysfunction is that it varies predictably with disease severity, or at least does so within a population of patients with cardiac disease, whilst remaining unrelated to disease severity in patients with other disease processes, for example respiratory disease.

Therefore within this chapter I have shown which variables are reflective of subjective measures of disease severity such as symptom status, and objective measures of disease severity such as echocardiographic variables and BNP, as well as the correlation between CPX variables and peak $\dot{V}_{O_2}$, which I use as a surrogate for overall exercise capacity. Finally I have analysed which CPX variables vary in their relation to these markers of disease severity by disease category.

7.5.1 NYHA classification

Symptom status, commonly described in heart failure patients using the New York Heart Association (NYHA) class, is a simple test but suffers from poor reproducibility (Goldman et al 1981, Raphael et al 2007). I have used it here because it requires almost no extra time or effort to record and in large populations allows us to view trends adequately despite this poor reproducibility. It is known to correlate poorly to ejection fraction yet moderately to peak $\dot{V}_{O_2}$ (Van den Broek 1992), although generally offers little extra information beyond peak $\dot{V}_{O_2}$ to aid in heart failure prognosis (Likoff et al 1987, Cohn et al 1993). Within this study peak $\dot{V}_{O_2}$, OUES and the $\dot{V}_E/\dot{V}_{CO_2}$ ratio were amongst the variables that best reflected symptom status regardless of disease category. The variables that changed significantly when divided by disease category were the $\dot{V}_{O_2}$ at the AT, OUES, $\dot{V}_E/\dot{V}_{CO_2}$ slope and the $\dot{V}_{O_2} – WR$ slope (which all became strongly correlated in the cardiac group and non-correlated in the COPD group), and the breathing reserve at AT (which became strongly correlated in the COPD group and remained non-correlated in the cardiac group).
7.5.2 Objective measures of respiratory and cardiac dysfunction

Thereafter I have used more objective measures of disease severity. Markers of respiratory severity were discussed in Chapter 6, where peak $\dot{V}_O_2$, the OUEP, $O_2$ pulse and breathing reserve all showed a significant relation to $FEV_1$ and $K_{CO_2}$ as measures of respiratory severity, and in the appropriate direction, i.e. worsening respiratory function related to a worsening of the CPX variable.

When we consider objective measures of cardiac dysfunction we typically think about systolic function. However it has long been established that echocardiographic indices of systolic dysfunction, namely ejection fraction, correlate poorly to exercise capacity (Franciosa et al 1981, Szlachcic et al 1985, Cohn et al 1988, Van den Broek et al 1992). However in a single study OUES correlated closely with ejection fraction in a group of patients with heart failure and reduced ejection fraction (Straburzyńska-Migaj et al 2010). Another problem is that patients with heart failure may have a normal ejection fraction; this group of patients, previously labelled as diastolic heart failure, would show an even poorer relationship between EF and objective exercise capacity. I limited the recruitment of heart failure patients to those with reduced ejection fractions, however heterogeneous abnormalities of cardiac function will still exist that make systolic dysfunction much more complex than just a reduced ejection fraction (for example impaired relaxation, valvular dysfunction, elevated pulmonary venous pressure). Unsurprisingly most CPX variables within my analysis did not correlate to ejection fraction, both across the whole group and when limited to CHF patients alone (results not shown). A more recently described marker of systolic function is the annular longitudinal velocities of systolic wall motion on tissue Doppler imaging. This has previously been shown to correlate more closely to peak $\dot{V}_O_2$ than ejection fraction does (Witte et al 2004), and within my analysis peak $\dot{V}_O_2$/kg, all measures of the OUES, double product and circulatory power all correlated closely with this marker; interestingly the correlation coefficients for OUES were higher than for peak $\dot{V}_O_2$. OUES also strongly correlated with left atrial volume and E:e’ as markers of diastolic function, as did the $\dot{V}_O_2 - WR$ slope.

OUES and peak $\dot{V}_O_2$ both had similar correlations with TAPSE, whilst variables of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship, DP and circulatory power were the most significantly correlated. DP and circulatory power were also very strongly correlated to the RV S’ wave, with OUES showing a stronger correlation than peak $\dot{V}_O_2$. In contrast to TAPSE, measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship did not correlate strongly with this measure of RV function. It has previously been shown that the $\dot{V}_E/\dot{V}_{CO_2}$ relationship related independently on multivariate analysis to TAPSE but not left ventricular abnormalities in patients with severe heart failure; both were strong predictors of death or
heart transplantation (Methvin et al 2011). On these analyses TAPSE overall correlates more closely with CPX variables than ejection fraction or fractional shortening do. This probably reflects the compliant right ventricle’s response to the constant burden of elevated pulmonary pressures secondary to left ventricular dysfunction. Whilst we see heterogeneous responses within the left ventricle, meaning that no single marker of “impairment” will strongly correlate to other global markers of heart failure, the right ventricular response is more uniform with dilatation and systolic impairment.

BNP measured from venous blood sampling reflects raised intra-cardiac pressures, and is typically elevated in patients with heart failure. It has been previously shown to correlate closely with OUES (Straburzyńska-Migaj et al 2010) and peak $\dot{V}_O_2$ (Williams et al 2004). My results largely agree with these studies; OUES, peak $\dot{V}_O_2$, all measures of the $\dot{V}_E$/CO, relationship, DP and circulatory power and $\dot{V}_O_2 - WR$ slope all closely correlate with BNP across the full study group. However when dividing by disease category these close correlations are lost in patients with COPD, and maintained within the cardiac population.

7.5.3 Correlation between peak $\dot{V}_O_2$ and other cardiopulmonary exercise variables

Finally I analysed the relationship between CPX variables; and the effect of disease aetiology on these relationships. Rather than analysing 15-20 variables against one another I chose peak $\dot{V}_O_2$, because this probably best reflects a patient’s overall exercise capacity. For any variable to be able to be considered “Ideal” it should correlate to peak $\dot{V}_O_2$. This relationship should be strong but not 100% (as otherwise it would not add further information to peak $\dot{V}_O_2$). Alternatively the variable could have a correlation close to 100% with peak $\dot{V}_O_2$ in cardiac patients, but little or no correlation in respiratory disease, although this would be highly unlikely to occur physiologically. Importantly breathing reserve, both at peak and AT did not correlate to peak $\dot{V}_O_2$. This is the most discrepant of variables but this lack of a relationship with peak $\dot{V}_O_2$ suggests that the magnitude of BR is not useful in predicting disease severity; it can only really act by determining if the lungs limit exercise through a threshold (for example 30% at peak) but the value generated gives us no further information, and probably has no relevance in serial tests. It is possible that the normal BR values seen in cardiac patients obscure a potential relationship with disease severity in COPD patients, however the correlation with peak $\dot{V}_O_2$ was not significant even when only COPD patients were considered. Almost all other CPX variables I have listed, with the exception of the $HR - \dot{V}_O_2$ slope, correlated to peak $\dot{V}_O_2$. For the various measures of the $\dot{V}_E$/CO, relationship
it was, as expected, an inverse correlation. Disease category affects the relationship between OUES and peak $\dot{V}_{O_2}$, but not most of the remaining CPX variables. Importantly OUES is much more closely correlated to peak $\dot{V}_{O_2}$ within the cardiac patients.

### 7.5.4 Variables relating to disease severity

Overall certain variables appear to display the necessary properties that would be consistent with reflecting disease severity. OUES reflected symptom scores in patients with cardiac disease, but not COPD, correlated closely with tissue Doppler measurements of left ventricular systolic and diastolic function, LA volume, right ventricular systolic function and BNP, did not relate to resting spirometry and diffusion capacity, and reflected exercise capacity (as measured by peak $\dot{V}_{O_2}$) in cardiac but not respiratory patients. The $\dot{V}_{O_2} - WR$ slope also related to NYHA class in cardiac patients, was correlated to diastolic although not systolic markers of cardiac dysfunction, and closely correlated with BNP. Within the cardiac patients it was closely correlated to peak $\dot{V}_{O_2}$/kg and percent of predicted peak $\dot{V}_{O_2}$. Similarly to the OUES, the $\dot{V}_E/\dot{V}_{CO_2}$ slope improved as FEV$_1$ worsened, however unlike OUES a decline in K$_{CO}$ was reflected by a rise (worsening) of the slope. All measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship showed a strong correlation with peak $\dot{V}_{O_2}$ (it is an inverse correlation but this would be expected given that a worsening slope rises, not falls) and a strong positive correlation with BNP in the cardiac group. $\dot{V}_E/\dot{V}_{CO_2}$ slope did not correlate strongly with any of the echocardiographic variables except TAPSE, but did strongly correlate to symptom score within the cardiac group. The $\dot{V}_E/\dot{V}_{CO_2}$ ratio behaved slightly differently in that when measured at the AT it also appeared to correlate to symptom score within the COPD group, and the ratio at nadir correlated to peak $\dot{V}_{O_2}$ within the COPD population. DP and circulatory power correlated to RV systolic function and BNP; DP also correlated with LV systolic function and circulatory power with symptom class. The OUEP and O$_2$ pulse correlated to peak $\dot{V}_{O_2}$ and to symptom score in both groups and so do not appear specifically related to symptom severity in heart failure. Breathing reserve at the AT correlated with symptom status in the COPD group, but not to peak $\dot{V}_{O_2}$. BR at peak exercise was correlated to BNP and ejection fraction in the full group of patients, however this was an inverse relationship, i.e. higher breathing reserves were correlated to higher BNP levels and reduced EF, suggesting another example of Simpson’s paradox due to the conflicting nature of the two diseases on the 2 variables being correlated. This hypothesis seems likely as BR at peak was not correlated to EF or BNP when analysed within respiratory and cardiac disease separately.
7.5.5 Limitations

The non-specific use of NYHA class in this study could be criticised. Firstly it is for use in heart failure, not COPD, but I have applied the same scoring system for my COPD patients. That was because the MRC dyspnoea score, although validated in COPD patients, runs on an ordinal scale from 1-5, so would statistically be very difficult to compare to a scale running from 1-4. The symptoms that help physicians categorise a heart failure patient with the NYHA classification are just as appropriate for a respiratory cause of breathlessness. Therefore I elected to use it for all patients; this decision was unfortunately made after the first few COPD patients had undergone the study protocol. A concern about NYHA class is the difficulty distinguishing between class II and III; confusion between these two classes leads to most of the poor reproducibility in this classification system. Within this analysis 82% of patients were within class II or III, which may restrict the classification system’s ability within my analyses.

Within chapter 6 I analyse the relationship between FEV$_1$ and CPX variables using multivariate regression. Within this chapter I use correlation with echocardiographic variables. That choice was made because regression implies direction and causation. It is conceptually likely that the obstructive abnormalities of respiration found in COPD patients cause the reduced exercise capacity and symptoms (although I accept COPD is far more complex than just obstructive spirometry). However abnormalities on echocardiography and BNP, for example, are features of heart failure, the same as exercise intolerance. Therefore there is no direction or causation and correlation is appropriate. The variables correlated against CPX variables, NYHA class, echocardiographic abnormalities and BNP, are largely unaffected by age, gender and weight, thus removing the need for a multivariate model.

7.5.6 Conclusion

OUES and then the $\dot{V}_{O_2} - WR$ slope appear to closely reflect disease severity, as measured by multiple variables, in cardiac, but not respiratory, patients. Other variables perform less well, and some such as the peak $\dot{V}_{O_2}$, OUEP and $O_2$ pulse, probably reflect disease severity in both conditions. The breathing reserve, whilst a good discriminator, does not reflect disease severity of cardiac or respiratory disease.
8.0 The Interventional Study – The Mitral Valve
8.1 Abstract

Certain cardiopulmonary exercise test (CPX) variables appear to show discriminant properties in identifying cardiac limitation. If these variables are specific for cardiac function then they should show rapid improvements following restoration of normal cardiac function following mitral valve surgery for severe mitral regurgitation or stenosis. Non-specific variables may deteriorate initially due to the long process of postoperative recovery and the impact of the surgery on the lungs. Minimally invasive mitral repair should show greater improvements in all CPX variables when compared to open repair. However these hypotheses have never been investigated.

21 patients (16 males, 5 females) with severe mitral regurgitation (n=20) or stenosis (n=1), were recruited prior to surgery. They underwent CPX testing on a bicycle ergometer. Following a familiarisation test, each patient underwent a personalised second test aiming for maximal exercise after approximately 10 minutes. This test was repeated approximately 2 months after surgery and again 6 months after surgery in 9 patients. Postoperative changes at 2 months were noted in FEV$\_1$ and total lung capacity. Transfer coefficient significantly deteriorated postoperatively in patients undergoing open repair, but not in those undergoing minimally invasive surgery.

Despite symptomatic improvement peak $\dot{V}_{O_2}$ was reduced at 2 months, and similar to baseline at 6 months when compared to baseline. $\dot{V}_{O_2}$ at the AT was reduced to a greater degree postoperatively than other variables and was not back to baseline at 6 months. OUES showed a borderline reduction at 2 months. Measures of the $\dot{V}_{E}/\dot{V}_{CO_2}$ relationship all showed non-significant trends to early (2 months) and sustained (6 months) improvement. Operation type, residual regurgitation and alterations to medication did not affect the degree of change in the majority of CPX variables.

Following mitral valve repair/replacement many variables take time to recover. Peak $\dot{V}_{O_2}$ and $\dot{V}_{O_2}$ at the AT deteriorated significantly after surgery, OUES less so. Measures of the relationship between minute ventilation and $CO\_2$ elimination showed a trend to early and sustained improvements.
8.2 Introduction

In previous chapters I have shown that certain variables appear to have increased sensitivity and specificity to detect abnormal cardiac physiology over others; importantly whilst remaining relatively resilient to abnormal respiratory function. OUES stands out as a variable with excellent reproducibility, excellent discriminant ability and with minimal correlation to abnormal spirometry. However, as defined within the aims of the study, an ideal variable will also appropriately reflect changes to a patient’s cardiac function.

This is a difficult criterion to investigate. It is not ethically acceptable to perform interventions on patients or healthy controls which may lead to permanent reductions in their cardiac function. It may be possible to perform acute, reversible interventions on specific populations of heart failure patients; this has been tested using patients with left ventricular assist devices with a transient reduction in pump speed (Jakovljevic et al 2010, Noor et al 2012). Ethically it is easier to perform studies on patients undergoing interventions to improve cardiac function; there are also common examples of this performed regularly in clinical practice.

Mitral valve reparative surgery is one such intervention. It is a relatively good model of an intervention that should acutely and persistently improve cardiac output and intra-cardiac pressures. Generally it is associated (so long as surgery is performed at an appropriate time) with good left ventricular function so recovery can be anticipated to be near absolute. Patients are often younger, in the degenerative form (where repair is commonly attempted) it is often an isolated disease and, because unwanted risk factors such as smoking and hypertension are not causative, typically patients will have no predisposition to coronary artery disease or respiratory abnormalities. Patients with mitral valve disease may be relatively asymptomatic (Messika-Zeitoun et al 2006), but even these patients can show improvements in their exercise capacity from surgery (Madaric et al 2007).

However there are limitations to this model, such as it doesn’t truly isolate improvements to the circulatory system without impacting on other organs, and despite amelioration of the regurgitation or stenosis of the mitral valve potentially permanent microscopic changes to the left ventricle and pulmonary vasculature may have occurred along with macroscopic abnormalities of the left atrium which predispose to atrial arrhythmias. The procedure may impact on other organs because the operation requires significant chest trauma (typically a midline sternotomy), and time on cardiopulmonary bypass and mechanical ventilation, which can all affect the lungs. Medium and long-term changes in respiratory function have been shown following coronary artery bypass surgery (Staton et al 2005) and heart transplantation (Ewert et al 1999).
Recovery time back to normal activity is considerable in patients undergoing mitral valve surgery and this, alongside reduced calorific intake (Nizami et al 1990), may lead to significant skeletal muscle loss. As discussed previously the muscle and heart have a complex interaction, and it is possible that gains to cardiac function could be offset by muscular changes associated with a post-surgical recovery state. This would delay the time taken to observe a benefit in exercise capacity. Therefore newer techniques, including minimally invasive mitral valve repair, are potentially better models as thoracic wall trauma is minimised, and overall recovery time is reduced potentially leading to less skeletal muscle atrophy.

Surgical correction of mitral valve disease has been shown to lead to rapid improvement of exercise capacity (Madaric et al 2007), although this response doesn’t appear uniform with significant time taken in other studies (Le Tourneau et al 2000, Kim et al 2004). In the former study improvements were seen at 4 months, however all patients had undergone minimally invasive surgery unlike the latter two studies.

No studies of CPX testing before or after surgical intervention measured more than a few variables. No statistical tests were used to compare if certain variables were affected to a greater or lesser extent.

Therefore I aim to show, in patients undergoing mitral valve surgery, whether CPX variables are differentially affected and to establish if surgical factors, including the use of minimally invasive surgery, can influence the change in these variables postoperatively.
8.3 Methods

8.3.1 Patient recruitment
Patients were recruited from Imperial College Healthcare NHS Trust (ICHNT) whilst awaiting surgical correction of severe mitral valve disease. They had all seen one of two surgeons who routinely performed mitral valve repairs and replacements at ICHNT and were involved in this research study. Potential patients were identified by a member of the cardiology or surgical team, and the patient was contacted asking them to be involved in the study. The decision for surgery was made by the patients’ own clinical teams; neither myself nor the results of this study influenced the decision to operate. Patients with either mitral regurgitation or stenosis were recruited, however I preferentially aimed to recruit patients undergoing repair, as this may be more physiological than a replacement valve. One of the two surgeons involved also performed some repairs via a minimally invasive access route (right minithoracotomy, typically through the 3rd or 4th intercostal space).

8.3.2 Patient testing preoperatively
Once patients were identified and consented, they underwent full lung function tests, a transthoracic echocardiogram, venous blood tests and 2 cardiopulmonary exercise tests; these are considered the baseline tests. The full description of these tests can be found within the Methods Chapter. The first test was for familiarisation and to identify the optimal protocol for each patient to achieve similar duration of exercise in all (approximately 10 minutes of incremental exercise). The results of this first CPX test were not considered. The results of the second test, performed at least 2 hours after the first test to allow for muscular recovery, were performed on the patient’s optimal protocol and were retained for analysis.

