Developing Markers of Mechanical Dyssynchrony in Heart Failure: Implications for Research and Clinical Practice in Cardiac Resynchronisation Therapy.

Punam Ashok Pabari

MB ChB, MRCP

A thesis submitted to Imperial College London

for the degree of

PhD

Imperial College of Science, Technology and Medicine

International Centre for Circulatory Health

59-61 North Wharf Road

London

W21LA

Originally submitted: January 2012

Resubmitted: July 2012

Word Count: 56, 941

Declaration

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

Punam Ashok Pabari

Abstract

This thesis studies techniques for maximising the effects of Cardiac Resynchronisation Therapy (CRT) which is a modern therapy for chronic heart failure.

The ability of echocardiographic parameters of mechanical dyssynchrony to identify suitable candidate patients for CRT, and subsequently to optimise the interventricular delay, is disputed. In this study, I initially perform a systematic review to clarify the nature and extent of this problem. I then perform detailed mathematical simulations to understand what is feasible in an ideal setting.

I then evaluate the realistic potential of several non-invasive approaches of optimising CRT devices by formal evaluation head-to-head. A fundamental requirement for any marker in optimisation of CRT devices, or in selection of patients for implantation, is that the marker must be reproducible and must change when the amount of dyssynchrony changes.

I perform detailed experiments in patients where I interrogated a panel of echocardiographic parameters to answer the questions relating to the sensitivity of each parameter and looked at methods to improve this. I look at the effect of spontaneous variability, and the impact on each echocardiographic method including 3D echocardiography, tissue Doppler imaging and pulsed Doppler techniques including velocity time integrals and pre-ejection times.

Finally, I perform an invasive study of the beat to beat physiological changes which occur after intra cardiac timings are altered, to evaluate alternative approaches to changes in cardiac performance that might be easier to automate.

Dedication

I dedicate this thesis to my parents, Ashok and Shobhna Pabari, who have been pillars of strength and support throughout my career. Their unwavering faith in me and continued encouragement has motivated me and enabled me to achieve my goals. I thank my sisters, Chandani and Roshni, who have always offered me help and support over the years.

Finally, I thank my husband, Jay Chhasatia, for his love, patience and confidence in me without which I would not have had the ability to succeed.

Acknowledgements

I have had the privilege of working alongside many great academics throughout this project. I owe heartfelt thanks to several people who have given me tremendous support over the past years, both intellectually and emotionally, and have worked tirelessly to help me with this research project.

My supervisor, Dr Darrel Francis who has believed in me, encouraging and motivating me from beginning to end. His enthusiasm and passion for clinical research have been inspirational and the precious time he sacrificed, tirelessly working with me to discuss ideas and critique my work has been invaluable.

Prof Jamil Mayet, who has provided sound advice and discussed many aspects of my project despite his busy clinical schedule.

Prof Alun Hughes, whose kindness and patience has been much appreciated.

I am grateful to my fellow research colleagues, physiologists and other members of the team who have given up time and energy despite having their own projects to complete. My warmest thanks to:

Dr Andreas Kyriacou, who has assisted me with experiments, data collection and been a supportive friend and colleague throughout.

Michela Moraldo who has practically assisted with experimental collection and analysis, plus been a true friend who has supported me emotionally throughout the up's and down's of completing a PhD.

Most importantly, thank you to the heart failure patients, who, despite their frailty, have given up time and energy to participate in these studies and contributed to the knowledge we have gained.

Finally, I would like to thank the British Heart Foundation for their generous financial support which has made this study possible.

Contents

Developing Markers of Mechanical Dyssynchrony in Heart Failure: Implications for Research and Clinical Practice in Cardiac Resynchronisation Therapy. 1

Declaration 2

Abstract 3

Dedication 4

Acknowledgements 5

Contents 7

List of Figures 14

List of Tables 18

List of Tables 18

List of Equations 20

List of Equations 20

Introduction 21

Introduction 21

1.1 Origin of the dyssynchrony concept 22

1.2 Development of criteria for selecting patients – problems and solutions 22

1.3 The PROSPECT Crisis 24

1.4 The next step… 25

1.5 The Problem 27

1.6 The different families for measuring mechanical dyssynchrony 27

1.6.1 Measurements focusing between the right and left ventricle 28

1.6.2 Measurements focusing within the left ventricle 28

1.6.3 Final pathway echocardiographic measurements for optimisation 29

1.7 How do we evaluate a proposed marker for assessing dyssynchrony and guiding optimisation? 30

1.8 Aims of this thesis 31

1.8.1 To evaluate and understand the current literature on cardiac resynchronisation therapy. 31

1.8.2 How can we ensure protocols for optimisation are not a waste of time, but in fact reliable? 31

1.8.3 Are echocardiographic markers plausible candidates for measuring dyssynchrony that can be measured by CRT? 31

1.8.4 Assess the effect of changing atrioventricular delay on invasive flow, invasive pressure and non-invasive pressure. 32

2 Methodology 33

2.1 Patient recruitment 33

2.2 Equipment 33

2.2.1 Echocardiography 33

2.2.2 Transducers....34
2.2.2.1 Transthoracic Probe 2D....34
2.2.2.2 Transthoracic probe 3D....34

2.3 Echocardiographic measurements used throughout the studies for optimisation processes. 36
2.3.1 Doppler assessment of VV delay....36
2.3.1.1 Ejection flow dyssynchrony – acquisition....36
2.3.1.2 Ejection flow dyssynchrony – analysis....37
2.3.1.3 Velocity-time Integral – acquisition....39
2.3.1.4 Velocity-time Integral – analysis....39

2.4 Segmental assessment of VV delay optimisation....40
2.4.1 3D echocardiography....40
2.4.1.1 3D echocardiography- acquisition....42
2.4.1.2 3D echocardiography- analysis....43
2.4.2 Tissue Doppler Imaging- technology....48
2.4.2.1 Tissue Doppler Imaging-acquisition....48
2.4.2.2 Tissue Doppler Imaging-analysis....48

2.5 QRS duration as a marker for optimisation of CRT....51

2.6 Pacemaker programmers....51

2.7 Invasive study....53
2.7.1 Invasive study - acquisition....53
2.7.2 Invasive pressure and flow measurements....53
2.7.3 Non-invasive blood pressure measurements - Finometer technology....54
2.7.4 Surface ECG....57
2.7.5 Acquisition system....58
2.7.6 Analysis software....58

2.8 Exclusion of ectopic beats....60

2.9 Methods of Analysis – Acquiring data....63
2.9.1 Single measurements....63
2.9.2 Averages of multiple measures....63

2.10 Calculation of Intraclass Correlation Coefficient....65

3 Limits to predictability of response to biventricular pacing from dyssynchrony indices, and systematic analysis of impact of study design on findings....67

3.1 Abstract....68
3.1.1 Background....68
3.1.2 Method and Results....68
3.1.3 Conclusions....69

3.2 Background....70

3.3 Methods....73
3.3.1 Quantitative separation of device-mediated, versus spontaneous, changes in LVEF....73
3.3.2 Data extraction from published studies....79
3.3.3 Data from my hospital's database assessing ejection fraction at repeated studies....82
3.3.4 Medcon search utility....83
3.3.4.1 Part I: Create report layout....83
3.3.4.2 Part II: Create search....83
3.3.5 Statistics....84

3.4 Results....85
3.4.1 Reported R^2 for echocardiographic response in EMRCTs and HSSCs....85

3.5 Impact of R^2 contraction factor arising from the response variable: the mandatory statistical ceiling on observable R^2 values....86
3.5.1 Impact of R^2 contraction factor arising from the dyssynchrony variable....90

3.6 True limits to correlation between baseline and delta response: Combined contraction factor from both sources of variability *92*
3.6.1 Echocardiographic data from my local hospital database 94

3.7 Discussion *99*
3.7.1 The fundamental ceiling on observable R^2 99
3.7.2 Minimising contraction in R^2 values 100
3.7.3 Opportunities and limits of replicate averaging 102
3.7.4 Why some HSSCSs report higher R^2 than EMRCTs (and higher than mathematically sustainable limits) 104
3.7.5 Clinical Implications 109

3.8 Limitations *111*

3.9 Conclusions *111*

4 The Puzzle of Optimisation: When is an optimisation not an optimisation? 113

4.1 Abstract *114*
4.1.1 Background 114
4.1.2 Method and Results 114
4.1.3 Conclusions 115

4.2 Background *116*
4.2.1 Information content 119

4.3 Published data *120*

4.4 Methods *123*
4.4.1 Observed measurement = underlying signal + superimposed noise 123
4.4.2 Simulation 123
4.4.3 Identification of optimum 124
4.4.4 Confidence intervals of the optimum 124