CPX tests were performed using a bicycle ergometer (Ergoline, GmbH, Bitz, Baden-Württemberg, Germany) on a COSMED Quark CPET System (COSMED S.r.l. Rome, Italy). The identification and measurement of CPX variables was performed as per directions from within the Methods Chapter and will not be discussed again here. However following on from the identification of a new set of contemporary reference equations (Gläser et al 2010, Chapter 4) percent of predicted values for peak $\dot{V}_{\text{O}_2}$, OUES and $\text{O}_2$ pulse were calculated using these SHIP equations described previously. The $\dot{V}_{\text{O}_2}$ at the AT will be presented as a percent of the predicted peak $\dot{V}_{\text{O}_2}$ using the SHIP reference equation for peak $\dot{V}_{\text{O}_2}$. All other variables will just be displayed giving their raw values. The $\dot{V}_E/\dot{V}_{\text{CO}_2}$ slope is displayed in 2 forms: slope 1 using data until the VCP; and slope 2 using data...
throughout exercise, in keeping with previous chapters, the $\dot{V}_E/\dot{V}_{CO_2}$ slope minus ratio difference was also calculated for each patient as per previous chapters.

### 8.3.3 Patient testing postoperatively

There was no involvement from myself during the perioperative and immediate postoperative period.

Patients were contacted over a month after their operation to gauge recovery and arrange follow-up. After discussion with the two surgeons involved it was agreed that 6 weeks should allow for enough recovery to make exercise testing safe; this was the minimum time from operation to follow-up for the study. On arrival patients underwent venous blood sampling, transthoracic echocardiography and full lung function testing, in the same manner as preoperatively. They each then underwent a single CPX test on the same protocol as the main test from the preoperative testing (i.e. not the familiarisation test).

Patients were then re-contacted at 5 months to arrange a further CPX test (on an identical protocol to the previous 2) at 6 months following surgery. Other tests were not repeated at this point, only CPX.

### 8.3.4 Echocardiography

Whilst transthoracic echocardiography is covered in detail within the Methods Chapter, there are specific measurements made for the mitral valve patients that warrant special mention. Wherever possible, quantitative measurements of regurgitant severity were attempted. This typically involves the use of the Proximal Isovelocity Surface Area (PISA) method, first described over 20 years ago (Recusani et al 1991). Full methodology for how I performed PISA is described in the Methods Chapter but briefly a hemisphere of colour Doppler on the ventricular aspect of the mitral valve leaflet tips occurs during systole in patients with MR. The size can be manipulated by altering the aliasing velocity in the direction towards regurgitant flow. Once a reasonable sized hemisphere is noted the radius of this hemisphere is measured, and recorded alongside the aliasing velocity. Continuous wave Doppler across the mitral valve generates a velocity-time integral (VTI) and peak velocity. These measurements allow for calculation of a regurgitant volume and effective orifice area of regurgitation (EROA). If forward stroke volume is known regurgitant fraction can also be calculated.

A semi-quantitative method of MR assessment is jet: left atrial area ratio. In the apical 4 chamber view, following optimisation of gain and depth settings, the frame is paused where the largest regurgitant jet area can
be seen. The area of the jet within the left atrium is measured (excluding the PISA on the ventricular aspect of the valve and any flow in the pulmonary veins) alongside the area of the left atrium; the ratio of these 2 areas is then calculated.

### 8.3.5 Hypotheses

I hypothesised that at 2 months we would see either no improvement or a deterioration in variables of exercise testing, but at 6 months we would see values superior to preoperative levels. I hypothesised that variables that are more reflective of cardiac function (based on the results of preceding chapters), for example OUES, would show a lesser reduction at 2 months compared with more non-specific variables, for example peak $\dot{V}_{O_2}$, which would be more negatively affected by the total body adaptations post-surgery (such as prolonged, enforced rest and recovery, reduced nutritional intake etc). I also hypothesised that the process of surgery would lead to some changes in respiratory function and that these would correlate to the reduction in exercise capacity at 2 months in non-specific variables, but not cardio-specific variables such as OUES.

### 8.3.6 Statistical analysis

Variables of gas analysis and lung function have previously been assessed for normal distribution. Normally distributed variables are displayed as mean ± standard deviation; non-normally distributed variables as median (25th, 75th percentiles). For preoperative versus postoperative changes of linear variables paired t-tests were used, unless the variable was non-normally distributed, in which case a Wilcoxon rank test was employed. For graphical representation of the change between baseline and 2 and 6 months each variable was standardised using the z-score method ($z$-score = change/standard deviation of variables within population). On these plots a leftward shift represents a worsening of a variable, regardless of whether this is an increase or decrease in the magnitude of the variable. Because z-scores standardise each variable these scores were directly compared, paired t-tests were employed to compare each variable against one another to see if there was a significant difference in the change between tests.

I analysed potential confounders for their influence on the change in a variable in 2 ways. Firstly patients were grouped by type of surgery (repair, minimally invasive repair, valve replacement), access site (sternotomy or thoracotomy), severity of residual regurgitation (none/trivial, mild, moderate, severe), surgeon, rhythm at baseline and rhythm postoperatively, and alterations in medications grouped by an increase, decrease or no
change in rate-control drugs and an increase, decrease or no change in anti-hypertensive medication. Each of these sets of groups were analysed against the change in each variable (the deltas) using an ANOVA model and confirmed using an interaction mixed linear model (interaction between change in variable and potential confounder).

Secondly for the continuous variables such as change in lung function, and length of time from surgery to follow-up, a multivariate regression model was employed. Change in FEV₁, and either change in D_{LCO} or K_{CO} (Hb) were used as measures of a change in lung function. The influence on the change in each CPX variable was assessed using this multivariate model.

A p value <0.05 was considered significant throughout.
8.4 Results

8.4.1 Patient recruitment and characteristics

26 patients were recruited with mitral valve disease requiring surgical correction. Of these 26, 1 died perioperatively, 1 had a postoperative stroke and 1 suffered multiple medical co-morbidities making follow-up impractical. A further 2 patients had delays to surgery (awaiting dental work) so that they still had not received the surgery by this analysis.

The study group was therefore on 21 patients (mean age 63.1 ± 9.3 years; 16 males, 5 females). All of these 21 patients had follow-up at approximately 2 months (median 61 days, IQR 53-71 days), with a further 9 having a second follow-up at approximately 6 months (median 211 days, IQR 188-233 days). Patient and surgical characteristics are shown in Table 8.1. Patient 5 was the only patient undergoing surgery for mitral stenosis; the others were all for mitral regurgitation. Patient 21 was the only patient with diabetes mellitus (Type 2).

Preoperatively, on transthoracic echocardiography, mean ejection fraction was 61.1 ± 9.6 %, 12 patients had an ejection fraction <60%, no patients had an ejection fraction ≤ 30 %. Left ventricular diastolic dimension was 55.5 ± 4.9 mm, systolic dimension was 37.7 ± 6.1 mm and left atrial volume was 150 ± 79 cm$^3$. Only 2 patients had a left ventricular systolic dimension ≥ 45 mm. Whilst all patients had echocardiography with internal dimension measurements, not all patients were able to have quantitative analysis of their mitral regurgitation severity, due to eccentric jets. All patients had previously undergone transoesophageal echocardiography; it was this test, rather than transthoracic echocardiography, that had ultimately made the decision for surgery. On 17 patients however I was able to make PISA measurements of their MR severity; effective orifice area (EROA) was 0.38 (0.27, 0.89) cm$^2$; regurgitant fraction was 51.4 (35.1, 67.0) %; and regurgitant volume was 58.5 (37.1, 108.1) mL. Mean jet area: left atrial area ratio was 42.7 ± 15.9 %. The patient with mitral stenosis had a mitral valve area of 0.99 cm$^2$ by pressure half time, and a mean transmitial pressure gradient of 11.3 mmHg.

12 patients underwent a traditional repair with midline sternotomy, 4 patients underwent repair via a minimally invasive approach, 3 patients underwent a valve replacement with a tissue valve, and 2 with a tilting disk metal valve. Largely they were uncomplicated, although one needed a repeat procedure 2 days after his first operation, and another patient had significant haemolysis requiring repeat transfusions, and eventually a repeat operation. The postoperative study haemoglobin for this patient was 6.6 g/dL; in all other patients it was 11.2 g/dL or higher. The results will be considered for the full group of 21, but then also with this single patient excluded, because of the potential influence of anaemia on exercise capacity.
<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age (yrs)</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
<th>NYHA class</th>
<th>Smoking status</th>
<th>Baseline rhythm</th>
<th>Peak $V_{O2}$ (%) predicted</th>
<th>Protocol (W/min)</th>
<th>Procedure performed</th>
<th>Rhythm post-op</th>
<th>MR grade</th>
<th>Significant complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>50</td>
<td>188</td>
<td>91</td>
<td>25.7</td>
<td>2</td>
<td>Never</td>
<td>SR</td>
<td>76.8</td>
<td>20</td>
<td>Repair</td>
<td>SR</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>52</td>
<td>174</td>
<td>72</td>
<td>23.6</td>
<td>1</td>
<td>Never</td>
<td>SR</td>
<td>87.0</td>
<td>20</td>
<td>Repair (MI)</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>63</td>
<td>158</td>
<td>88</td>
<td>35.3</td>
<td>3</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>98.6</td>
<td>12</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>63</td>
<td>175</td>
<td>82</td>
<td>26.8</td>
<td>2</td>
<td>Never</td>
<td>AF</td>
<td>64.2</td>
<td>12</td>
<td>Repair</td>
<td>AF</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>76</td>
<td>161</td>
<td>47</td>
<td>18.1</td>
<td>2</td>
<td>Never</td>
<td>AF</td>
<td>75.0</td>
<td>8</td>
<td>Tissue valve</td>
<td>AF</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>76</td>
<td>163</td>
<td>67</td>
<td>25.2</td>
<td>1</td>
<td>Never</td>
<td>SR</td>
<td>76.0</td>
<td>10</td>
<td>Repair</td>
<td>SR</td>
<td>3</td>
<td>Requiring transfusions</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>66</td>
<td>170</td>
<td>62</td>
<td>21.5</td>
<td>2</td>
<td>Ex-smoker</td>
<td>AF</td>
<td>52.1</td>
<td>10</td>
<td>Tissue valve</td>
<td>AF</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>63</td>
<td>187</td>
<td>90</td>
<td>25.7</td>
<td>1</td>
<td>Never</td>
<td>AF</td>
<td>77.3</td>
<td>15</td>
<td>Tissue valve</td>
<td>AF</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>76</td>
<td>185</td>
<td>100</td>
<td>29.1</td>
<td>2</td>
<td>Ex-smoker</td>
<td>AF</td>
<td>77.9</td>
<td>12</td>
<td>Repair (MI)</td>
<td>AF</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>50</td>
<td>193</td>
<td>85</td>
<td>22.8</td>
<td>3</td>
<td>Never</td>
<td>SR</td>
<td>27.9</td>
<td>15</td>
<td>Repair</td>
<td>Flutter</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>59</td>
<td>188</td>
<td>81</td>
<td>22.9</td>
<td>2</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>86.6</td>
<td>20</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>58</td>
<td>188</td>
<td>91</td>
<td>25.7</td>
<td>2</td>
<td>Ex-smoker</td>
<td>AF</td>
<td>69.0</td>
<td>15</td>
<td>Tilting disk</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>65</td>
<td>182</td>
<td>78</td>
<td>23.5</td>
<td>1</td>
<td>Never</td>
<td>SR</td>
<td>71.2</td>
<td>12</td>
<td>Repair (MI)</td>
<td>SR</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>69</td>
<td>172</td>
<td>62</td>
<td>21</td>
<td>2</td>
<td>Never</td>
<td>SR</td>
<td>73.9</td>
<td>15</td>
<td>Repair</td>
<td>SR</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>44</td>
<td>175</td>
<td>85</td>
<td>27.8</td>
<td>2</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>74.9</td>
<td>15</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>73</td>
<td>167</td>
<td>77</td>
<td>27.6</td>
<td>1</td>
<td>Never</td>
<td>SR</td>
<td>84.9</td>
<td>12</td>
<td>Repair (MI)</td>
<td>SR</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>68</td>
<td>157</td>
<td>64</td>
<td>26</td>
<td>3</td>
<td>Ex-smoker</td>
<td>AF</td>
<td>65.5</td>
<td>8</td>
<td>Repair</td>
<td>AF</td>
<td>0</td>
<td>Recurrent chest infections</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>54</td>
<td>171</td>
<td>60</td>
<td>20.5</td>
<td>1</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>101.3</td>
<td>15</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>F</td>
<td>71</td>
<td>164</td>
<td>71</td>
<td>26.4</td>
<td>1</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>97.4</td>
<td>10</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>68</td>
<td>160</td>
<td>61</td>
<td>23.8</td>
<td>3</td>
<td>Ex-smoker</td>
<td>SR</td>
<td>71.2</td>
<td>8</td>
<td>Repair</td>
<td>SR</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>62</td>
<td>169</td>
<td>77</td>
<td>26.9</td>
<td>2</td>
<td>Never</td>
<td>AF</td>
<td>61.9</td>
<td>12</td>
<td>Tilting disk</td>
<td>AF</td>
<td>0</td>
<td>Repeat surgery 2 days later</td>
</tr>
</tbody>
</table>

Table 8.1: Patient baseline and operative characteristics. 21 patients had undergone mitral valve surgery for severe regurgitation (n=20) or stenosis (n=1). NYHA=New York Heart Association; SR=sinus rhythm; AF=atrial fibrillation; MI=minimally invasive; MR=mitral regurgitation.
8.4.2 Impact of surgery on respiratory function

All patients underwent full lung function testing on recruitment prior to surgery, and then again after surgery on the same day as the 2 month CPX was performed. The pre and post results for some of the important lung function variables can be seen in Table 8.2. FEV₁ as an absolute value and percent predicted deteriorated postoperatively (-0.15 L and -5.9 % respectively). The FVC deteriorated to a similar extent (-0.23 L and -7.1 % respectively), and the FEV₁:FVC ratio was not significantly altered. Total lung capacity significantly deteriorated by quite a large volume (-0.66 L and -11.1 %), and the measures of diffusion capacity all went down (corrected for predicted values, haemoglobin and for alveolar volume – the transfer coefficient). Although the numbers are small for patients undergoing minimally invasive surgery (n=4) they did not have a significantly reduced diffusion capacity unlike patients undergoing open repair (n=12) (ΔK CO (Hb) = -0.02 in minimally invasive patients, p=0.72; -0.16 in open repair patients, p=0.002).

8.4.3 Other changes postoperatively and echocardiographic findings

There was a significant fall in haemoglobin postoperatively (mean change -1.5 g/dL, p<0.001), which was still significant even after exclusion of the single patient with haemolysis (mean change -1.3 g/dL, p<0.001). There were no significant changes in creatinine, eGFR, sodium, bicarbonate or BNP.

1 patient was graded to have severe residual regurgitation following the operation. This was the same patient who had haemolysis and soon after he underwent a repeat operation. 2 patients had the appearances of moderate mitral regurgitation postoperatively, but no patients had restrictive opening of the mitral valve. There was a significant reduction in left ventricular diastolic dimension (-7.1 mm, p<0.001) and diastolic volume (-27.8 cm³, p=0.001), but not systolic dimension (-1.4 mm, p=0.28) or volume (+1.9 cm³, p=0.67). Ejection fraction but not left atrial volume fell significantly (-9.9 %, p<0.001, -21.4 cm³, p=0.13 respectively).

1 patient reverted to sinus rhythm from their preoperative rhythm of AF, whilst another changed from sinus rhythm to atrial flutter. All other patients retained the rhythm of their baseline tests.

Largely symptoms were non-significantly improved at 2 months with an average NYHA class 1.9 ± 0.7 preoperatively and 1.6 ± 0.7 postoperatively, with 10 not changing class, 8 improving and 3 worsening.
<table>
<thead>
<tr>
<th>Lung function variable</th>
<th>Preoperative (mean ± sd)</th>
<th>Postoperative (mean ± sd)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV₁ (L)</td>
<td>2.49 ± 0.70</td>
<td>2.34 ± 0.75</td>
<td>0.07</td>
</tr>
<tr>
<td>FEV₁ (% predicted)</td>
<td>85 ± 20</td>
<td>79 ± 20</td>
<td>0.02</td>
</tr>
<tr>
<td>FVC (L)</td>
<td>3.63 ± 1.00</td>
<td>3.40 ± 1.04</td>
<td>0.03</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>99 ± 24</td>
<td>92 ± 23</td>
<td>0.01</td>
</tr>
<tr>
<td>FEV₁:FVC ratio</td>
<td>69 ± 10</td>
<td>69 ± 10</td>
<td>0.82</td>
</tr>
<tr>
<td>PEF</td>
<td>7.77 ± 1.96</td>
<td>7.41 ± 2.28</td>
<td>0.15</td>
</tr>
<tr>
<td>TLC</td>
<td>5.92 ± 1.67</td>
<td>5.26 ± 1.57</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>RV</td>
<td>2.19 ± 0.92</td>
<td>1.94 ± 0.54</td>
<td>0.09</td>
</tr>
<tr>
<td>FRC</td>
<td>3.35 ± 0.99</td>
<td>3.10 ± 0.76</td>
<td>0.08</td>
</tr>
<tr>
<td>Dlco</td>
<td>6.90 ± 2.07</td>
<td>5.94 ± 2.05</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Dlco (% predicted)</td>
<td>77 ± 16</td>
<td>66 ± 16</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Dlco (Hb)</td>
<td>7.08 ± 2.18</td>
<td>6.23 ± 1.99</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>KCO</td>
<td>1.29 ± 0.23</td>
<td>1.17 ± 0.26</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>KCO (% predicted)</td>
<td>94 ± 16</td>
<td>86 ± 18</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>KCO (Hb)</td>
<td>1.32 ± 0.24</td>
<td>1.23 ± 0.24</td>
<td>0.01</td>
</tr>
<tr>
<td>V₄ₐ</td>
<td>5.39 ± 1.43</td>
<td>5.13 ± 1.38</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 8.2: Comparison of lung function variables before and after mitral surgery. FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity; PEF = peak expiratory flow; TLC = total lung capacity; RV = residual volume; FRC = function residual capacity; Dlco = Diffusing capacity of the lung for carbon monoxide; Hb = haemoglobin; KCO = transfer coefficient; V₄ₐ = alveolar volume.
8.4.4 Change in CPX variables at 2 months

All 21 patients had a CPX performed postoperatively at approximately 2 months (median 61 days, IQR 53-71 days); although in two patients they did not have their CPX test until at least 4 months had passed (Table 8.3).