4.5 Results *126*
4.5.1 Impact of information content on consistency of detecting optimum, using a single beat at each setting 126
4.5.2 Size of confidence interval of the observed optimum 129

4.6 Impact of averaging multiple replicates on reproducibility *129*
4.6.1 Apparent versus true size of improvement on optimisation 132
4.6.2 Apparent change in optimum over time 135

4.7 Discussion *137*
4.7.1 Illusion 1: “We have selected the true underlying optima” 137
4.7.2 Illusion 2: “The optimisation increased flow (or pressure) by X and was therefore worthwhile” 138
4.7.3 Illusion 3: “The optimum has changed between X months and now” 139
4.7.4 Illusion 4: “We should not waste time making multiple replicate measurements at each setting in clinical practice” 139
4.7.5 Illusion 5: “We should optimise using whatever measurement method we are most familiar with” 140
4.7.6 Illusion 6: “Between separate beats, variability in my laboratory is only X%, therefore this measure is suitable for use in optimisation” 141

4.8 A simple method of calculating information content of a cardiovascular measure used clinically for optimisation *142*

4.9 How many replicates are really needed in clinical practice? *144*

4.10 Clinical Implications *148*
4.10.1 Practical Recommendations 150

4.11 Conclusions *151*

5 Validating markers of mechanical dyssynchrony by experimental manipulation of interventricular timings: What is needed to make them a reasonable prospect for Cardiac Resynchronisation Therapy selection? 152

5.1 Abstract 153
5.1.1 Background 153
5.1.2 Methods and Results 153
5.1.3 Conclusions 154

5.2 Background 155

5.3 Methods 157
5.3.1 Study Population 157

5.4 Protocol 158

5.5 Echocardiography 158
5.5.1 Modalities used for optimising VV delay 159
5.5.1.1 LVOT – Velocity Time Integral 159
5.5.1.2 Ejection flow dyssynchrony 159
5.5.1.3 Tissue Doppler imaging 159
5.5.1.4 3D Echocardiography: Systolic dyssynchrony index 160
5.5.2 Calculating the agreement of selecting the optimal VV delay setting within the same modality 163
5.5.3 Calculating Kappa within the same modality 164

5.6 Calculating the scatter of detected optima between repeated optimisations for each separate modality 165

5.7 Quantification of the genuine effect of interventricular delay versus random variability 165

5.8 Results 166
5.8.1 Assessment of each of the 6 dyssynchrony markers as candidates for VV optimisation using single measurements 166
5.8.1.1 Probability of selecting the same optimum on replicate optimisations 166
5.8.1.2 Probability of two different optimisation methods agreeing on an optimum 168
5.8.1.3 The Scatter of VV optima obtained at repeated optimisations 170
5.8.1.4 Quantification of genuine change versus biological variation: Intraclass correlation coefficient 172

5.9 Multiple Replicate Measurements 174
5.9.1 The effect of multiple measurements on the scatter of VV optima obtained 174
5.9.2 The effect of multiple measurements on the intraclass correlation coefficient 176
5.9.3 Responders and Non-responders 178

5.10 Discussion 180
5.10.1 Why this study design, rather than any other, to assess a dyssynchrony marker? 180
5.10.2 Can anything be done to improve the situation? 182

5.11 What do these results mean for clinical VV delay optimisation? 182
5.11.1 Why are our results “worse” than commonly supposed 185
5.11.2 Would more training help? 186

5.12 Limitations 187

5.13 Conclusions 188

6 Rapid method for evaluating and improving plausibility of tissue Doppler protocols for assessment of dyssynchrony in patients being considered for cardiac resynchronisation therapy 189

6.1 Abstract 190
6.1.1 Background 190
6.1.2 Methods and Results 190
6.1.3 Conclusions 191

6.2 Background 192

6.2.1 Which tissue Doppler model is correct for assessing dyssynchrony? 194

6.3 Methods 195
6.3.1 Tissue Doppler dyssynchrony assessment 195
6.3.1.1 2, 6 or 12-segment model 195
6.3.1.2 Earliest to latest or standard deviation of wall timings 195

6.4 Study population 196

6.5 Measurements 197
6.5.1 Image Acquisition 197
6.5.2 Image Analysis 198

6.6 Results 200

6.7 Using single measurements versus averages of paired measurements 200
6.7.1 Measuring to onset of S wave or to peak of S wave. 205
6.7.2 Comparison between differences of absolute values and standard deviation calculations of the timings of 6 and 12 segment models of tissue Doppler analysis 209
6.7.3 2 versus 6 versus 12 segments 211

6.8 Discussion 213
6.8.1 Beat-to-beat variability taboo 213
6.8.2 Inability to detect resynchronisation when it occurs 214
6.8.3 Implications for protocol designs and reproducibility 216

6.9 Limitations 217

6.10 Conclusions 218

7 Comparison of the reproducibility of left ventricular outflow tract velocity time integral (LVOT VTI) versus peak velocity (V_{max}) as targets for optimisation of VV delay in cardiac resynchronisation therapy. 219

7.1 Abstract 220
7.1.1 Background 220
7.1.2 Methods 220
7.1.3 Results 220
7.1.4 Conclusions 221

7.2 Background 222

7.3 Methods 226
7.3.1 Study Population 226
7.3.2 Protocol 227
7.3.3 Measurements 227

7.4 Analysis 227
7.4.1 Determination of Optimal VV delay 228
7.4.2 Calculation of the effect of averaging replicate optimisation sequences on the scatter of identified optima 228
7.4.3 Calculation of Intraclass Correlation Coefficient 232
7.4.4 Calculation of within-patient agreement of replicate optimisations 232
7.4.5 Calculation of between method agreement of VV optimisation 233
7.4.6 Calculation of time taken to perform analysis of VTI and V_{max}. 233
7.4.7 Statistical Analysis 233

7.5 Results 234
7.5.1 Relative characteristics of VTI based optimum versus V_{max} based optimum 234
7.5.2 Consequences of averaging replicate measurements 234
7.5.3 Within-patient agreement of VV optima obtained on replicate optimisations 236
7.5.4 Between method agreement of VV optimisation 236
7.5.5 Responders and non-responders 238
Time taken to perform optimisation analysis 239

7.6 Discussion 240
7.6.1 How can an incomplete picture be better? 241

7.7 The importance of assessing signal to noise ratio when planning optimisation algorithms ……242
7.7.1 Averaging multiple measurements 243
7.7.2 How to have confidence in the method of optimisation 243

7.8 Limitations....244

7.9 Conclusion....245

8 Invasive pressure and flow measurements during AV delay improvement reveal a compensatory peripheral vasodilator response which attenuates the initial blood pressure increment: implications for the design of AV optimisation protocols....246

8.1 Abstract 247
8.1.1 Introduction 247
8.1.2 Methods and Results....247
8.1.3 Conclusion....248

8.2 Background 249

8.3 Methods....252
8.3.1 Protocol design elements for reliability of results 252
8.3.1.1 Invasive acquisition**.**....252
8.3.1.2 Signal and noise. 252
8.3.1.3 Minimising noise: replication and selection of variables**.** 252
8.3.2 Maximising signal: only two AV delays. 253

8.4 Study population....254

8.5 Measurements....255
8.5.1 (a) Invasive measurements 255
8.5.2 (b) Non-invasive blood pressure measurements 255
8.5.3 Pacing protocol....256
8.5.4 Total Peripheral Resistance 257
8.5.5 Data Acquisition and Analysis 257
8.5.6 Statistics....257

8.6 Results 258
8.6.1 Patient characteristics 258

8.7 Phases of Response to change in AV delay from 40 to 120 ms....258
8.7.1 Comparison between invasive and non-invasive blood pressure measurements following the change in AV delay from 40-120ms 261
8.7.2 Reverse transitions: 120 ms to 40 ms 264
8.7.3 Mechanisms: Delayed change in Total Peripheral Resistance....267

8.8 Discussion 269
8.8.1 Relevance to quantitative optimisation of CRT....269

8.9 Immediate and delayed effects of changes in pacemaker settings....270
8.9.1 Is AV delay optimisation dead in the wake of SMART AV?....273

8.10 Limitations....274

8.11 Conclusions 275

9 Synthesis 276

9.1 The puzzle of dyssynchrony and response 276

9.2 The puzzle of optimisation....278

9.3 Biological variability of optimisation using a spectrum of echocardiographic parameters ……281

9.4 The effect of AV delay optimisation on invasive pressure and flow measurements: implications for the design of AV optimisation protocols....283