Generally CPX variables worsened when comparing with preoperative values (Figure 8.1). \( \dot{V}_{O_2} \) at the AT (p<0.001), peak \( \dot{V}_{O_2} \) (p<0.001 all measures), double product and circulatory power were the most significantly reduced. Peak \( \dot{V}_{O_2} \) was reduced by a mean value of 2.1 mL/min/kg, or 9 % predicted. OUES showed a small, but significant, drop between the 2 tests (-0.14, p=0.047; -6.3% predicted, p=0.04) but not when adjusted for weight (p=0.10). The OUEP (p=0.01) and peak minute ventilation (p=0.004), but not breathing reserve, were also significantly lower on test 2. Mean peak watts attained was lower during the second test (106 W vs 115 W); when combined with the reduction in peak \( \dot{V}_{O_2} \) this led to a non-significantly lower \( \dot{V}_{O_2} - WR \) slope (-0.66, p=0.065). The \( O_2 \) pulse did not change. \( \dot{V}_E/\dot{V}_{CO_2} \) slope throughout exercise and until the VCP, and the \( \dot{V}_E/\dot{V}_{CO_2} \) ratio at nadir and AT all showed a non-significant trend to lower, improved values at the second test. The \( HR - \dot{V}_{O_2} \) slope was significantly flatter on the second test (p=0.02 by Wilcoxon rank test).

Following the removal of the single patient with severe postoperative mitral regurgitation and severe anaemia the majority of the results did not significantly alter. However OUES and percent of predicted OUES were no longer significantly different between the 2 tests (-0.12, p=0.08; -5.4% predicted, p=0.07 respectively). The \( \dot{V}_E/\dot{V}_{CO_2} \) slope throughout the whole test became significantly lower than baseline after the removal of that single patient (-3.3, p=0.04 by Wilcoxon rank test).

The z-scores of certain variables were compared against one another. Peak \( \dot{V}_{O_2} \) (mL/min/kg) showed a significantly smaller reduction between baseline and 2 months than \( \dot{V}_{O_2} \) at the AT (p=0.025). However peak \( \dot{V}_{O_2} \) showed a significantly greater deterioration at 2 months when compared to OUES (p=0.002), the \( O_2 \) pulse (p<0.01), \( \dot{V}_E/\dot{V}_{CO_2} \) slope (p=0.0001) and ratio (p<0.001).
<table>
<thead>
<tr>
<th></th>
<th>Baseline test (n=21)</th>
<th>Δ at 2 month test (n=21)</th>
<th>Δ at 6 month test (n=9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak $V_{O_2}$ (mL/min)</td>
<td>1472 ± 425</td>
<td>-175 ± 202*</td>
<td>1 ± 254‡</td>
</tr>
<tr>
<td>Peak $V_{O_2}$/kg</td>
<td>19.4 ± 4.5</td>
<td>-2.1 ± 2.5*</td>
<td>0.2 ± 3.1‡</td>
</tr>
<tr>
<td>Peak $V_{O_2}$ (% predicted)</td>
<td>74.8 ± 16.3</td>
<td>-9.0 ± 9.5*</td>
<td>0.4 ± 10.1‡</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>1022 ± 287</td>
<td>-173 ± 153*</td>
<td>-81 ± 282‡</td>
</tr>
<tr>
<td>AT (% of pred peak $V_{O_2}$)</td>
<td>52.3 ± 12</td>
<td>-9.3 ± 7.2*</td>
<td>-2.6 ± 10.9‡</td>
</tr>
<tr>
<td>OUES</td>
<td>1.65 ± 0.64</td>
<td>-0.14 ± 0.30*</td>
<td>-0.02 ± 0.52</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>21.7 ± 7.3</td>
<td>-1.5 ± 3.9</td>
<td>0.26 ± 6.43</td>
</tr>
<tr>
<td>OUES (% predicted)</td>
<td>76.7 ± 23.8</td>
<td>-6.3 ± 13.1*</td>
<td>0.8 ± 21.4</td>
</tr>
<tr>
<td>OUEP</td>
<td>35.4 ± 4.1</td>
<td>-1.8 ± 3.02*</td>
<td>0.73 ± 4.22</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat)</td>
<td>10.8 ± 3.0</td>
<td>0.22 ± 2.46</td>
<td>0.44 ± 1.24</td>
</tr>
<tr>
<td>$O_2$ Pulse (% predicted)</td>
<td>76.5 ± 20.0</td>
<td>0.9 ± 17.7</td>
<td>4.0 ± 8.1</td>
</tr>
<tr>
<td>$\dot{V}<em>{E}/\dot{V}</em>{CO_2}$ slope 1</td>
<td>33 (28.3, 36.1)</td>
<td>-0.82 ± 6.47</td>
<td>-3.87 ± 12.12</td>
</tr>
<tr>
<td>$\dot{V}<em>{E}/\dot{V}</em>{CO_2}$ slope 2</td>
<td>35.9 (32.4, 39.9)</td>
<td>-2.49 ± 8.09</td>
<td>-5.60 ± 12.16</td>
</tr>
<tr>
<td>$\dot{V}<em>{E}/\dot{V}</em>{CO_2}$ ratio nadir</td>
<td>31.5 (29.9, 34.2)</td>
<td>-0.41 ± 4.1</td>
<td>-2.32 ± 5.99‡</td>
</tr>
<tr>
<td>$\dot{V}<em>{E}/\dot{V}</em>{CO_2}$ ratio AT</td>
<td>32.6 (31, 36.8)</td>
<td>-0.24 ± 3.9</td>
<td>-1.66 ± 6.13‡</td>
</tr>
<tr>
<td>$\dot{V}<em>{E}/\dot{V}</em>{CO_2}$ slope - ratio</td>
<td>-0.4 (-1.9, 1)</td>
<td>0.4 ± 3.54</td>
<td>1.55 ± 6.46</td>
</tr>
<tr>
<td>RER at peak</td>
<td>1.15 ± 0.11</td>
<td>0.06 ± 0.11*</td>
<td>0.04 ± 0.11</td>
</tr>
<tr>
<td>PetCO$_2$ at AT (mmHg)</td>
<td>36 ± 3.9</td>
<td>0.48 ± 3.37</td>
<td>2.11 ± 4.78‡</td>
</tr>
<tr>
<td>HR at peak (bpm)</td>
<td>143 ± 26</td>
<td>-19 ± 26*</td>
<td>0 ± 33‡</td>
</tr>
<tr>
<td>DP (mmHg bpm)</td>
<td>23699 ± 7151</td>
<td>-3978 ± 6904*</td>
<td>-827 ± 6302‡</td>
</tr>
<tr>
<td>Peak circulatory power</td>
<td>251789 ± 111271</td>
<td>-42280 ± 73837*</td>
<td>-13393 ± 70314‡</td>
</tr>
<tr>
<td>Peak oxygen saturations (%)</td>
<td>97 (96, 98)</td>
<td>-0.63 ± 3.61</td>
<td>1.13 ± 2.17</td>
</tr>
<tr>
<td>Breathing reserve at AT (%)</td>
<td>69.3 (63.3, 77.9)</td>
<td>1.26 ± 5.59</td>
<td>0.32 ± 8.31</td>
</tr>
<tr>
<td>Breathing reserve (%)</td>
<td>36.9 (26.2, 42.3)</td>
<td>1.11 ± 15.1</td>
<td>-5.16 ± 22.3</td>
</tr>
<tr>
<td>Minute ventilation (L/min)</td>
<td>63.9 ± 16.8</td>
<td>-7.08 ± 10.05*</td>
<td>-1.15 ± 14.6</td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ slope</td>
<td>8.7 (8.2, 9.6)</td>
<td>-0.66 ± 1.55</td>
<td>0.02 ± 1.37</td>
</tr>
<tr>
<td>Peak work rate (W)</td>
<td>115 ± 41</td>
<td>-9 ± 17*</td>
<td>2 ± 15‡</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ slope</td>
<td>0.055 (0.043, 0.070)</td>
<td>-0.015 ± 0.024*</td>
<td>-0.011 ± 0.022</td>
</tr>
</tbody>
</table>

Table 8.3: CPX variables preoperatively and at an average of 2 and 6 months postoperatively in patients undergoing mitral valve surgery for severe mitral regurgitation (n=20) or stenosis (n=1). * p<0.05 baseline vs 2 months; † p<0.05 baseline vs 6 months; ‡ p<0.05 2 months vs 6 months
Figure 8.1: Change in 12 variables between a preoperative CPX and a CPX performed, on average, 2 months postoperatively. The change has been converted in a z-score to allow for comparison between variables. A deviation to the left indicates a worsening of the variable (so for the $\dot{V}_E/\dot{V}_{CO_2}$ slope the positive score indicates a decrease in magnitude).
8.4.5 Change in CPX variables at 6 months

9 patients came back for a 6 month CPX test. Table 8.3 shows the difference in the major variables between baseline and this 6 month time point. No variables were significantly different from baseline (Figure 8.2), however there remains a trend to an improvement in the $\dot{V}_E/\dot{V}_CO_2$ slope, which was first noted at 2 months.

It is interesting to see, although the numbers are too small to achieve statistical significance, that the $\dot{V}_O_2$ at the AT remained markedly reduced compared to other variables. The near zero difference in peak $\dot{V}_O_2$ suggests exercise capacity was back to preoperative levels. When 6 month tests were compared to 2 months, peak $\dot{V}_O_2$ had significantly improved but $\dot{V}_O_2$ at the AT, OUES, OUEP and $O_2$ pulse did not change. $\dot{V}_E/\dot{V}_CO_2$ ratio at nadir and AT had significantly decreased whilst the $P_{ET}CO_2$ at the AT had significantly increased. Peak heart rate, double product and circulatory power were significantly higher at 6 months than 2 months. A higher work rate was achieved but the $\dot{V}_O_2 - WR$ slope was not significantly altered.

8.4.6 Influencing factors on postoperative CPX variables

Surgical factors were assessed for their influence on CPX variables. Valve replacements had a significantly greater drop in OUEP compared to valve repairs ($p=0.008$). Minimally invasive surgery (n=4) did not significantly alter any results when compared with open repair and replacement. Ignoring the single patient with grade 3 residual MR, the presence of grade 1 or 2 regurgitation did not alter CPX variables significantly. Two surgeons operated, with no significant differences their results. Looking more closely at the patients with minimal access surgery, they actually displayed the largest decline in peak $\dot{V}_O_2$ (−223mL/min, −10.5%, compared with −188mL/min, −10.2% for open repair, and −105mL/min, −4.9% for replacement) although they did have the highest exercise capacity at baseline, so perhaps would expect to show the largest absolute difference.
Figure 8.2: Change in 12 variables between a preoperative CPX and a CPX performed, on average, 6 months postoperatively. The change has been converted in a z-score to allow for comparison between variables. A deviation to the left indicates a worsening of the variable (so for the $V_E/V_{CO_2}$ slope the positive score indicates a decrease in magnitude).
Only 2 patients had a change in rhythm between the 2 tests, one from sinus to atrial flutter, and one from atrial fibrillation to sinus. Analysing from baseline rhythm patients in AF had a borderline significantly smaller drop in OUES/kg and percent of predicted OUES, with a significantly greater deterioration in OUEP. However the patients with AF at recruitment had lower OUES values at baseline so a smaller decline after intervention would be expected. When assessing the impact of medication, I analysed for the addition of rate-control, no change to rate control and a reduction in rate-control medication. Unsurprisingly the postoperative changes in double product and peak heart rate varied to different degrees in these 3 groups. Patients with a reduction in rate-control showed a borderline smaller decline in OUES/kg than the other two groups (p=0.048); no other variables showed a significant influence from rate-control medications. When patients were grouped into more, less or the same amount of antihypertensive medication, in none of the variables was the change between tests affected.

A regression model including the change in FEV₁, DLCO and the length of time from surgery to testing showed significant relationships between the change in DLCO and peak VO₂, V̇O₂ at the AT, OUES, V̇E/V̇CO₂ slope and ratio, and P_{ET}CO₂ at the AT. However after correction for haemoglobin and alveolar volume, the K_{CO} (Hb) was not significantly related to the change in any CPX variables. The change in FEV₁ was related to the change in the OUEP, V̇E/V̇CO₂ ratio and the breathing reserve at peak. Length of time from operation to first follow-up CPX did not significantly affect the change in CPX variables.

**8.4.7 Changes in patients with reduced exercise capacity**

5 of the 21 patients had a normal exercise capacity preoperatively as defined by a peak V̇O₂ ≥ 85% predicted. I considered the group with reduced preoperative exercise capacity separately. In these remaining 16, peak V̇O₂ was reduced at 2 months by 146 ± 213 mL/min, 1.6 ± 2.6 mL/min/kg and 7.5% ± 10.3 (p<0.03 for all), which were all non-significantly smaller reductions than seen across the full group. OUES, OUES/kg and percent predicted OUES were all non-significantly reduced at 2 months compared with baseline in these 16 patients (-0.11, p=0.23; -0.95/kg, p=0.40; -4.4%, p=0.24 respectively), a smaller change than that seen across the full cohort. Double product, circulatory power, peak work rate and HR - V̇O₂ slope also displayed non-significantly smaller reductions at 2 months in patients with reduced exercise capacity compared to those within normal exercise capacity.
8.5 Discussion

I have shown that 2 months following mitral valve surgery for severe regurgitation or stenosis, there were large, significant declines in peak $\dot{V}_{O_2}$, anaerobic threshold, double product and circulatory power. There are significantly smaller reductions in OUES, which are almost non-significant from baseline. The ventilatory equivalents for CO$_2$ show an early and persistent, albeit non-significant, improvement following surgery. At 6 months most changes, with the principal exception of the $\dot{V}_{O_2}$ at the AT, have returned to close to the baseline values. At this time point $\dot{V}_{O_2}$ at the AT has a z-score of approximately -0.3, that is 0.3 standard deviations (of AT within the whole population) worse than baseline, although this does not reach significance. The $\dot{V}_E/\dot{V}_{CO_2}$ slope is almost +0.5 z-scores, but with each variable, the presence of only 9 patients so far returned for 6 month follow-up makes statistical significance difficult to obtain. It is also interesting to see that minimally invasive surgery does not show an improvement in exercise capacity, although there are suggestions that lung function may be improved.

8.5.1 Reductions in exercise capacity pre- and postoperatively

Organic mitral regurgitation is known to limit exercise capacity, but even in the presence of severe regurgitation it is not uncommon for patients to have preserved exercise tolerance (Messika-Zeitoun et al 2006). Conversely patients who believe themselves to be asymptomatic are often limited in exercise capacity compared with matched controls (Olaf et al 2012). 7 patients within my analysis were self-described as asymptomatic (NYHA class 1) whilst only 5 had a peak $\dot{V}_{O_2} > 85\%$, and only 3 patients met both these criteria (i.e. were both self-described as asymptomatic and had a near normal peak $\dot{V}_{O_2}$). A reduction in exercise capacity in mitral regurgitation is important as it is associated with increased clinical events such as death, heart failure and atrial arrhythmias (Messika-Zeitoun et al 2006), and also predicts a worse symptom class a year after surgery (Kim et al 2003). It would be expected that exercise capacity should improve following surgical correction of mitral regurgitation; however previous studies have shown vastly different results. A single study has shown significant improvements in peak $\dot{V}_{O_2}$ four months after minimally invasive mitral valve repair for regurgitation (Madaric et al 2007); interestingly improvements were most marked in the least symptomatic patients. In contrast two studies have shown no change in exercise capacity following open surgery. One study of patients undergoing open repair showed no overall difference in peak $\dot{V}_{O_2}$ at one year, with a decline if residual regurgitation persisted (Kim et al 2004). Another study, in which 40% of patients underwent replacement and
60% repair, showed no improvement in peak $\dot{V}_{O_2}$ after an average of 7 months (Le Tourneau et al 2000). It is difficult to know if the different modes of surgical correction, or underlying patient characteristics lead to such dramatically different results. Madaric et al did not recruit any patients with AF; either this or the minimally invasive nature of the surgery could explain the differences found. Whilst it is not the purpose of my study to detail the change in exercise capacity, purely to look at how various CPX variables behave differently, it’s interesting to note that my results show more similarities with the latter studies, rather than the former. At 2 months, albeit much earlier than the Madaric et al study, I show a significant deterioration in most variables. At 6 months they have returned to baseline and no better, although small numbers at this time point make it difficult to draw firm conclusions. Also in contrast to the Madaric et al study I find that it is probably the most symptomatic patients in my study who benefit the most. Could the differences be explained by their universal adoption of minimally invasive repair? Unfortunately it is impossible to draw any conclusions from my minimally invasive patients, as they number only 4.

What is interesting is that despite subjective opinion that symptoms are improved (0.3 improvement in NYHA class), objective measures of maximal exercise capacity do not agree.