9.5 Recommendations for future studies and clinical implementation....288

10 Appendices 291

10.1 Appendix 1: Invitation letter 291

10.2 Appendix 2: Patient information leaflet 292

10.3 Appendix 3: Consent form 296

10.4 Appendix 4: Determining the contraction factor 297

11 References 300

List of Figures

Figure 2-1: Photograph of the Philips IE33 machine used for data acquisition (top panel), with the 2D probe (middle panel) and the 3D probe (bottom panel).....35

Figure 2-2 Measuring aortic pre-ejection time. 38

Figure 2-3 An example of a velocity time integral tracing of a pulsed wave Doppler image. 39

Figure 2-4 Segmental division of the left ventricle. In the bottom left corner we can see the ventricle divided up into 16 segments (plus apex), each of which corresponds to a different coloured section of the ventricle. 41

Figure 2-5 An example of drawing around the endocardium which is shown here for the 4 chamber view. This is repeated for 2 chamber and 3 chamber, which then results in the 3D reconstructed ventricle. 44

Figure 2-6 The dyssynchrony curves of a normal synchronous ventricle as seen when analysis is performed. All segments reach the time to minimum volume at a similar time. An SDI value is then calculated. 46

Figure 2-7 Schematic diagram showing the segmental analysis with each coloured line representing a segment of the ventricle. 47

Figure 2-8 An example of a TDI trace where the region of interest is the basal septum which is where the cursor is positioned in the top panel. 50

Figure 2-9 A photograph of the Boston Guidant pacemaker used in the study (left panel) and a close up of the programmer screen (right panel) as programmed during one of my experiments. 52

Figure 2-10 A Finometer unit (top panel) which has the connections at the front as described: the main unit is strapped to the back of the patients arm with the cuff around the index finger (bottom panel)...55

Figure 2-11 An example of the ECG acquisition which occurs simultaneously with the invasive measurements 57

Figure 2-12 A snap shot of the data acquisition system on Labview where all data is recorded.....59

Figure 2-13 showing ectopic and a normal beat with VTI measured. We can see the final beat in the frame (on the right) is a normal beat following the ectopic. The difference in VTI is visibly noticeable. 60

Figure 2-14 An example of an ectopic beat whilst recording ECG, invasive blood pressure and non invasive blood pressure. 61

Figure 2-15 Schematic diagram showing how we measured single beats, average-of-two and averages-of-three to ensure each beat was independent. We then calculated the standard deviation from the optimum derived. 64

Figure 2-16 Intuitive description of intraclass correlation coefficient, as applied to dyssynchrony measures. 66

Figure 3-1 The variance of the intervention group is only partly due to the intervention. 74

Figure 3-2 How inherent variability in two measures reduces the maximum achievable R^2 between them. 77

Figure 3-3 Systematic search strategy 80

Figure 3-4 Step by step method to search Medcon to perform searches for echocardiograms as performed in this study 83

Figure 3-5 Comparison of highly skilled single centre studies (HSSCSs) and externally monitored randomised controlled trials (EMRCTs) for reported R^2 values between baseline dyssynchrony markers and echocardiographic markers of response to biventricular pacing 108

Figure 4-1: An example of clinical data from typical patient undergoing three separate Doppler optimisation processes (#1, #2 and #3) a few minutes apart, using one heartbeat of velocity-time integral as the measurement to be maximised. In this patient, the small differences in velocity-time integral between the three optimisation processes are enough to cause different AV delay settings to be identified as apparently optimal on the three occasions. 117

Figure 4-2 Observed measurements are composed of underlying true difference between settings ("signal", top panel) and beat-to-beat variability ("noise", middle panel) which may be small (left) or large (right) relative to the signal. 125

Figure 4-3 Scatter between successive optima increases when noise variance is increased (Top panel) and decreases when signal magnitude is increased (Middle panel). Information content, encompassing the relative sizes of signal and noise, has a powerful effect on the scatter between successive optima. The bottom panel shows the effect of information content (the proportion of variance that arises from signal) on the scatter between successive optima. 127

Figure 4-4 Impact of switching from single measurements to average of multiple measurements for the optimisation process. 131

Figure 4-5 Relation between variability and apparent benefit of optimisation. 133

Figure 4-6 Impact of information content on the probability of falsely detecting a change in optimum. 136

Figure 4-7 Calculation of information content using raw clinical data from a single patient. 143

Figure 4-8 The effect of heart rate on information content from datasets of published studies regardless of method used for AV optimisation 146

Figure 5-1 A shell is created with all segments numbered. 162

Figure 5-2 Agreement between replicate VV optimisations by the same method. Inset shows method of calculating this value for 3D SDI. 167

Figure 5-3 Example of how agreement between methods was calculated (3D Systolic dyssynchrony index versus Aortic pre ejection time) 169

Figure 5-4 Calculation of scatter of optimum VV delays in a representative patient using repeat measurements of aortic pre-ejection time as the marker of dyssynchrony. 171

Figure 5-5 Six replicate measures at each VV delay using aortic pre-ejection time as the dyssynchrony marker. 173

Figure 5-6 The scatter of VV optima obtained using each dyssynchrony parameter at rest with single measurements and averaging paired replicate measurements 175

Figure 5-7 The Intraclass correlation coefficient of each modality for single and averaged paired measurements. 177

Figure 5-8: The intra class correlation coefficient between responders and non-responders showed no significant difference across all modalities 178

Figure 5-9: The scatter of VV optima obtained when replicate optimisations are performed showed no significant differences between responders and non-responders across all modalities tested 179

Figure 6-1 Demonstration of measuring from onset of QRS to onset of S wave and from onset of QRS to peak of S wave 198

Figure 6-2 showing a tissue Doppler image of the 4 chamber view, with the basal septal wall highlighted. 199

Figure 6-3 The ICC increases significantly across all patients and all measurement methods, when averages of paired measurements are used in comparison to single measurements. 202

Figure 6-4 The scatter of optima is significantly narrower when averages of pairs of measurements are used compared to single measurements. 204

Figure 6-5 The width of 95% confidence intervals which will surround the optimum detected are shown above for each possibility of tissue Doppler measurements used. The graphical representation

shows the dramatic decline in this width when averages of pairs are used. Numerically we can see the smallest width with 12 segment earliest to latest using paired measurements. 215

Figure 7-1: An example patient comparing the scatter of the values for optimal VV delay between velocity time integral and peak velocity. 231

Figure 7-2 The difference in the VV optimum obtained between VTI and V_{max} across all patients. In 58% of cases the optimum was the same for both VTI and V_{max}. For 76% of cases the optimum by VTI and Vmax were within 20ms of each other. It should be noted that "± 20 ms" permits, in most cases, the second optimum to be any of 3 of the 6 tested settings. 237

Figure 7-3 Time taken to analyse single measures, 2's and 3's as used for the single, averages-of-two and averages-of-three calculations. 239

Figure 8-1 Beat-to-beat changes in invasive blood pressure when AV changed from 40 ms to 120 ms across all patients. After the transition the four phases are shown: (1) the first beat, (2) the early plateau (beats 2-11), (3) a progressive decline and (4) the late plateau (beats 20-80). 259

Figure 8-2 Change in invasive flow, shown beat-by-beat, as AV delay is changed from 40 to 120ms across all patients.....260

Figure 8-3 Beat-to-beat data for invasive and Finometer (non-invasive) systolic blood pressure across all patients during a programmed change in AV delay from 40 to120 ms, averaged across all subjects. 262

Figure 8-4 Relationship between change in beat-to-beat invasive SBP and Finometer SBP. 263

Figure 8-5 Beat-to-beat changes in invasive blood pressure when AV changed from 120 to 40 ms across all patients. There is a drop in invasive blood pressure to −14.4 mmHg ±2.3 mmHg below baseline (p=0.0004) in the period from beats 2 to 11 265

Figure 8-6 Change in invasive flow, shown beat-by-beat, as AV delay is changed from 120 to 40ms across all patients.....266

Figure 8-7 Beat-to-beat changes in Total Peripheral Resistance across all patients for change in AV delay from 40ms to 120ms (top panel) and 120ms to 40ms (bottom panel).....268

Figure 8-8 Asymmetrical impact of AV-lengthening and AV-shortening. 272

List of Tables

Table 3-1 A comparison of the baseline dyssynchrony variables found to predict response in externally monitored randomised controlled trials with those found highly skilled single centre studies. Correlation coefficients and coefficient of determination (R^2) are presented. 86

Table 3-2 Calculation of the contraction factor for ΔLVEF, ΔLVESV and ΔLVEDV in externally monitored randomised controlled trials assessing biventricular pacing. The maximum possible R^2 for any parameter to the ΔEF in each study is presented for each study. 87