8.5.2 Changes in CPX variables

It is to be expected that different variables will behave differently following surgery as they are all influenced to different degrees from the numerous processes that go on before, during and after surgery, such as improvement in cardiac function, potential deterioration in lung function from cardio-pulmonary bypass and mechanical ventilation, and the reduction in muscle bulk and power associated with a prolonged period of recovery. Madaric et al showed improvements in peak $\dot{V}_{O_2}$ and $O_2$ pulse, but not the AT. There have been a handful of studies following balloon valvuloplasty for mitral stenosis that have similarly shown a smaller improvement in the AT when compared to the improvement noted in the peak $\dot{V}_{O_2}$ (Douard et al 1997, Meurin et al 2005). Meanwhile Wright et al showed a reasonable improvement in both peak $\dot{V}_{O_2}$ (+2.5 mL/min/kg) and $\dot{V}_{E}/\dot{V}_{CO_2}$ slope (-4.5), although only the former reached statistical significance, with no mention of the AT (Wright et al 2003).

Evidence that the muscle plays a strong role in the delay of improvement in CPX variables comes from work by Yasu et al, who showed that despite an immediate improvement in symptoms following balloon valvuloplasty
for mitral stenosis it took a month until peak $\dot{V}_{O_2}$ improved. These prolonged improvements were closely correlated to muscular changes on magnetic resonance spectroscopy of the forearm, including intracellular pH on exercise and phosphocreatine recovery (Yasu et al 1996). In the study by Meurin et al an early exercise rehabilitation program following valve repair led to an increase in peak $\dot{V}_{O_2}$ (+22%) and a smaller increase in the AT (+16%) when compared before and after the program. However there was no control group, no tests performed before surgery, and tests were done on average at 3 and then 6 weeks postoperatively. This is a very early period postoperatively and we would expect to see large changes between these two time points; at 3 weeks patients have only recently become fully ambulant again, compared with 6 weeks when their leg muscles have had a much longer time to recover strength and bulk. When Douard et al compared two groups after balloon valvuloplasty, one following routine care, and the other undergoing an exercise rehabilitation program, the latter showed significant increases in peak $\dot{V}_{O_2}$ and work rate achieved (Douard et al 1997). AT non-significantly increased in the exercise program group by 17%, with only a 1% improvement in the control group; $\dot{V}_E/\dot{V}_{CO_2}$ slope was significantly decreased in both groups but to a greater extent in the exercise regime group.

The valvuloplasty model is relatively unique in that it can immediately rectify some of the abnormal physiology associated with mitral valve disease without the long recovery time seen with an open repair or replacement. Despite immediate restoration of valvular function (mean gradient fell from 12.7 to 3.4 in the control arm of the Douard et al study) anaerobic threshold appears unchanged at 3 months unless exercise rehabilitation occurred. It seems that AT may lag behind other variables, perhaps because it reflects muscular adaptations which take time, and I see similar results in my study, with AT being the most negatively affected variable at both 2 and 6 months. I have used a z-score method to allow comparison between variables (change in each variable between tests is normalised for the standard deviation of that variable) and peak $\dot{V}_{O_2}$, AT, DP and circulatory power have the largest negative change at 2 months. However most of these appear to recover by 6 months -although total patient numbers are too small to draw statistically significant conclusions - with the exception of AT, which although not significant shows a trend to a continued deterioration from baseline.

OUES behaves in an interesting manner. It does not improve immediately postoperatively, but neither does it show the significant declines noted with peak $\dot{V}_{O_2}$, AT and some of the other variables. Surprisingly changes postoperatively were less marked in patients with atrial fibrillation than in those in sinus rhythm.
8.5.3 $V_{E}/V_{CO_2}$ relationship

The main group of variables which appear to show an early and persistent benefit are variables of the $V_{E}/V_{CO_2}$ relationship. This may help us to explain the patients’ symptomatic improvement. In their day-to-day lives patients don’t typically stress themselves maximally, and may not reach their anaerobic threshold whilst only undertaking activities of daily living, so a decline in peak $V_{O_2}$ may go unnoticed. In contrast improved ventilatory efficiency means that work of breathing for any level of activity will be reduced, and this may be noted as an improvement in symptoms by the patient. Why does the $V_{E}/V_{CO_2}$ relationship improve? This relationship is largely felt to reflect 2 major components in a heart failure patient; abnormal ventilation: perfusion matching; and primary hyperventilation possibly driven from the ergoreflex. In mitral valve disease high left atrial pressures drive abnormal pulmonary vasoregulation leading to ventilation: perfusion mismatch, this could potentially be reversed soon after amelioration of the valve lesion. Although there are medium-term reductions in lung function, as I show with a reduction in $K_{CO}$ and FEV$_1$ postoperatively, because the lungs capacity for diffusion of carbon dioxide is far greater than for oxygen (Wagner et al 1971) it is possible that these reductions do not affect the $V_{E}/V_{CO_2}$ relationship, whilst variables involving oxygen uptake may be more affected. Therefore muscular adaptations after surgery (which take time) are not necessary to improve the coupling of $CO_2$ elimination to ventilation which is solely affected by ventilation: perfusion mismatch.

8.5.4 Effect of surgery on lung function

Although not part of this study’s aims, it is interesting to note how lung function is affected by cardiac surgery. There have been a number of studies on lung function following heart transplantation (Ewert et al 1999, Ewert et al 2000) and in patients undergoing coronary artery bypass graft (CABG) surgery, although there is little data beyond the first few days postoperatively (Staton et al 2005). I am unaware of any studies specifically looking at the long-term effects to the lungs from mitral valve surgery. This differs from heart transplant recipients in that patients are normally less symptomatic, with fewer abnormalities of resting lung function preoperatively. Heart transplantation therefore possibly represents the extreme end of the spectrum. Conversely CABG surgery is
often performed on patients with structurally normal hearts, and operative time can be shorter in duration when compared to mitral valve repair. Therefore we may expect long-term abnormalities in lung function following mitral surgery to lie somewhere between CABG surgery and heart transplantation. I have shown an 8% reduction in $K_{CO}$ and 11% reduction in $D_{LCO}$, which at 2 months are smaller reductions than noted in patients comparing before with 6 and 12 months after heart transplantation (Ewert et al 1999). I also show small but significant changes in FEV$_1$ (-6% predicted) and FVC (-7% predicted) but not the ratio of these 2 values, suggesting that there has been the addition of a small restrictive, but not obstructive, insult to the lungs.

Although the numbers are small there was a 7-fold smaller reduction in $K_{CO}$ for my patients undergoing minimally invasive rather than open repair. Larger studies will be necessary to decide if this is significant.

8.5.5 Limitations

Unfortunately total numbers of patients are small, especially when trying to look for differences between surgical techniques and looking for differences at 6 months.

There are many confounding influences that we cannot account for. Postoperative recovery times, rehabilitation routines, nutrition etc will all vary between patients and may influence results but are not easily quantifiable.

For calculation of breathing reserve the spirometry performed preoperatively was used for the baseline test, and the spirometry performed postoperatively (at 2 months) was used for both the 2 month and 6 month test. It is possible that spirometry could have improved between these 2 time points; this would therefore have led to an underestimation of breathing reserve at the 6 month mark. It was, however, unfortunately not practical to repeat full lung function tests a third time.
8.5.6 Conclusions

Two months following mitral valve surgery peak $\dot{V}_{O_2}$ is depressed compared with preoperative measures. By 6 months peak $\dot{V}_{O_2}$ has returned to baseline. $\dot{V}_{O_2}$ at the AT shows significantly greater deterioration at 2 months compared with peak $\dot{V}_{O_2}$. In contrast OUES shows a small deterioration at 2 months, which is also back to baseline at 6 months. At 2 months this small decline in OUES is significantly smaller than the decline in peak $\dot{V}_{O_2}$ and $\dot{V}_{O_2}$ at the AT. Markers of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship show no such deterioration at 2 months, with a trend to large improvements at 6 months. OUES and $\dot{V}_E/\dot{V}_{CO_2}$ slope or ratio may be good variables to assess the changes in cardiac function. It does not appear that the increased influence of lung abnormalities on peak $\dot{V}_{O_2}$ accounts for its greater decline at 2 months when compared to OUES and $\dot{V}_E/\dot{V}_{CO_2}$ slope and ratio, so it is possible that these other variables are less affected by the peripheral changes associated with a postoperative state than peak $\dot{V}_{O_2}$. 
9.0 The Interventional Study – Cardiac Resynchronisation Therapy
Cardiac resynchronisation therapy (CRT) has been shown to improve prognosis and exercise capacity in patients with heart failure. Given that certain cardiopulmonary exercise test (CPX) variables appear to be more specific for cardiac function than others, these should show rapid improvements following CRT. Non-specific variables should improve to a lesser degree because they also require improvements within other organs. Conversely the sudden reduction in cardiac function, via the transient switching off of CRT, should reduce all variables equally as the heart becomes centrally limiting. I will investigate these hypotheses.

10 patients with heart failure were recruited prior to CRT implantation (pre-post study), and a further 13 with existing CRT (on-off study). They underwent CPX testing on a bicycle ergometer. Following a familiarisation test, each patient in the pre-post study underwent a personalised second test aiming for maximal exercise after approximately 10 minutes. This test was repeated approximately 2 months after CRT implantation. Patients in the on-off study underwent two identical, personalised tests in a random order, one with CRT active and one deactivated. QRS duration was on average 24.3 ms shorter with CRT.

The $\dot{V}E/\dot{V}CO_2$ ratio, $P_{ET}CO_2$ at the AT and respiratory frequency improved after implantation. The change in these variables, $\dot{V}E/\dot{V}CO_2$ slope and OUES were all better than the change in peak $V_o_2$ which non-significantly deteriorated post-implant.

When CRT was transiently switched off there were no changes to any variables. When all patients were considered together the act of resynchronisation only significantly affected the $\dot{V}O_2 - WR$ slope, reducing it by 0.64.

CRT was not associated with an improvement in exercise capacity as measured by peak $\dot{V}O_2$. There was a dissociation between various variables with ventilatory equivalents and slopes for $\dot{V}CO_2$ and OUES responding more positively to CRT than peak $\dot{V}O_2$ did. I propose that this may be because peak $\dot{V}O_2$ requires full body adaptations post-CRT which take longer to manifest.
9.2 Introduction

Cardiac resynchronisation therapy (CRT) has been shown to improve symptoms, exercise capacity and reduce hospitalisations (Cazeau et al 2001, Abraham et al 2002, Cleland et al 2005).

Improvements in exercise capacity seem consistent throughout the published literature. Cazeau et al showed a significant average improvement in peak \( \dot{V}_{O_2} \) of 8% following 3 months of CRT (Cazeau et al 2001). Larger changes in peak \( \dot{V}_{O_2} \) were noted in other studies; including a 28% increase at 3-6 months (Chwyczko et al 2008), 26% in a study of a single left ventricular lead (no right ventricular lead) after 12 months (Blanc et al 2004), and a 17% improvement in another study (Madaric et al 2007). In patients with pre-existing CRT, significant reductions in exercise capacity have been noted when removing the CRT component transiently (Salukhe et al 2008, Laveneziana et al 2009).

In most of these studies only peak \( \dot{V}_{O_2} \) was measured, although Chwyczko et al showed improvements in the anaerobic threshold, \( O_2 \) pulse and \( \dot{V}_E/\dot{V}_{O_2} \) slope; Laveneziana et al showed significant improvements in the \( O_2 \) pulse, \( \dot{V}_{O_2} - WR \) slope, respiratory frequency and ventilatory equivalents when CRT was “on”. Another study showed significant increases in OUES in responders to CRT (Berger et al 2011). However what none of these studies have done is to compare the response of different variables.

To prove my overall hypothesis that an ideal variable of cardiovascular function on CPX exists, it should reliably and consistently improve following an intervention to the heart designed to improve cardiac function. CRT supplies us with an excellent model, because cardiac output should improve and procedural recovery is minimal. Therefore any muscular maladaptation that occurs in mitral surgery patients owing to the long recovery time, should not occur in CRT patients. Because certain variables, such as the anaerobic threshold and peak \( \dot{V}_{O_2} \), may be more reflective of total body physiology, these may not improve as significantly following CRT implantation (because further changes are required within the body to improve them) compared with more specific variables. Variables more specific for the heart, such as OUES, would show improvements independent of the changes to maximal exercise capacity.

Conversely the sudden impairment of cardiac function through an intervention would affect all variables equally as the heart becomes the “lowest common denominator”. It is unethical to perform an intervention causing lasting cardiac impairment, however the transient deactivation of CRT by switching to an alternative pacing mode, has been shown, in the studies by Salukhe et al and Laveneziana et al, to be safe and result in sudden and
significant reductions in exercise capacity. Similar experiments have been conducted on patients with a left ventricular assist device for severe heart failure, where a reduction in pump speed (consistent with a decline in circulatory support) led to an immediate decline in peak $\dot{V}_O_2$ (Jakovljevic et al 2010, Noor et al 2012).

I aim, through the investigation of CPX testing, both before and, on average, 2 months after CRT implantation (pre-post study) as well as in patients with pre-existing CRT in 2 states, CRT-on and CRT-off (on-off study), how different variables behave, and how they relate to the overall improvement in exercise capacity.
9.3 Methods

9.3.1 Patient Recruitment

Patients were recruited for the pre-post study from Imperial College Healthcare NHS Trust (ICHNT) whilst awaiting the implantation of a cardiac resynchronisation therapy pacemaker. The indication for this had to include left ventricular systolic impairment with a broad QRS duration on ECG (>120 milliseconds). The decision for CRT had been made by one of the heart failure or electrophysiology physicians at ICHNT.

Separately patients with pre-existing CRT were recruited for the on-off study. These were either patients at ICHNT or the Royal Brompton and Harefield NHS Foundation Trust. The patients from the latter organisation had been approached by their pacing team to take part in research studies into CRT optimisation at ICHNT; I approached them after their acceptance. Patients with pre-existing CRT from my institution were either recruited from the aforementioned optimisation studies, the pacing clinic or heart failure clinics. The presence or absence of a defibrillator component to the pacemaker did not affect recruitment (i.e. the patient could have CRT-P or CRT-D).

9.3.2 Patient testing

Once patients were identified and consented for the pre-post study, they underwent full lung function tests, a transthoracic echocardiogram, venous blood tests and 2 cardiopulmonary exercise tests; these are considered the baseline tests. The full description of these tests can be found within the Methods Chapter. The first test was for familiarisation and to identify the optimal protocol for each patient to achieve similar duration of exercise in all (approximately 10 minutes of incremental exercise). The results of this first CPX test were not considered. The results of the second test, performed at least 2 hours after the first test to allow for muscular recovery, were performed on the patient’s optimal protocol and were recorded for analysis.

CPX tests were performed using a bicycle ergometer (Ergoline, GmbH, Bitz, Baden-Württemberg, Germany) on a COSMED Quark CPET System (COSMED S.r.l. Rome, Italy). The identification and measurement of CPX variables was performed as per directions from within the Methods Chapter and will not be discussed again here. However following on from the identification of a new set of contemporary reference equations (Gläser et al 2010, Chapter 4) percent of predicted values for peak $\dot{V}_{O_2}$, OUES and $O_2$ pulse were calculated using these SHIP equations described previously. The $\dot{V}_{O_2}$ at the AT will be presented as a percent of the predicted peak $\dot{V}_{O_2}$ using the SHIP reference equation for peak $\dot{V}_{O_2}$. All other variables will just be displayed giving their raw
values. $\dot{V}_E/\dot{V}_{CO_2}$ slope is displayed in 2 forms: slope 1 using data until the VCP; and slope 2 using data throughout exercise, in keeping with previous chapters, the $\dot{V}_E/\dot{V}_{CO_2}$ slope minus ratio difference was also calculated for each patient as per previous chapters.

Patients were contacted over a month after their CRT implantation to gauge recovery and arrange follow-up. It was agreed that 4 weeks should allow for enough recovery to make exercise testing safe, this was the minimum time from procedure to follow-up for the study. On arrival patients underwent venous blood sampling and a single CPX test on the same protocol as the main test from the preoperative testing (i.e. not the familiarisation test).

Patients within the on-off study had 3 CPX tests; the familiarisation test, one test on their optimal protocol with the CRT switched off, and one test on their optimal protocol with the CRT on. No more than 2 tests were performed on a single day, so this study required 2 visits. The order of the second and third tests was random. Typically these 2 tests were performed on the same day so differed from the other parts of the Observational study. Patients in the on-off study also had baseline full lung function tests, venous bloods and echocardiography. For CRT-on, the device was allowed to work at the pre-specified settings by the patients’ clinical teams. Upper tracking rates and defibrillator (if appropriate) thresholds were recorded to allow me to stop exercise if necessary in the event of inappropriately high heart rates during exercise. For CRT-off the patients’ devices were interrogated for an underlying rhythm (i.e. not pacing dependent). If they had a stable underlying rhythm (all patients) then the patients had their pacemaker set to DDI-40, so that they would resume pacing in the event of bradycardia, but would otherwise remain intrinsic. It was decided that in the absence of a stable intrinsic rhythm the patients would have the left ventricular lead component deactivated so that they only paced with their right ventricular lead; this would accentuate dyssynchrony. These changes to the patients’ pacemakers were made a few minutes prior to this CPX test, and returned back to normal immediately after. The patient underwent a “sham” turning off of their CRT for the CRT-on test so that they remained blinded to the order of tests. This was not possible for myself or the other operator as we had to, for safety, watch the ECG during exercise.

QRS durations whilst unpaced (pre-implant or CRT-off) and paced (post-implant and CRT-on) were measured using Iconico screen calipers (Iconico, New York, USA) on digitised ECG recordings at rest.
9.3.3 Hypotheses

I hypothesised that because CRT was a procedure requiring minimal recovery (and therefore minimal risk of loss of muscle bulk) we would not see a reduction in exercise capacity between pre-implantation and post-implantation. I hypothesised that we would see a small improvement in exercise capacity because the short time between implantation and retesting meant that muscular adaptations would be minimal. However cardiac specific variables (as based on previous chapters) such as OUES would show greater improvements because they may be more independent of the muscles’ impact on exercise.

Conversely the on-off experiment would show a similar drop in all variables because the muscles would not have had time to deteriorate from just a few minutes of dyssynchrony and so all variables would only be influenced by the change in cardiac output.