Table 3-3 R^2 contraction factor for non-echocardiographic markers of response to biventricular pacing. The maximal achievable R^2 column shows the maximum R^2 that any predictor could find when correlated against any of these outcome measures.. 89

Table 3-4 Test-retest variability of dyssynchrony markers within individuals, in populations who are candidates for biventricular pacemaker implantation. Most studies did not report test-retest reproducibility. 91

Table 3-5 The true limit to the observed R^2 for the correlation between ΔLVEF and dyssynchrony markers is a product of their two R^2 contraction factors. 93

Table 3-6 The distribution of EF for patients who had 2, 3, 4, 5 or 6 scans performed. 95

Table 3-7 The scatter (SD) between the first scan and each subsequent scan. This shows that the scatter does not change significantly whether the first scan is compared to the 2^{nd} or to the 6^{th}, hence the passage of time is not a factor. Instead it is the variability of measurement which is causing the wide scatter (SD).....96

Table 3-8 The scatter (SD) of ΔEF between immediately successive scans is shown below.....98

Table 3-9 Differences in study design between highly skilled single centre studies (HSSCS) and externally monitored randomised control trials (EMRCTs). 107

Table 4-1 : Information content calculated from published studies of optimisation (Auricchio et al. 1999b; van Geldorp et al. 2011; Whinnett et al. 2006a).....121

Table 4-2 Effect of signal and noise on the information content, and on widths of 95% confidence limits.....130

Table 4-3 Impact of information content on the size of the apparent benefit of optimisation.....134

Table 4-4 The number of replicates required when optimisation is performed to reduce the scatter of AV optima obtained to a range of acceptable confidence intervals.....145

Table 5-1 Calculation of the number of replicates required to provide reliable optima with a 95% confidence internal of ± 10ms, for each method. 184

Table 6-1 Across all patients there was a statistically significant improvement in the intraclass correlation coefficient when averages of paired measurements were used compared to single measurements. 201

Table 6-2 The scatter of the optima obtained across all patients decreases significantly when averages of paired measurements are used compared to single measures. 203

Table 6-3 Intra class correlation coefficients of measuring to the onset of systolic wave and to the peak of systolic wave. There is no significant difference in the intra class correlation whether you measure to the onset of systolic wave or to the peak. 206

Table 6-4 Scatter of optimum measuring to the onset of systolic wave and to the peak of systolic wave. 208

Table 6-5 Measuring earliest to latest versus standard deviation between wall timings. The Intraclass correlation coefficient and the scatter of the optima are not significantly different when the earliest to latest timings are used, or whenever the standard deviation of timings are used for 6 or 12 segment TDI models. 210

Table 6-6 The effect of measuring 2, 6 or 12 segments for TDI analysis..... 212

Table 7-1: Data from one patient showing all raw measurements of VTI from 6 separate optimisation sequences. This demonstrates how we averaged raw data as pairs and triplets, and calculated the between optima scatter (SD)..... 229

Table 7-2 The scatter and intraclass correlation coefficient for both V_{max} and VTI..... 235

List of Equations

Equation 3-1 R^2 contraction factor caused by inherent variability in LVEF, where σ is standard deviation and σ^2 is variance. 75

Equation 3-2 The combined contraction factor which takes into consideration the R^2 contraction factor due to the dyssynchrony marker and the response marker 76

Equation 4-1 119

Equation 4-2 142

Equation 5-1 164

Equation 5-2 164

Equation 5-3 164

Introduction

This thesis addresses the mechanistic question and practical challenge for the contemporary management of heart failure. Heart failure is a growing burden in the developed world, with an overall prevalence of 2-3% which rises further after 70 years of age (Dickstein et al. 2008). The number of people living with heart failure will continue to increase due to a greater number of people surviving myocardial infarctions and an aging population (Cowie et al. 1999). Developments in the management of heart failure have led to treatment of these patients moving progressively from drug treatments alone, i.e. ACE inhibitors (1987; Flather et al. 2000), beta blockers (Bristow et al. 1996), angiotensin II receptor blockers (Pitt et al. 2000) and aldosterone antagonists (Pitt et al. 1999; Pitt et al. 2001); to the addition of cardiac resynchronisation devices (Bristow et al. 2004; Cazeau et al. 1994; Cleland et al. 2005) and now the beneficial effects of these devices can be maximised by fine tuning the different programmable settings (Duvall et al. 2010; Ellenbogen et al. 2010; Perego et al. 2003; Vanderheyden et al. 2005; Whinnett et al. 2006a)