9.3.4 Statistical analysis

Variables of gas analysis and lung function have previously been assessed for normal distribution. Normally distributed variables are displayed as mean ± standard deviation; non-normally distributed variables as median (25th, 75th percentiles). For pre-implant versus post-implant changes of continuous variables paired t-tests were used, unless the variable was non-normally distributed, in which case a Wilcoxon signed rank test was employed. For graphical representation of the change in each variable it was standardised using the z-score method (z-score = change/standard deviation of variables within population). On these plots a leftward shift represents a worsening of a variable, regardless of whether this is an increase or decrease in the magnitude of the variable. Because z-scores standardise each variable these scores could be directly compared, paired t-tests were employed to compare each variable against one another to see if there was a significant difference in the change between tests.

Unpaired t-tests were performed between the 2 study groups. Where variables were non-parametric, the Kruskall-Wallis test was performed. Linear regression was employed to look for relations between continuous variables.

A p value <0.05 was considered significant throughout.
9.4 Results

9.4.1 Patient recruitment and characteristics

17 patients awaiting CRT implantation were recruited. 2 of these patients did not have a left ventricular lead implanted or switched on. 2 patients had problems with lead displacement, so did not have a prolonged period of biventricular pacing. 2 patients did not return for follow-up, and 1 patient had his procedure delayed multiple times and still had not undergone CRT prior to this analysis, leaving only 10 patients.

15 patients who had previously undergone CRT were assessed comparing the on-off state. 1 patient with alcoholic cardiomyopathy and long-standing CRT was found on echocardiogram to have near normal systolic function. He was abstinent from alcohol and in the CRT-off state his QRS duration was noted to be narrow (<120ms). He was hence excluded from all analyses. A further patient did not attend his second visit. He had already undertaken 2 tests (familiarisation plus CRT-on) so his data was retained within the reproducibility and Observational studies but not within the Interventional study. This left 13 patients for the CRT on-off study.

For the 10 patients tested before and after implantation (mean age 66.9 ± 10.4 years) the average time to testing was 52 days (IQR 42-69 days), however there were 2 patients as significant outliers due to unavoidable delays (returned 154 and 160 days post-implantation). For the 13 patients with pre-existing CRT (mean age 67.8 ± 8.5 years), 11 undertook both tests on the same day (therefore both on a different day to the familiarisation test), 1 patient undertook both tests within 4 weeks and 1 patient undertook both tests within 5 weeks. Order was randomised for each patient; 8 undertook CRT-off first, 5 CRT-on first. Of the 2 patients performing the 2 tests on different days the order was different. Details of the 23 patients are shown in Table 9.1. The first 10 patients are from the pre-post study, the last 13 from the on-off study. Patient 1 was in atrial fibrillation; all other patients were in sinus rhythm. Patients 3 and 21 had type 2 diabetes mellitus. On echocardiography only ejection fraction from the principal variables differed between these two study groups (30.8 ± 7.2 % pre-post study; 40.9 ± 9.2 % on-off study, p=0.01 between studies). One patient within the CRT on-off study group had reasonable systolic function, however they had a broad QRS duration on CRT-off so were retained within the study. Mean left ventricular internal dimension in diastole was 55 ± 8 mm, left ventricular diastolic volume 165 ± 59 cm³, left ventricular systolic volume 107 ± 47 cm³, fractional shortening 16.6 ± 7.7 %, left atrial volume 86 ± 35 cm³ and TAPSE 18.8 ± 4.5 mm.

Mean QRS duration was 169 ± 21 milliseconds pre-implant and 176 ± 34 milliseconds in the second group when in the off-state.
<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age (yrs)</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
<th>NYHA class</th>
<th>Smoking status</th>
<th>FEV₁ (% predicted)</th>
<th>Peak $V_{O₂}$ (% predicted)</th>
<th>Protocol (W/min)</th>
<th>Order of tests</th>
<th>Indication for implant</th>
<th>Time from implant (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>81</td>
<td>170</td>
<td>83</td>
<td>28.6</td>
<td>3</td>
<td>Ex-smoker</td>
<td>81</td>
<td>54.5</td>
<td>6</td>
<td></td>
<td>IHD</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>63</td>
<td>157</td>
<td>61</td>
<td>24.7</td>
<td>1</td>
<td>Never</td>
<td>121</td>
<td>85.5</td>
<td>10</td>
<td></td>
<td>DCM</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>60</td>
<td>149</td>
<td>74</td>
<td>33.3</td>
<td>2</td>
<td>Never</td>
<td>111</td>
<td>64.6</td>
<td>8</td>
<td></td>
<td>DCM</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>80</td>
<td>175</td>
<td>80</td>
<td>26.1</td>
<td>2</td>
<td>Current</td>
<td>96</td>
<td>94.2</td>
<td>10</td>
<td></td>
<td>DCM</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>78</td>
<td>178</td>
<td>85</td>
<td>26.8</td>
<td>2</td>
<td>Never</td>
<td>89</td>
<td>80.7</td>
<td>12</td>
<td></td>
<td>DCM</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>63</td>
<td>190</td>
<td>119</td>
<td>33</td>
<td>3</td>
<td>Ex-smoker</td>
<td>91</td>
<td>56.3</td>
<td>15</td>
<td></td>
<td>IHD</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>72</td>
<td>173</td>
<td>109</td>
<td>36.4</td>
<td>2</td>
<td>Ex-smoker</td>
<td>120</td>
<td>90.8</td>
<td>12</td>
<td></td>
<td>IHD</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>64</td>
<td>167</td>
<td>82</td>
<td>29.6</td>
<td>2</td>
<td>Ex-smoker</td>
<td>86</td>
<td>58.8</td>
<td>10</td>
<td></td>
<td>IHD</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>50</td>
<td>183</td>
<td>79</td>
<td>23.6</td>
<td>2</td>
<td>Current</td>
<td>90</td>
<td>48.7</td>
<td>12</td>
<td></td>
<td>IHD</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>58</td>
<td>172</td>
<td>87</td>
<td>29.4</td>
<td>1</td>
<td>Ex-smoker</td>
<td>90</td>
<td>76.9</td>
<td>15</td>
<td></td>
<td>DCM</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>73</td>
<td>174</td>
<td>82</td>
<td>27.1</td>
<td>2</td>
<td>Never</td>
<td>109</td>
<td>54.0</td>
<td>10</td>
<td>off-on</td>
<td>IHD</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>71</td>
<td>177</td>
<td>67</td>
<td>21.4</td>
<td>3</td>
<td>Never</td>
<td>59</td>
<td>37.5</td>
<td>10</td>
<td>on-off</td>
<td>DCM</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>84</td>
<td>159</td>
<td>63</td>
<td>24.9</td>
<td>1</td>
<td>Ex-smoker</td>
<td>130</td>
<td>98.7</td>
<td>10</td>
<td>on-off</td>
<td>IHD and heart block</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>66</td>
<td>173</td>
<td>78</td>
<td>26.1</td>
<td>2</td>
<td>Ex-smoker</td>
<td>33</td>
<td>63.5</td>
<td>12</td>
<td>off-on</td>
<td>DCM</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>62</td>
<td>162</td>
<td>82</td>
<td>31.2</td>
<td>2</td>
<td>Never</td>
<td>103</td>
<td>73.1</td>
<td>10</td>
<td>on-off</td>
<td>DCM</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>69</td>
<td>175</td>
<td>70</td>
<td>22.9</td>
<td>2</td>
<td>Ex-smoker</td>
<td>81</td>
<td>69.2</td>
<td>12</td>
<td>off-on</td>
<td>IHD, AoV disease</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>53</td>
<td>179</td>
<td>116</td>
<td>36.2</td>
<td>3</td>
<td>Ex-smoker</td>
<td>87</td>
<td>52.7</td>
<td>8</td>
<td>on-off</td>
<td>IHD</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>66</td>
<td>165</td>
<td>85</td>
<td>31.2</td>
<td>2</td>
<td>Never</td>
<td>76</td>
<td>89.3</td>
<td>10</td>
<td>off-on</td>
<td>DCM</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>55</td>
<td>173</td>
<td>93</td>
<td>31.1</td>
<td>1</td>
<td>Never</td>
<td>88</td>
<td>74.2</td>
<td>15</td>
<td>off-on</td>
<td>DCM</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>77</td>
<td>168</td>
<td>65</td>
<td>23</td>
<td>2</td>
<td>Never</td>
<td>69</td>
<td>66.3</td>
<td>8</td>
<td>on-off</td>
<td>CHB, pacing induced LVSD</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>67</td>
<td>183</td>
<td>122</td>
<td>36.4</td>
<td>2</td>
<td>Ex-smoker</td>
<td>79</td>
<td>65.7</td>
<td>15</td>
<td>on-off</td>
<td>DCM</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>M</td>
<td>74</td>
<td>186</td>
<td>112</td>
<td>32.4</td>
<td>2</td>
<td>Never</td>
<td>84</td>
<td>85.0</td>
<td>15</td>
<td>off-on</td>
<td>IHD</td>
<td>5</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>64</td>
<td>172</td>
<td>79</td>
<td>26.7</td>
<td>2</td>
<td>Current</td>
<td>68</td>
<td>66.0</td>
<td>12</td>
<td>off-on</td>
<td>IHD, AVR</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 9.1: Patient baseline characteristics. Patients 1-10 were recruited prior to CRT implantation; patients 11-23 had longstanding CRT and were tested with CRT on vs off. AoV = aortic valve disease, CHB = complete heart block, AVR = aortic valve replacement, LVSD = LV systolic dysfunction.
9.4.2 Changes between pre- and post-implantation

There were no significant changes in haemoglobin, creatinine, eGFR, sodium, bicarbonate or BNP between pre-implantation and post-implantation. 8 patients described themselves as having responded to CRT, 1 other was unsure, whilst 1 patient felt there had been no change and had had repeated chest infections post-implant. There was a non-significant reduction in NYHA class (mean change – 0.20, p=0.17). 2 patients, despite both taking 2.5mg of Bisoprolol, exceeded the upper tracking rate of the pacemaker and assumed an intrinsic rhythm towards the end of exercise. Heart rates did not closely approach the defibrillator thresholds and neither intrinsic rhythm led to a deterioration in symptoms, so exercise was not halted. The potential impact of the intrinsic rhythm is shown in Figure 9.1.

2 patients had an increase in beta-blocker dose, and 1 patient a decrease.

Mean change in QRS duration when comparing post-implant or the on-state, with pre-implant or the off-state was a 24.3 millisecond reduction with CRT (range -33 to 74 milliseconds).

9.4.3 Comparing the 2 groups

There were no differences in age, height, weight, BMI, smoking status, and beta-blocker use at baseline between those undergoing CRT and those in the on-off study. Other baseline variables compared include similar BNP values (p=0.56) and NYHA class (p=0.90).

CPX tests from patients recruited into the on-off study were compared with those in the pre- post-implant study; results are shown in Table 9.2. At baseline/CRT-off the only significant differences between the groups were a \( \dot{V}_E/\dot{V}_{CO_2} \) slope two 10.9 higher in the pre-implant group (p=0.03 by t-test, p=0.03 by Kruskall-Wallis), \( \dot{V}_E/\dot{V}_{CO_2} \) ratio at the AT 7.1 higher in the pre-implant group (p=0.02 by t-test, p=0.05 by Kruskall-Wallis), an end-tidal CO\(_2\) at the AT 4.9 mmHg lower in the pre-implant group (p=0.01), minute ventilation 16.7 L/min higher in the pre-implant group (p=0.03 by t-test), and a 7 breaths per minute higher respiratory frequency in the pre-implant group (p=0.02).
Figure 9.1: Sparkline plots of the change of $\dot{V}_{O_2}$ over time for the 10 patients post-implantation. The top two patients breached the upper tracking rate on their CRT devices, and assumed an intrinsic rhythm; this is denoted by the black arrowhead. It can be seen that in these two patients $\dot{V}_{O_2}$ reaches a plateau and may start to fall whereas this phenomenon is not noted in the other eight patients who remained in a paced rhythm throughout.
<table>
<thead>
<tr>
<th></th>
<th>Pre-Post Study</th>
<th>On-Off Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline test</td>
<td>Δ at 2 month test</td>
</tr>
<tr>
<td>Peak $\dot{V}_{O_2}$ (mL/min)</td>
<td>1334 ± 340</td>
<td>-109 ± 198</td>
</tr>
<tr>
<td>Peak $\dot{V}_{O_2}$ (mL/min/kg)</td>
<td>15.6 ± 3.1</td>
<td>-1.02 ± 1.94</td>
</tr>
<tr>
<td>Peak $\dot{V}_{O_2}$ (% predicted)</td>
<td>71 ± 16</td>
<td>-4.4 ± 10.1</td>
</tr>
<tr>
<td>AT (mL/min)</td>
<td>973 ± 229</td>
<td>-71 ± 151</td>
</tr>
<tr>
<td>AT (% of pred peak $\dot{V}_{O_2}$)</td>
<td>52.9 ± 13.9</td>
<td>-3.1 ± 7.8</td>
</tr>
<tr>
<td>OUES</td>
<td>1.40 ± 0.44</td>
<td>-0.02 ± 0.28</td>
</tr>
<tr>
<td>OUES/kg</td>
<td>16.5 ± 4.7</td>
<td>-0.07 ± 3.27</td>
</tr>
<tr>
<td>OUES (% predicted)</td>
<td>66 ± 18</td>
<td>0.1 ± 14.1</td>
</tr>
<tr>
<td>OUEP</td>
<td>31.2 ± 4.8</td>
<td>-0.3 ± 2.93</td>
</tr>
<tr>
<td>$O_2$ Pulse (mL/beat)</td>
<td>12.2 ± 3</td>
<td>-0.16 ± 1.4</td>
</tr>
<tr>
<td>$O_2$ Pulse (% predicted)</td>
<td>80 ± 19</td>
<td>-2.3 ± 11.9</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 1</td>
<td>37.9 (33.1, 48.9)</td>
<td>-3.1 ± 6.7</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope 2 *</td>
<td>37.9 (35.4, 58.6)</td>
<td>-3.6 ± 7.6</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio nadir</td>
<td>33.7 (32.1, 39.4)</td>
<td>-0.7 ± 2.6</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ ratio AT **</td>
<td>37.7 (33.6, 44.5)</td>
<td>-2.7 ± 3.3</td>
</tr>
<tr>
<td>$\dot{V}<em>E/\dot{V}</em>{CO_2}$ slope - ratio</td>
<td>-3.4 (-0.9, -0.5)</td>
<td>2.3 ± 5.4</td>
</tr>
<tr>
<td>RER at peak</td>
<td>1.14 ± 0.1</td>
<td>0.04 ± 0.11</td>
</tr>
<tr>
<td>$P_{ET}CO_2$ at AT (mmHg) **</td>
<td>31.5 ± 4.6</td>
<td>3 ± 3.4</td>
</tr>
<tr>
<td>HR at peak (bpm) **</td>
<td>120 ± 26</td>
<td>-11 ± 8</td>
</tr>
<tr>
<td>DP (mmHg bpm)</td>
<td>18693 ± 6008</td>
<td>-1684 ± 3636</td>
</tr>
<tr>
<td>Peak circulatory power</td>
<td>212741 ± 92111</td>
<td>-20964 ± 57858</td>
</tr>
<tr>
<td>Peak $O_2$ saturations (%)</td>
<td>98.5 (96, 100)</td>
<td>-1 ± 4.5</td>
</tr>
<tr>
<td>BR at AT (%) †</td>
<td>68.9 (66.7, 72.3)</td>
<td>5 ± 5.8</td>
</tr>
<tr>
<td>Breathing reserve (%)</td>
<td>38.4 (26.5, 42.9)</td>
<td>8.2 ± 13.4</td>
</tr>
<tr>
<td>$\dot{V}_{O_2} - WR$ slope †</td>
<td>9.5 (8.4, 9.8)</td>
<td>-1.19 ± 1.05</td>
</tr>
<tr>
<td>Peak work rate (W)</td>
<td>100 ± 26</td>
<td>-4 ± 19.1</td>
</tr>
<tr>
<td>$HR - \dot{V}_{O_2}$ slope</td>
<td>0.043 (0.04, 0.06)</td>
<td>0.007 ± 0.008</td>
</tr>
</tbody>
</table>

Table 9.2: CPX results for the pre- and post-implant patients, and the CRT on-off patients. Pre-implant (column 2) and CRT-off (column 4) are considered the two baseline states, the results of post-implant (column 3) and CRT-on (column 5) are displayed as a change (Δ) from the baseline states. * p <0.05 baseline pre-post study vs on-off study † p <0.05 baseline vs 2 months pre-post study ‡ p <0.05 on versus off.
9.4.4 CPX results – The pre-post study

10 patients underwent two identical CPX tests, one prior to implantation of CRT and a second an average of 52 days (IQR 42–69 days) after. On average there was a non-significant reduction in peak \( \dot{V}_O_2 \) of 1.02 mL/min/kg, or 4.4% of predicted. \( \dot{V}_O_2 \) at the AT showed a similar, non-significant, reduction of 3.1%. OUES showed trivial, non-significant differences between tests (-0.02 kg\(^{-1} \), +0.7% predicted). There were suggestions that ventilatory control in heart failure was improved by CRT. All measures of the \( \dot{V}_E/\dot{V}_{CO_2} \) relationship improved, although only the \( \dot{V}_E/\dot{V}_{CO_2} \) ratio at the AT showed statistically significant improvements post-implant (p=0.03). The \( P_{ET}CO_2 \) at the AT was 3 mmHg higher post-implant (p=0.02), with a reduction in respiratory frequency (-4.6 breaths per minute, p=0.03). Breathing reserve was significantly improved by +5% at the AT and non-significantly by +8% at peak (p=0.02 and p=0.09 respectively). Heart rate was significantly lower at rest and peak (p=0.01 and p=0.002 respectively). The \( \dot{V}_O_2 - WR \) slope was significantly lower (worse) post-implant with the value falling in 8 of the 10 patients (p=0.001). There were no significant differences in duration, peak work rate or peak RER. Figure 9.2 shows the change in a few principal variables.

When the change in each of the major variables was compared to the change in peak \( \dot{V}_O_2 \) (standardised by z-scores) the change in OUES (p=0.005), OUES/kg (p=0.02), percent of predicted OUES (p=0.02), \( \dot{V}_E/\dot{V}_{CO_2} \) slope (p=0.03) and ratio at nadir (p=0.04), and \( P_{ET}CO_2 \) at the AT (p=0.005) were all better than for peak \( \dot{V}_O_2 \).

9.4.5 CPX results – The on-off study

All patients had an intrinsic rhythm capable of sustaining exercise, therefore all patients within this part of the study were set to DDI-40, and none of them required RV pacing during their CRT-off test.

There were no significant differences between patients when tested with their intrinsic rhythm (CRT-off) versus full biventricular pacing (CRT-on) (Figure 9.3). Importantly there were no significant differences between the groups in resting (66 bpm CRT-off versus 65 bpm CRT-on, p=0.77) and peak heart rates (100 bpm CRT-off versus 103 bpm CRT-on, p=0.18) and the \( HR - \dot{V}_O_2 \) slope (0.038 CRT-off versus 0.036 CRT-on, p=0.62), suggesting that any contribution from rate-responsive pacing was not significant. When the change in each of the major variables was compared to the change in peak \( \dot{V}_O_2 \) (standardised by z-scores) there were no significant differences within the on-off group.
When all the patients from both studies were considered together only the $\dot{V}_{O_2} - WR$ slope was significantly different with an average reduction of 0.64 (p=0.02) between pre-implant/off-CRT and post-implant/on-CRT. Only the $\dot{V}_{O_2} - WR$ slope’s z-score was significantly different from peak $\dot{V}_{O_2}$ when both groups’ results were considered together.

The change in variables between on-off and pre-post were significantly different for peak $\dot{V}_{O_2}$ (p<0.05), $\dot{V}_E/\dot{V}_{CO_2}$ slope throughout exercise (p=0.03 by Kruskall-Wallis), $\dot{V}_E/\dot{V}_{CO_2}$ ratio at the AT (p=0.01 by Kruskall-Wallis), $P_{ET}CO_2$ at the AT (p=0.003), breathing reserve at the AT and peak (p=0.01 and p=0.03 respectively) and $\dot{V}_{O_2} - WR$ slope (p<0.05).

Because 6 patients finished the tests with breathing reserves <30% (i.e. could be considered to have a component of respiratory limitation) the analyses were rerun without these patients, with no change in results. Change in the QRS duration in the patients following implantation or after switching CRT on related only to the change in $O_2$ pulse and $\dot{V}_{O_2} - WR$ slope between the 2 tests and to no other variables. Both of these relationships were negative, i.e. the greater the improvement in QRS the less the improvement in $O_2$ pulse and greater deterioration in $\dot{V}_{O_2} - WR$ slope. Baseline (pre-implant or CRT-off) QRS duration did not relate to exercise capacity.
Figure 9.2: Change in 13 variables between a pre-implant CPX and a CPX performed on average 2 months post-implant. The change has been converted in a z-score for comparison between variables. A deviation to the left indicates a worsening of the variable following implantation (so for the $\dot{V}_B/\dot{V}_{CO_2}$ slope a positive score indicates a decrease in magnitude).
Figure 9.3: Change in 13 variables between a CPX performed with CRT enabled and another with CRT disabled. The change has been converted in a z-score for comparison between variables. A deviation to the left indicates a worsening of the variable with CRT enabled (so for the $\dot{V}_E/\dot{V}_{CO2}$ slope a negative score indicates an increase in magnitude).
9.5 Discussion

Cardiac resynchronisation therapy for patients with heart failure does not appear to significantly improve exercise capacity within my cohort of patients receiving CRT. If it does, it requires longer than 2 months for these improvements to be noticeable. If improvements in exercise capacity will occur and were not identified during this time frame, this action does not appear to be significantly influenced by the acute haemodynamic improvements associated with the act of resynchronisation, given we see no early improvements, and removing it (by switching LV pacing off) does not affect exercise capacity. The state of resynchronisation may, without affecting peak exercise capacity, exert beneficial effects on ventilatory efficiency.

9.5.1 Changes in ventilation

Despite the numbers being small there are strong trends to improvements in all measures of the $\dot{V}_e/\dot{V}_{CO_2}$ relationship following implantation of CRT. There are also significant increases in $P_{ET}CO_2$ at the AT and reductions in respiratory frequency. Minute ventilation at anaerobic threshold and peak exercise are reduced to a greater amount on average than oxygen uptake at these points, suggesting less ventilation is required to perform a similar amount of work; this may explain the improvement in subjective symptoms of exercise intolerance following implantation. It has long been established that patients with heart failure have elevated ventilatory equivalents, often wrongly referred to as ventilatory efficiency. Historically an elevated minute ventilation to carbon dioxide ratio was believed to be caused principally by an increase in ventilation: perfusion mismatch, hence it was considered to be synonymous with ventilatory inefficiency (Wasserman et al 1997). However more recently it has been proven that a large contribution to the elevated ratio of ventilation to carbon dioxide excretion is caused by a reduced arterial $CO_2$ set-point (Wensel et al 2004). Increased ventilation is required to maintain arterial carbon dioxide levels lower than average. It has become clear that the muscle is central to this pathophysiological state; ergoreceptors within the muscle, sensitive to hydrogen ions, drive ventilation (Scott et al 2002). It may be that although peripheral lactate levels are not elevated, lactate contained within the muscle’s local environment triggers the ergoreflex. Typically patients with heart failure display a flattened end-tidal $CO_2$ profile during incremental exercise; they do not show the rise from baseline to a plateau around anaerobic metabolism before the fall due to worsening systemic acidosis. End-tidal $CO_2$ is reflective of arterial $CO_2$ levels, although it becomes less closely correlated to arterial levels the more abnormal a patient’s physiology is. I have shown that, even in this small population of patients, a significant rise in end-tidal $CO_2$ at the anaerobic threshold occurs after CRT. Whilst we cannot infer mechanistic conclusions from these results, they may
suggest a reduction in the ergoreflex with a resetting of the CO\textsubscript{2} set-point at a level closer to that seen in healthy adults.

It is interesting to note that in the mitral intervention study very similar results were seen, with small but noticeable improvements in the ventilatory equivalents for carbon dioxide early after intervention.

Whilst the ventilatory equivalent measures for carbon dioxide show improvement the same cannot be said for oxygen. The oxygen uptake efficiency plateau (OUEP) is conceptually similar to the $\dot{V}_E/\dot{V}_{CO_2}$ ratio at the nadir, only as an inverse relationship (Sun et al 2012). It is a recent addition to the panoply of CPX variables, although its own inverse, the $\dot{V}_E/\dot{V}_{O_2}$ ratio, has been around for much longer. It has been postulated as a marker of cardiovascular fitness, and within the seminal papers describing its use was measured as the highest total 90 second continuous average during exercise. This “plateau” typically happens around the anaerobic threshold; in fact it is almost identical to averages of the OUE at the AT (Sun et al 2012). I believe the authors used a 90 second average to minimise the influence from cycles of oscillatory breathing which may be up to 90 seconds in duration. There are significant physiological differences with how the body manages oxygen and carbon dioxide; whilst it is easily to alter arterial carbon dioxide levels through adjusting ventilatory patterns, this is less possible with oxygen. Should improvements in the $\dot{V}_E/\dot{V}_{CO_2}$ relationship stem principally from improvements to ventilation: perfusion matching we would expect to see similar improvements (albeit a rise in OUEP corresponds to a fall in $\dot{V}_E/\dot{V}_{CO_2}$) in OUEP. We do not, in either CRT patients or in those post mitral valve repair. This lack of similarity between these two conceptually similar groups of variables perhaps more than any other result in this study shows the variability in the body’s response to improving cardiac haemodynamics.

### 9.5.2 Lack of improvement in exercise capacity

Both arms of this study failed to show an improvement in exercise capacity despite the pre-post study observing a subjective improvement in symptoms. These findings do not agree with published trials. Cazeau et al, in one of the early studies into CRT performed a cross-over design of 3 months of active CRT pacing, and 3 months inactive (Cazeau et al 2001). Overall there was a significant 8% improvement at the end of the periods of pacing, however it appeared that this benefit was greater for those who had the inactive state first (+1.6 mL/min/kg compared to +0.6 mL/min/kg). This would suggest that following a period of CRT, benefits are exerted that are independent of the act of acute resynchronisation; in this respect the results are comparable to mine. Other studies have shown a 28% increase in peak $\dot{V}_{O_2}$ at 3-6 months post-implant compared to pre-implant
(Chwyczko et al 2008), a 26% increase in peak $\dot{V}_{O_2}$ in a study examining the benefit of a single left ventricular lead (no right ventricular lead) after 12 months (Blanc et al 2004) and a 17% improvement in peak $\dot{V}_{O_2}$ in a study that showed benefits to exercise capacity correlated to the degree of improvement in mitral regurgitation (Madaric et al 2007). Only the former of these 3 studies commented on other CPX variables, showing large improvements in the $\dot{V}_{E}/\dot{V}_{CO_2}$ slope and ratio at the AT. Berger et al showed a significant improvement in OUES in responders to CRT therapy, but no change in non-responders. However many of the variables of severity in the responder group were worse than the non-responders at baseline, for example peak $\dot{V}_{O_2}$, and it may be that this is an example that patients with greater severity of disease benefit more (Berger et al 2011). It was interesting to see that OUES predicted response to CRT better than peak $\dot{V}_{O_2}$ and $\dot{V}_{E}/\dot{V}_{CO_2}$ slope (which non-significantly improved in both groups); this may be further evidence for the ability of OUES to have increased specificity for cardiac impairment. Salukhe et al showed a significant mean increase of 1.6 mL/min/kg in patients when their device was switched from AAI mode to full CRT (i.e. a similar experiment to our on-off study) (Salukhe et al 2008). In a further study looking at the on-off state of CRT there were significant improvements in the on-state seen in peak $\dot{V}_{O_2}$, O$_2$ pulse, $\dot{V}_{O_2} - WR$ slope and ventilatory equivalents (Laveneziana et al 2009). Respiratory frequency was significantly reduced with CRT on, which is similar to the results seen in my pre-post study group, although this was not replicated in my on-off study.

It is difficult to understand why I see such different results from the above studies. My sample size calculations were based on the results from these studies, so it looks like I am seeing a completely different effect rather than an underpowered study. It may be my study population differs significantly from those studied before. Certainly a few of my patients display signs of respiratory limitation (BR <30%) but the results with these patients removed do not change. Two patients went into an intrinsic rhythm during exercise (went above the upper tracking rate of CRT), with the resultant effect that final peak $\dot{V}_{O_2}$ appeared to underestimate the apparent trend that $\dot{V}_{O_2}$ was on prior to this; however this still resulted in similar peak $\dot{V}_{O_2}$s to pre-implant, not an effect large enough to alter the full group analysis. The ejection fraction within my pre-post cohort does not appear to differ largely from these previous studies. The QRS benefit seen within my patients from CRT is only 24 ms, perhaps these patients do not have a sufficient degree of correctable dyssynchrony. 15 of the 23 patients had a baseline OUES value below the threshold identified by Berger et al to be highly specific and sensitive for the detection of response to CRT. Interestingly both patients pre-implant with an OUES above this threshold self-described as responders. Perhaps a sensitive indicator that I have a group of patients that has not responded is the change in BNP between baseline and post-implant (+64pg/mL, p=0.71). The CARE-HF study showed large improvements
in N-terminal pro-BNP values at 3 months and at 18 months (over 50% reduction at this time point) in the CRT group (Fruhwald et al 2007). Significantly smaller improvements were seen in the control arm, suggesting a component of the improvement may have come from the intensification of medical therapy. This is still vastly different from my non-significant increase in BNP post-implant.

9.5.3 Differences between variables

Because of very small changes in all variables between the on-state and off-state (Figure 9.3) it is unsurprising that there are no significantly better or worse z-scores of change than peak \( \dot{V}_O_2 \). These results are also seen for all 23 patients together. However from the pre-post study we see that the z-score of change for all measures of OUES, \( \dot{V}_E/\dot{V}_CO_2 \) slope, breathing reserve, and the \( P_{ET}CO_2 \) at the AT improved to a significantly greater degree (or deteriorated to a significantly lesser degree) than peak \( \dot{V}_O_2 \). This is further support for the hypothesis that these variables are distinct from peak \( \dot{V}_O_2 \), and give us information beyond maximal exercise capacity. As suggested in previous chapters, interventional improvements in cardiac function will not improve maximal exercise capacity and anaerobic threshold without rehabilitation. That is not routinely offered to the recipients of CRT. The potential to improve exercise capacity through the medium of exercise rehabilitation may be limited by cardiac function; by improving cardiac function you raise the ceiling of what’s possible, but without work to improve muscular strength and endurance, ultimately gains won’t occur. The improvements seen in OUES and \( \dot{V}_E/\dot{V}_CO_2 \) slope compared to peak \( \dot{V}_O_2 \) may give us hints that improved exercise capacity is possible.

9.5.4 Limitations

Post-implant echocardiograms were not performed. I did not feel that an echocardiogram at this time point would add further information. 2 months is a short time period to expect reverse remodelling within the ventricle; any improvements in ejection fraction at this time point would be through the visual phenomenon of a resynchronised ventricle, not a true improvement in muscular contraction and systolic function. It could be criticised that the EF of the on-off group was higher than the pre-post group, however, again, because the echocardiogram in the on-off group was performed in the on-state, we would expect higher EFs because the act of resynchronisation improves how we visually assess EF. It is known that for any given EF the presence of interventricular mechanical delay (IVMD) improves prognosis and the greater the IVMD the lower the mortality (Richardson et al 2007, Cleland et al 2008). This is because dyssynchrony itself does not reflect reduced systolic
function of the individual myocytes. For any given EF with no IVMD there will be a greater burden of myocyte damage than in an equivalent ventricle with the same EF and significant IVMD. Therefore I do not believe that the higher EFs seen in my on-off group represent a set of patients with less severe heart failure than the pre-post group. Similar NYHA class and BNP values between my 2 groups supports that.

Final numbers for analysis were lower than anticipated. Sample size calculations required 44 patients within the mitral and CRT studies combined; this has been achieved, however it was disappointing that from 17 patients recruited prior to implantation only 10 returned at 2 months. This illustrates some of the difficulties in performing research on patients undergoing procedures for clinical purposes.

9.5.5 Conclusions

CRT implantation in a small group of patients did not improve exercise capacity. Transient deactivation of CRT in patients with long-term devices in situ did not improve exercise capacity. The model of on-off did not show any significant differences between variables, however the pre-post study did show variability in the response of different CPX variables.

Measures of ventilatory control, namely the $\dot{V}_E/\dot{V}_{CO_2}$ relationship, end-tidal CO$_2$ at the AT and respiratory frequency all improved following implantation. The change from baseline in these variables and the OUES was significantly better than the change from baseline in peak $\dot{V}_{O_2}$. These variables may better reflect the improved haemodynamics following CRT implantation than peak $\dot{V}_{O_2}$, which requires concurrent enhancements in other organ systems (namely the muscle) to show improvements.
10.0 Synthesis
10.1 Overview

Diagnosis and prognosis are two of the key roles for cardiopulmonary exercise testing. Furthermore, changes in objective exercise capacity as measured by CPX are felt to reflect disease progression or improvement following intervention. Yet it has never been reliably proven that CPX can distinguish the cause of a patient with exercise limitation via the algorithms published, and the reasons for the different prognostic powers of different variables on exercise testing has never been explained. Peak $\dot{V}O_2$ is the ubiquitous variable measured during a test and is well known to specialists within exercise physiology and beyond. But whilst it is an excellent variable to describe overall exercise capacity, its universality limits its use in many patients where multiple medical conditions coexist. In a young adult with dilated cardiomyopathy alone it will be an excellent marker for prognosis and therefore decision making on heart transplantation (Mancini et al 1991), but in a typical older patient with cardiac, and respiratory disease, can we be sure that abnormalities and changes over time in peak $\dot{V}O_2$ relate only to the heart? It was the aim of this thesis to identify if a variable exists that will be specific and sensitive to cardiovascular changes, limiting the influence from a common co-morbidity in cardiac patients that also severely limits exercise, respiratory disease.

4 criteria for an ideal variable were investigated: the ability to discriminate the two disease states; good reproducibility; appropriately alters with interventions to the heart; and finally relates to disease severity in cardiac, but not respiratory, disease.

I recruited patients with COPD, chronic heart failure and severe mitral valve disease, and measured objective variables of cardiac and respiratory function in all and then compared CPX tests in all as the Observational study. All patients underwent two baseline tests, to optimise individual testing conditions and to allow for comparison of the reproducibility of the various variables. Most of the patients recruited with CHF or mitral valve disease were undergoing a clinical intervention (CRT or valve repair) with the intention of improving the heart. These patients comprised the Interventional study, in which I compared the response of individual CPX variables to improvements in cardiac physiology.

Whilst many variables were examined throughout the thesis for comprehensiveness, only a small number were considered likely candidates. Amongst them was the $\dot{V}O_2$ at the anaerobic threshold (AT), measures of the $\dot{V}E/\dot{V}CO_2$ relationship, the O$_2$ pulse, the oxygen uptake efficiency plateau (OUEP), interactions between breathing reserve and RER, circulatory power, and measures of oxygen uptake kinetics; which included the oxygen uptake efficiency slope (OUES), oxygen uptake - work rate ($\dot{V}O_2 - WR$) slope and the heart rate -
oxygen uptake \((HR - \dot{V}_O_2)\) slope. I will go through many of these variables to describe their strengths and weaknesses.

### 10.2 Respiratory exchange ratio and breathing reserve

I will discuss these 2 variables together. The breathing reserve (BR) is a well established variable used to identify patients with respiratory limitation; a value < 30% has been suggested as a suitable cut-off (Eschenbacher et al. 1990, Messner-Pellenc et al. 1994, Milani et al. 2004). It is typically calculated using the patient’s FEV\(_1\), measured during spirometry, and the maximum minute ventilation at peak exercise. As such it is reliant on data obtained both during exercise and before. Spirometry is dependent on technique and effort and so FEV\(_1\) and therefore the BR, could easily be underestimated. This would have the effect of identifying a greater proportion of patients as respiratory limited. Another problem is that patients with heart failure commonly have abnormalities of spirometry; combined with the greater ventilatory requirements of exercise as seen with elevated ventilatory equivalents, means that heart failure patients may appear to have a degree of ventilatory limitation rather than respiratory disease per se. Patients with significant respiratory limitation typically stop exercise when the mechanical act of ventilation becomes discomforting. This can occur even though oxygen delivery to the exercising muscles may be normal and anaerobic metabolism has not occurred. Therefore the respiratory exchange ratio (RER) is typically lower in patients with COPD than CHF.

I have shown that both BR and RER display a good ability to discriminate patients with cardiovascular disease from COPD. BR measured at peak exercise was slightly superior to BR measured at the anaerobic threshold; both were significantly superior to the RER and although there was a slight improvement in the AUC with the addition of RER to BR, it was a non-significant improvement. BR at peak exercise was overall the best variable at discriminating patients with COPD from patients with CHF. However the aim of the thesis is to identify an ideal variable for cardiovascular dysfunction. For BR and RER patients with CHF and mitral valve disease behave more like healthy adults, with COPD patients behaving differently. Therefore these variables discriminate COPD from CHF or normal adults. BR could still be useful for the management of respiratory patients, so long as it correlates well with markers of disease severity. Symptom score correlated reasonably well with BR at the AT, but not at peak exercise. Importantly the 2 measures of the BR were amongst only a small handful of variables that did not correlate to peak \(\dot{V}_O_2\). BR at AT and peak had excellent test-retest reliability as measured by ICC. It is unsurprising that BR at peak had a poor coefficient of variation, as this is a...
measure that performs poorly when a variable can assume negative values. The reproducibility of BR was not affected by age, gender, BMI, disease category, time interval between tests or a change in ramp protocol.

Overall RER appears to add nothing beyond BR. The product of the two behaves similarly to BR alone, and is more complex. BR may be beneficial in identifying patients with respiratory limitation, and it is reproducible, but its magnitude adds nothing in terms of identifying severity of disease or symptom status. For patients with CHF it is unlikely it will be a useful variable in serial testing.

### 10.3 Oxygen uptake efficiency slope

The oxygen uptake efficiency slope (OUES) has been shown to be superior when predicting prognosis in patients with heart failure than both peak $\dot{V}_{O_2}$ and the $\dot{V}_E/\dot{V}_{CO_2}$ slope (Davies *et al* 2006). Pilot data from this thesis showed that OUES is much less sensitive to respiratory abnormalities than peak $\dot{V}_{O_2}$ or the $\dot{V}_E$ at the AT. Population data from over 1200 healthy adults also confirms that OUES is less affected by a number of lung function variables than peak $\dot{V}_{O_2}$, although more so than the $\dot{V}_E$ at the AT. Using predictive equations for FEV$_1$, correcting for age, gender and height, removed almost all of the relationship between FEV$_1$ and OUES. However there still remained a small relation and it was necessary to take this into consideration in the predictive equations generated from this population. In results similar to the pilot data, within the Observational cohort of 96 patients with CHF, mitral valve disease and COPD, OUES was actually inversely related to FEV$_1$, so that a worsening in FEV$_1$ led to an improvement in OUES; almost certainly however this is just a statistical anomaly due to the combined analysis of two discordant groups, cardiac and respiratory disease. Within this analysis the OUES interestingly did not relate to the $K_{CO}$, despite $K_{CO}$ being affected by heart failure as well as primary lung disease. Within the analysis absolute OUES was different between the groups, but this difference was more noticeable when adjusted for weight or as a percent of predicted. This was because there were significantly more males, and heavier individuals in the CHF group. OUES/kg and as a percent of predicted showed excellent discriminatory ability between CHF and COPD, second only to BR. What was also interesting was that the addition of the mitral patients, on average the group with the best exercise capacity, did not measurably reduce the area-under-curve on ROC curve analysis. A surprising result from the reproducibility study was that markers at peak exercise (peak $\dot{V}_{O_2}$ and the O$_2$ pulse) had the best test-retest reliability as measured by ICC, better than measurements of slopes which we would imagine would be superior. However whilst the ICC and CoV for OUES was not as strong as for peak $\dot{V}_{O_2}$ and the O$_2$ pulse, they were still very good,
and importantly there was not a significant mean difference in OUES between test 1 to 2, unlike peak $V_{O_2}$, i.e. there was no learning effect, and change in protocol intensity did not affect reproducibility. This may make it a more useful variable to serially measure in patients who are naive to exercise testing on their first visit. A slight weakness of OUES is that it appears to be less reproducible in heavier individuals and in those with COPD and CHF compared to mitral valve disease. However in all these groups the ICC was still $>0.88$. There remains no evidence for the comparison of OUES on treadmill versus bicycle ergometer testing, which until studied may restrict the applicability of these results to bicycle tests.

Perhaps disappointingly OUES did not appear to be affected much in either direction following interventions on the heart. The reduction at 2 months post-mitral valve surgery was less than for peak $V_{O_2}$ and the $V_{O_2}$ at the AT, which it could be argued reflects a lesser effect from the peripheral effects of recovery from major surgery on this variable. Interestingly the reduction in OUES was of a smaller, now non-significant, magnitude when patients with good preoperative exercise capacity were excluded, which perhaps indicates that in patients where their cardiac disease does not significantly limit their OUES it can still be affected by external components (surgical recovery), but in those where cardiac function is the driver to an abnormal OUES, the improvement to the heart allows for early improvements in the OUES. However at 6 months, when we would expect recovery to be complete and patients to have values better than those preoperatively, the OUES was unchanged from the baseline test. The results in patients undergoing CRT were very similar, with small, non-significant changes in OUES, although again OUES showed less deterioration than peak $V_{O_2}$ and the $V_{O_2}$ at the AT. It is difficult to understand why any variables worsened as CRT is a procedure unlikely to require recovery, muscular changes should therefore be minimal, and any cardiovascular changes should only be positive. In patients where the CRT was transiently turned off there were no real changes in any variables, at odds with previous work which had shown significant improvements in peak $V_{O_2}$ in a similar style of study. OUES had been shown to predict response to CRT, although my results would not support that.

Finally OUES does appear to be a good measure of disease severity. As shown above it doesn’t strongly relate to respiratory dysfunction, but correlates well to peak $V_{O_2}$, especially in patients with heart failure. Similarly it correlates to symptom scores in patients with heart failure but unlike other major variables has an inverse relationship in patients with COPD. Tissue Doppler measures of systolic and diastolic function as well as right ventricular function, correlated better with OUES than peak $V_{O_2}$ and almost all other variables, whilst B-natriuretic peptide, a highly objective measure of heart failure severity correlated strongly with OUES.
Overall therefore OUES is a reproducible, easily measured variable that correlates closely to disease and symptom severity in patients with cardiac disease but not respiratory disease, and appears to discriminate patients with cardiovascular limitation from respiratory limitation better than almost any other. Ideally it should be either weight adjusted or corrected using predictive equations. I believe it is the single best, and therefore ideal, variable identified through this thesis. In a heart failure patient with respiratory co-morbidities the serial measurement of OUES allows a way to isolate the contribution to exercise limitation from the heart.

10.4 Oxygen uptake at the anaerobic threshold

It has long been believed that this variable best discriminates these two common disease states. Patients with respiratory limitation are believed to have relatively normal ATs, the problem is that their lungs either stop them reaching it, or do so shortly after. In contrast the AT is reached early in patients with cardiovascular limitation. A threshold of 40% of a patient’s predicted peak $\dot{V}_{O_2}$ has been described as a useful discriminant.

Within my Observational group, the $\dot{V}_{O_2}$ at the AT was a poor discriminator between cardiac and respiratory disease and only 14 from 96 patients (2 with COPD) had a value below the suggested threshold of 40%. The $\dot{V}_{O_2}$ at the AT appears to be affected to a greater degree postoperatively than any other variable, with a large deterioration at 2 months that is still to return to baseline at 6 months. It, like many other variables, was unaffected by the implantation of CRT. It did relate strongly to symptom score in CHF patients but not COPD, but was less closely correlated than other variables to BNP, and was not correlated to echocardiographic variables of systolic and diastolic function. It was not as reproducible as peak $\dot{V}_{O_2}$ or the OUES when comparing the 2 tests performed by each patient, and showed a potential learning effect with a statistically significant systematic increase from test 1 to test 2, especially noted if the two tests were performed on the same day. It did however show almost excellent reproducibility amongst healthy adults.

I believe the $\dot{V}_{O_2}$ at the AT is principally a index of the muscle. Any chronic disease can, and will reduce it. It is known that endurance athletes see greater improvements in the $\dot{V}_{O_2}$ at the AT over peak $\dot{V}_{O_2}$ with training. Postoperatively following an operation with a considerable recovery time, mitral valve surgery, this variable is the most adversely affected. My results show no ability to discriminate and the accepted algorithms for distinguishing limiting physiology, within which the $\dot{V}_{O_2}$ at the AT is so central, should be reconsidered.
10.5 Oxygen uptake to work rate slope

The oxygen uptake to work rate (\(\dot{V}_O_2 - WR\)) slope, another measure of oxygen uptake kinetics with similarities to the OUES, has been suggested to be a marker of oxygen delivery and utilisation, with significant reductions in patients with cardiovascular disease. One of its perceived strengths is that the value shows very little variability in healthy adults, with minimal influence from age, gender and body size. My results from the SHIP data agree with the marginal influence of these variables on the slope, however in older adults there was an increased spread of values. A criticism of a seminal study into the specificity of a decrease in the slope for CHF was that these patients appeared to have shorter exercise durations than the controls; as I and others have shown, this would lead to a reduction in the slope independent of the patient’s physiology. My data goes some way to confirming their conclusion however, irrespective of test duration; in my Observational cohort this slope was a reasonable discriminator between CHF and COPD with similar exercise durations between the groups. The \(\dot{V}_O_2 - WR\) slope wasn’t strongly related to respiratory variables, and unlike many other variables its ability to discriminate was strengthened by the inclusion of mitral valve patients to the CHF cohort, suggesting it reflects cardiac physiology more specifically than just heart failure physiology. However it is poorly reproducible, and its magnitude can be heavily influenced by the intensity of the protocol used. Following intervention on the mitral valve there is an early deterioration, with a return to baseline levels, performing similarly to many of the other variables. When CRT is implanted the \(\dot{V}_O_2 - WR\) slope is the only principal variable that is reduced, despite an overall improvement in the patients’ symptom status. At baseline it does correlate to symptom score and markers of diastolic function. It is strongly correlated, although not as strongly as OUES, to BNP.

Overall it shows reasonable ability to discriminate but its other limitations, namely its poor reproducibility, make it an unsuitable candidate for the ideal variable. It must also be remembered that it can only be measured using bicycle ergometry.

10.6 \(O_2\) Pulse and the heart rate to oxygen uptake slope

Oxygen uptake is equal to the product of cardiac output and arterial-venous oxygen difference. Cardiac output is the product of stroke volume and heart rate, so \(\dot{V}_O_2\) divided by heart rate measures the product of stroke volume and arterial-venous oxygen difference. It is therefore taken as a surrogate for stroke volume, and so is expected to be reduced in patients with cardiac limitation. My results from the Observational study show the O₂ pulse was highest in the CHF patients and lowest in COPD, although these differences were no longer significant when a
percentage of predicted is calculated and it shows no ability to discriminate. The likely explanation for the higher absolute O₂ pulse in the CHF group is the use of rate-limiting drugs, principally beta-blockers, as following correction for the use of beta-blockers the between groups differences were no longer present. A reduction in heart rate would allow increased diastolic filling which should therefore lead to increased stroke volumes (Starling’s law).

The O₂ pulse correlated to symptom score as well in COPD as CHF patients, and to peak \( \dot{V}_O_2 \), although it did not correlate to BNP or echocardiographic variables. It has excellent reproducibility. Following mitral valve surgery it was non-significantly improved with no changes in patients undergoing CRT. It is possible that the non-significant changes after mitral surgery reflect a slight improvement in forward stroke volume with a reduction in the regurgitant fraction; peak heart rate at 6 months was unchanged and a reduction or increase in rate-limiting drugs did not affect the change in O₂ pulse in this group.

A limitation to the O₂ pulse is that submaximal effort will reduce its magnitude; the \( HR - \dot{V}_O_2 \) slope minimises this limitation because it measures a slope of the heart rate relationship to \( \dot{V}_O_2 \), rather than just the ratio at peak. It appears to be largely independent of respiratory abnormalities, both in health and disease, but is less reproducible than many other variables. It does not appear to correlate to most markers of disease or symptom severity, and importantly behaves differently between the two cardiac groups. It was significantly lower in the CHF patients and higher in the mitral valve disease patients, with the COPD patients in the middle. This bivariate distribution severely limits its ability to diagnose and monitor cardiac disease. The AUC on ROC curve analysis for all cardiac patients is very poor (0.53) and only moderate if mitral valve patients are excluded (0.63). It goes down after mitral surgery, which in a similar manner to the non-significant rise in O₂ pulse, may reflect the increase in forward stroke volume and reduced dependency on heart rate.

These two variables each have specific uses and strengths, but unfortunately do not appear to fulfil the necessary criteria for an ideal variable.

10.7 Ventilatory equivalents

There are a number of ways to measure the ventilatory response, typically the relationship of minute ventilation to carbon dioxide elimination, either as the regression slope throughout exercise or as instantaneous ratios at predetermined times such as the AT, is used. All measures are generally reproducible, with coefficients of
variation showing better reproducibility than intraclass correlation coefficients. Their test-retest reliabilities are worse in females and heavier individuals. They do not show a significant difference in mean values between tests i.e. no systematic bias or learning effect. However the $\dot{V}/\dot{V}_{CO_2}$ slope and $\dot{V}/\dot{V}_{CO_2}$ ratio do not show a strong ability to discriminate between the different aetiologies of disease, the principal criterion required by this study. These variables may be useful in determining an effect of intervention above and beyond all others however. They correlate to BNP and symptom status, and in both the CRT implantation and mitral studies showed strong trends towards improvement. It is difficult to reconcile the lack of benefit noted in many variables despite subjective suggestion of symptom improvement; these variables may give us an explanation. They represent the ventilatory work required to perform a certain degree of external work, improvements will lead to an improved sensation of breathlessness during exercise. This is what is of most importance to patients, not their peak exercise capacity. Improvements in cardiac physiology may lead to peripheral changes to the autonomic system, neurohormonal interactions and the ergoreflex which all reduce the peripheral drive to ventilation. There may also be improvements in ventilation to perfusion matching within the lungs, due to better blood flow distribution. These physiological adaptations reduce the drive to ventilation, yielding improved symptoms.

What I don’t think my results answer is which of the variables is superior. Ratios appear to have greater reproducibility but were less discriminant. The slope shows greater improvement after intervention, perhaps because it is influenced from ventilatory control during aerobic and importantly anaerobic metabolism, the latter where improvements in physiology may have the greatest bearing. Measuring the slope throughout exercise rather than the traditional approach of censoring data at the ventilatory compensation point, showed greater differences after intervention but given that very few patients went significantly past the VCP the significance of these results is difficult to assess. I still feel conceptually that stopping at the VCP is physiologically correct; however within this thesis the evidence for differences noted using both methods are at least presented.

Rather than making CO$_2$ elimination the focus, we could look at the relationship between oxygen uptake and ventilation. Recently a variable called the oxygen uptake efficiency plateau has been described (OUEP) which is the highest 90 second average of the ratio of $\dot{V}_{O_2}$ to $\dot{V}_E$. It was a highly reproducible variable but, unlike the $\dot{V}_E/\dot{V}_{CO_2}$ variables within my analysis, worse values were found in the COPD group, not the CHF. It showed moderate discriminant ability, but again to detect COPD, not CHF. It also behaved differently than the $\dot{V}_E/\dot{V}_{CO_2}$ variables after interventions, with early deteriorations noted. Although conceptually similar variables, these results show the difference in how ventilation interacts with oxygen uptake or carbon dioxide elimination.
Overall these variables, mainly those involving the $\dot{V}_E/\dot{V}_{CO_2}$ relationship, may be able to provide information in patients undergoing interventions, but the lack of discriminant ability between COPD and CHF does not make them good candidates at tracking changes specific to cardiac function over time.

### 10.8 Double product and circulatory power

These two variables are derived from systolic blood pressure (SBP) at peak exercise. DP is the product of heart rate and SBP, whilst circulatory power is the product of SBP and peak $\dot{V}_{O_2}$. DP was an excellent discriminator of CHF and COPD, but this was significantly reduced after the addition of mitral valve patients. This may just reflect the greater burden of heart rate and blood pressure reducing drugs taken by the CHF group compared to the other groups. Circulatory power showed poor ability to discriminate. Surprisingly given the accepted poor reproducibility of manually measured blood pressure and the difficulties measuring it during exertion, both these variables showed good test-retest reliability. DP correlated to systolic function on echocardiography, whilst both correlated to BNP. Unsurprisingly circulatory power correlated to peak $\dot{V}_{O_2}$, as it is calculated using peak $\dot{V}_{O_2}$. DP also correlated quite strongly to peak $\dot{V}_{O_2}$. On the Interventional studies they showed similar results to peak $\dot{V}_{O_2}$ with early deteriorations after mitral surgery that had not fully resolved by 6 months, with non-significant reductions after CRT implantation.

The poor discriminant ability of these variables, especially circulatory power, unfortunately make them poor candidates for an ideal variable of cardiac function.

### 10.9 Limitations

Limitations of each separate study are shown within each Chapter.

A major limitation is my assumption that principal disease categorisation also equals limiting physiology. Left ventricular impairment is not always symptomatic and may be an almost incidental finding. This is even more likely following CRT therapy when reverse remodelling may lead to long-term improvements in cardiac function. However I have wherever possible tried to take this into account, by looking separately at patients with reduced exercise capacity and removing from the cardiac group those with low breathing reserves. The study design with prospective echocardiography, BNP measurement and full lung function allows for patients to be
grouped differently than their prevailing diagnosis. However it is entirely possible that their known diagnosis biases my later judgement on whether they belong within that category or another.

Ideally these studies would be replicated in patients with new diagnoses of heart failure or COPD, prior to medical involvement. The testing of patients with long-standing diagnoses will influence the results in a number of ways, principally with the prescription of medications. There were naturally large differences in the numbers and types of medications between groups; these will certainly affect variables that are heart rate and blood pressure derived, and may affect other variables in ways we cannot determine. A diagnosis of heart disease is likely to lead to sedentary behaviour as this is often wrongly believed to be safer than regular exercise. It is difficult to ascertain how patients’ previous and current experience with exercise would differ between groups, although it is not in doubt that it will significantly affect almost all variables on exercise testing.

It could be argued that because I was aware the groups the patients were in I could somehow bias the exercise tests. However the protocols for performing the tests and calculating results were predefined. Only the AT and VCP were subjectively identified. I performed the test analysis remote to the test, patient details were also anonymised, to minimise the chance that I would understand which disease category the patient belonged to whilst measuring, and therefore inadvertently biasing my positioning of either of these thresholds.

It can be argued that my research group has a vested interest in OUES based on previous work they have published showing its prognostic superiority over the \( \dot{V}_E/\dot{V}_{CO_2} \) slope and peak \( \dot{V}_{O_2} \). It would be interesting for these results, certainly of the Observational study using ROC curve analysis, to be replicated by an independent study group.

Finally I have shown how variables change following interventions to the heart, but not to interventions to the lungs. An ideal cardiac variable would show no difference between 2 tests performed with different lung function, whilst other variables would change. A respiratory Interventional study was attempted; patients with COPD requiring tiotropium, a long-acting anticholinergic agent, had this and other long-acting inhalers withheld for 72 hours prior to a repeat CPX test, and then another test the following day after reinstitution of the inhalers. Unfortunately the first two patients became markedly breathless during the 72 hour period and the study was halted.
10.10 Conclusions

Breathing reserve appears to discriminate patients with COPD better than any other variable. However its magnitude has no value beyond being used as a threshold for discrimination. Measures of the $\dot{V}_E/\dot{V}_{CO_2}$ relationship were not discriminant but may reflect improvements in cardiac physiology following cardiac interventions. OUES, a highly reproducible variable, shows excellent ability to discriminate CHF from COPD, with a strong relationship to markers of disease and symptom severity in cardiovascular, but not respiratory, disease. It appears to be affected to a lesser degree by the peripheral changes brought about through recovery post-surgery than other variables. In a patient with concomitant cardiac and respiratory disease, the OUES is a variable that appears to specifically reflect the burden of disease from the heart alone.
11.0 Bibliography


12.0 Appendix 1 – Data analysis within Excel2007/2010

12.1 Identification of AT and slopes
12.2 Measurement of OUES and foreshortened OUES
12.3 Measurement of $\frac{V_{E}}{V_{CO_2}}$ slope up to the VCP and post-VCP

<table>
<thead>
<tr>
<th>Phase 2 (t)</th>
<th>VE</th>
<th>VT</th>
<th>ECO2</th>
<th>CO2</th>
<th>VE/VT</th>
<th>VE/CO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>10.49</td>
<td>10.50</td>
<td>10.038</td>
<td>9.0375</td>
<td>10.49</td>
<td>10.50</td>
</tr>
</tbody>
</table>

![Graph showing VE/CO2 slope with equations and R^2 value]
13.0 - Appendix 2

13.1 Participant Information Leaflet – Valve disease

Identifying an Ideal Cardiopulmonary Exercise Test Parameter
Chief Investigator: Dr Roland Wensel

What is the purpose of the study?

We want to learn how best to use a test called a Cardiopulmonary Exercise Test. This is a stress test involving cycling on an exercise bicycle, or walking on a treadmill, whilst doctors measure your breaths. We want to know how people with different diseases of the heart and lungs have different readings on this test to decide how best to treat patients with a range of heart and lung conditions.

Why have I been chosen?

You can take part in our research as you are awaiting valvular surgery.

It is not compulsory to take part and even if you initially agree to can withdraw at any time. Your usual treatment and follow up will not be affected should you choose not to take part. Participation is entirely voluntary.

What happens if I chose to take part?

You will meet with one of the research team to arrange suitable times for you to come for the studies, on this initial meeting blood tests will be taken from a vein in your arm which will be about 10-20ml (2-4 teaspoons). We will then arrange for you to come for the lung function tests, echocardiography studies and exercise tests.

The echocardiography studies take up to 1 hour. We will ask you to lie on a bed and we will attach several pieces of equipment including ECG leads. We will take images of your heart using an echo scanner. During the same visit we will arrange for you to have lung function testing which will also take about 1 hour. This involves controlled breathing into a machine under the instruction of a professional to record the ability of your lungs.

You will also undergo exercise testing on multiple visits (between 2 and 4, before and after surgery). This will involve cycling on an exercise bicycle with a mouthpiece in. This records your breathing and the amount of oxygen or carbon dioxide you breathe in and out. You are also attached to an ECG machine and have regular blood pressure readings. You will be expected to cycle until you are unable to continue. The bicycle will get progressively harder to cycle. We aim for the length of time for you to be cycling to be 10-15 minutes, total time will be less than 1 hour. You decide when you have cycled enough and you are unable to continue, the test stops when you stop cycling.

Expenses and Payments

Any travel and costs for meals will be reimbursed up to an amount of £50 per person and will vary based on individual expenses.

What are the side effects of taking part?

There are no drugs or invasive tests. We will require blood tests during your first visit, and rarely once more after then. Some people find having echocardiography uncomfortable. If this is the case we will take breaks and / or stop as required.

Occasionally, the blood pressure monitor can cause tingling on the finger. This goes away when removed.

A cardiopulmonary exercise test is tiring, and it is expected that you will be out of breath at the end. We will allow you plenty of time for recovery, and supply you with drinks and food as you need. The mouthpiece causes some dribbling, we will cover your clothes with an apron if required.
What if something goes wrong?

We will take every care in the course of this study, and the risk involved is minimal. Imperial College London holds insurance policies which apply to this study. If you experience serious and enduring harm or injury as a result of taking part in this study, you may be eligible to claim compensation without having to prove that Imperial College is at fault. This does not affect your legal rights to seek compensation. If you are harmed due to someone’s negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been treated during the course of this study then you should immediately inform the Investigator, Dr Anthony Barron (07932 648274) or Dr Roland Wensel. The normal National Health Service complaint mechanisms are also available to you. If you are still not satisfied with the response, you may contact the Imperial AHSC Joint Research Office, Imperial College London, St Mary's Hospital, 41 Praed Street, Paddington, London, W2 1NY (02033126484).

What are the benefits?

The research may not benefit you directly. However, we will have improved knowledge regarding the best possible use of this test for the future management of patients with a variety of heart conditions. This may impact on who and when patients undergo procedures and operations to the heart. This may prevent unnecessary treatments.

Will my taking part in this study be kept confidential?

Data collected about you will be entered on to a computer, in an anonymous format and any information taken from this investigation that can be identified will remain confidential. If you agree, we would contact your GP to inform them you are participating and will provide them with information regarding the study. At times during the study professional bodies (for example the NHS Research and Development team or Imperial College) may wish to audit our data to ensure that the study is being carried out appropriately. Any non-essential information regarding your involvement will not be shown. These auditors will be appropriate professionally trained individuals and all information will be kept confidential. Our research is in collaboration with experts in Germany, they will be given access to the results of your tests, but again your personal details will be kept anonymous.

What happens once the research stops?

The results may be presented at meetings and published so that it can be used to help treatment of other patients with heart failure, lung disease, arrhythmias and valve disease. Your participation will be kept strictly confidential. Should you wish to receive information regarding the outcomes of the study, this can be sent to you at the conclusion of the study.

Who has reviewed the study?

This study has been reviewed by the South East London Research Ethics Committee. If you have any further questions please do not hesitate to contact:

Dr Anthony Barron on 07932 648274

Thank you for taking the time to consider participating in this study.
What is the purpose of the study?
We want to learn how best to use a test called a Cardiopulmonary Exercise Test. This is a stress test involving cycling on an exercise bicycle, or walking on a treadmill, whilst doctors measure your breaths. We want to know how people with different diseases of the heart and lungs have different readings on this test to decide how best to treat patients with a range of heart and lung conditions.

Why have I been chosen?
You can take part in our research as you have COPD/lung disease.
It is not compulsory to take part and even if you initially agree to can withdraw at any time. Your usual treatment and follow up will not be affected should you chose not to take part. Participation is entirely voluntary.

What happens if I chose to take part?
You will meet with one of the research team to arrange suitable times for you to come for the studies, on this initial meeting blood tests will be taken from a vein in your arm which will be about 10-20ml (2-4 teaspoons). We will then arrange for you to come for the lung function tests, echocardiography studies and exercise tests.

The echocardiography studies take up to 1 hour. We will ask you to lie on a bed and we will attach several pieces of equipment including ECG leads. We will take images of your heart using an echo scanner. During the same visit we will arrange for you to have lung function testing which will also take about 1 hour. This involves controlled breathing into a machine under the instruction of a professional to record the ability of your lungs.

You will also undergo exercise testing on multiple visits (usually 2). This will involve cycling on an exercise bicycle with a mouthpiece in. This records your breathing and the amount of oxygen or carbon dioxide you breathe in and out. You are also attached to an ECG machine and have regular blood pressure readings.
You will be expected to cycle until you are unable to continue. The bicycle will get progressively harder to cycle. We aim for the length of time for you to be cycling to be 10-15 minutes, total time will be less than 1 hour. You decide when you have cycled enough and you are unable to continue, the test stops when you stop cycling.

Expenses and Payments
Any travel and costs for meals will be reimbursed up to an amount of £50 per person and will vary based on individual expenses.

What are the side effects of taking part?
There are no drugs or invasive tests. We will require blood tests during your first visit, and rarely once more after then. Some people find having echocardiography uncomfortable. If this is the case we will take breaks and / or stop as required.

Occasionally, the blood pressure monitor can cause tingling on the finger. This goes away when removed.
A cardiopulmonary exercise test is tiring, and it is expected that you will be out of breath at the end. We will allow you plenty of time for recovery, and supply you with drinks and food as you need. The mouthpiece causes some dribbling, we will cover your clothes with an apron if required.
What if something goes wrong?

We will take every care in the course of this study, and the risk involved is minimal. Imperial College London holds insurance policies which apply to this study. If you experience serious and enduring harm or injury as a result of taking part in this study, you may be eligible to claim compensation without having to prove that Imperial College is at fault. This does not affect your legal rights to seek compensation.

If you are harmed due to someone’s negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been treated during the course of this study then you should immediately inform the Investigator, Dr Anthony Barron (07932 648274) or Dr Roland Wensel. The normal National Health Service complaint mechanisms are also available to you. If you are still not satisfied with the response, you may contact the Imperial AHSC Joint Research Office, Imperial College London, St Mary’s Hospital, 41 Praed Street, Paddington, London, W2 1NY (02033126484).

What are the benefits?

The research may not benefit you directly. However, we will have improved knowledge regarding the best possible use of this test for the future management of patients with a variety of heart conditions. This may impact on who and when patients undergo procedures and operations to the heart. This may prevent unnecessary treatments.

Will my taking part in this study be kept confidential?

Data collected about you will be entered on to a computer, in an anonymous format and any information taken from this investigation that can be identified will remain confidential.

If you agree, we would contact your GP to inform them you are participating and will provide them with information regarding the study.

At times during the study professional bodies (for example the NHS Research and Development team or Imperial College) may wish to audit our data to ensure that the study is being carried out appropriately. Any non-essential information regarding your involvement will not be shown. These auditors will be appropriate professionally trained individuals and all information will be kept confidential. Our research is in collaboration with experts in Germany, they will be given access to the results of your tests, but again your personal details will be kept anonymous.

What happens once the research stops?

The results may be presented at meetings and published so that it can be used to help treatment of other patients with heart failure, lung disease, arrhythmias and valve disease.

Your participation will be kept strictly confidential.

Should you wish to receive information regarding the outcomes of the study, this can be sent to you at the conclusion of the study.

Who has reviewed the study?

This study has been reviewed by the South East London Research Ethics Committee.

If you have any further questions please do not hesitate to contact:

Dr Anthony Barron on 07932 648274

Thank you for taking the time to consider participating in this study.
13.3 Consent form

CONSENT FORM

Title of Project: Identifying an Ideal Cardiopulmonary Exercise Test Parameter
Chief Investigator: Dr Roland Wensel

Name of Researcher: Dr Anthony Barron

Please tick box

1. I confirm that I have read and understand the information sheet version 1 and dated 01/03/2010 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my medical care or legal rights being affected.

3. I understand that relevant sections of any of my medical notes and data collected during the study, may be looked at by responsible individuals from Imperial College Healthcare NHS Trust, Imperial College and Regulatory bodies, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.

4. I understand the risks and benefits of the study and have had any questions answered. I am aware of who to contact should I have any further queries.

5. I agree to my GP being informed of my participation in the study.

Centre Number: 
Study Number: 
Patient Identification Number for this trial:
6. I agree to my personal data being stored safely for a total of 10 years after completion of the study and for further contact and invitation to future studies (at my discretion for inclusion at this time) only by the host institution (Imperial College and Imperial College Healthcare NHS Trust). Otherwise it will be destroyed at the end of the 3 years.

7. The compensation arrangements have been discussed with me.

8. I agree to take part in the above study.

____________________  __________________  __________________
Name of Patient        Date                  Signature

____________________  __________________  __________________
Name of Person taking consent (if different from researcher) Date                  Signature

____________________  __________________  __________________
Researcher             Date                  Signature
13.4 Research Ethics Committee acceptance

South East London REC 5
South London REC Office (5)
Ranken House
Queen Elizabeth Hospital
Stadium Road
Woolwich
London
SE18 4QH

Tel: 0208 8366740
Fax: 0208 8364862

19 August 2010

Dr Roland Wensel
Consultant Cardiologist
International Centre for
Circulatory Health, Imperial College
59 North Wharf Road, London
W2 1LA

Dear Dr Wensel

Study title: Identifying the Ideal Parameter of the Cardiopulmonary Exercise Test to Distinguish Between the Cardiovascular and Respiratory Components of Functional Limitation and to detect Relevant Physiological Changes in Function

REC reference: 10/H0805/36
Protocol number: N/A
Amendment number: 1
Amendment date: 05 August 2010

The above amendment was reviewed at the meeting of the Sub-Committee held on 17 August 2010 by the Sub-Committee in correspondence.

Ethical opinion

Favourable opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

<table>
<thead>
<tr>
<th>Document</th>
<th>Version</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notice of Substantial Amendment (non-CTIMPs)</td>
<td>1</td>
<td>05 August 2010</td>
</tr>
<tr>
<td>Covering Letter</td>
<td></td>
<td>05 August 2010</td>
</tr>
<tr>
<td>IRAS project filter questions</td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.
R&D approval

*All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.*

Statement of compliance

*The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.*

<table>
<thead>
<tr>
<th>10/H0805/36:</th>
<th>Please quote this number on all correspondence</th>
</tr>
</thead>
</table>

Yours sincerely

**Esther Richman**  
Committee Co-ordinator

E-mail: estherrichman@nhs.net

**Enclosures:**  
List of names and professions of members who took part in the review

**Copy to:**  
Lucy Parker, Imperial College

**South East London REC 5**

**Attendance at Sub-Committee of the REC meeting on 22 August 2010**

<table>
<thead>
<tr>
<th>name</th>
<th>Profession</th>
<th>position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carol Jones</td>
<td>Management Consultant</td>
<td>Chair</td>
</tr>
<tr>
<td>Niall MacCrae</td>
<td>Health service researcher</td>
<td>Vice-chair</td>
</tr>
</tbody>
</table>
13.5 Permission for use of figure

September 10, 2007
Anthony Barron

**Journal Name Year Citation Item(s) used**
Arch Intern Med 1988 148:2221-2226 Figure

**Intended** Material will be used in a paper on cardiopulmonary exercise testing.

**Permission Granted:**
Thank you for your interest in JAMA, Archives and AM News. Rights granted herein are non-exclusive and limited to one time only reproduction as specified in this request, in printed format in the English Language. Your credit line must include the name of the publication, issue date, volume and page number, as well as “Copyright © (Year of Publication), American Medical Association. All Rights reserved.”

Best wishes,

Rhonda Bailey Brown
Permission Assistant
Publishing Operations
AMA
14.0 Publications

Reduced confounding by impaired ventilatory function with oxygen uptake efficiency slope and VE/VCO\textsubscript{2} slope rather than peak oxygen consumption to assess exercise physiology in suspected heart failure.
Congest Heart Fail. 2010 Nov-Dec;16(6):259-64.

The role for cardiopulmonary exercise testing in patients with atrial septal defects: a review.
Int J Cardiol. 2012 Nov 15;161(2):68-72

First-in-man safety evaluation of renal denervation for chronic systolic heart failure: primary outcome from REACH-Pilot study.
Int J Cardiol. 2013 Jan 20;162(3):189-92
15.0 Abstracts

Test-retest repeatability of cardiopulmonary exercise test parameters in patients with cardiac or respiratory disease.

Barron AJ, Dhutia N, Hughes AD, Francis DP, Wensel R.
European Society of Cardiology Heart Failure 2013 Congress

Red Cell Distribution Width Relates to Exercise Capacity in Chronic Heart Failure but Provides Prognostic Information Incremental to Peak Oxygen Consumption.

American Heart Association 2011 Congress

Attenuations in Serum Albumin Over Time Powerfully Predict an Amplified Risk of Death in Chronic Heart Failure.

American Heart Association 2011 Congress