In this thesis I have studied the evolving sciences of dyssynchrony assessment and optimisation of resynchronisation devices. I seek to understand whether the extensive claims in these areas should be accepted at face value. Improving ability to select patients for CRT by identifying those who are likely to respond would be clinically valuable if true. I have studied this in a critical quantitative literature review in Chapter 3. Once implanted, ensuring that we are getting the most from the device by changing the atrioventricular (AV) or ventriculo-ventricular (VV) delay is important *however* as described in Chapter 4, it is essential that we have a tool which is able to detect the small changes which occur when intra-cardiac settings are altered in a reproducible manner.

1.1 Origin of the dyssynchrony concept

In normal subjects, the ventricular walls contract together simultaneously – easily observable be eyeball alone (Holloway et al. 2011). In heart failure however, it is equally evident by eyeball observation alone that some patients have very disrupted ventricular wall contraction which is dyssynchronous.

Left bundle branch block (LBBB) is common in heart failure (Baldasseroni et al. 2002; Hawkins et al. 2007; Xiao et al. 1991) and it is conceptually easy to understand why this disruption of normal electrical conduction would lead to disrupted mechanical contraction. Many studies demonstrated the link between LBBB and poor survival in heart failure, with a rise in mortality as the QRS duration becomes more prolonged (Baldasseroni et al. 2002; Hawkins et al. 2007; Hochleitner et al. 1990; Shamim et al. 1999; Stewart et al. 2011).

As a consequence of the observation, innovative electrophysiologists introduced the concept of pacing the ventricles simultaneously at two distinct sites to alleviate the effect of LBBB (and of long PR intervals) (Auricchio et al. 1999a; Cazeau et al. 1994; Nelson et al. 2000).

1.2 Development of criteria for selecting patients – problems and solutions

The dramatic success of biventricular pacing seen in many studies propelled the field into the limelight and in its wake came the assumption that mechanical dyssynchrony – a frequent counterpart of LBBB – must surely be the true lesion being targeted. In parallel with studies reporting the mechanical counterpart of electrical dyssynchrony (LBBB), large scale clinical trials organised by manufacturers demonstrated clear effect on quality of life, 6 minute walk test, ejection fraction and symptoms (Cazeau

et al. 2001; Cleland et al. 2005; Higgins et al. 2003; Kapetanakis et al. 2011; Linde et al. 2002; Sutton et al. 2006a; Young et al. 2003) from CRT in patients with heart failure. CARE HF closed the debate on whether CRT was beneficial: the survival enhancement was large, unequivocal and incontrovertible.

Initial selection criteria included wide QRS duration, and prolonged PR intervals (Bristow et al. 2004; Cazeau et al. 1994). Entry criteria for the majority of the studies included a long QRS duration, a reduced ejection fraction, and NYHA III/IV. There has also been work on NYHA class II (Adabag et al. 2011; Linde et al. 2010; Moss et al. 2009a) patients and those with a narrower QRS duration (Beshai et al. 2007; Cazeau et al. 2008; Ghio et al. 2004; van Bommel et al. 2010a) which have had varied results.

At this stage the very large number of papers reporting associations between mechanical dyssynchrony and physiological response became the driver to encourage detailed assessment of patients for mechanical dyssynchrony as part of their evaluation for CRT implantation. Our centre was also caught up in this excitement and published recommended protocols for mechanical dyssynchrony assessment (Lane et al. 2004). A forest of papers soon followed. They showed that the identification of mechanical dyssynchrony could predict response (Bax et al. 2003a; Chung et al. 2008a; Kapetanakis et al. 2005; Kapetanakis et al. 2011; Yu et al. 2004a). The concept of identifying patients with dyssynchrony was championed by many, assisting in identifying patients who would benefit most from CRT implantation. Many techniques became available and are still continuing to come to light, with the reported ability to detect mechanical dyssynchrony in a manner that predicts response. Mostly at the forefront was tissue Doppler imaging and real time 3-dimensional echocardiography.

Unfortunately many patients in real life do not show a clear, favourable response (Auricchio & Prinzen 2011) as the published reports lead us to believe. Many individual centres could not reproduce reliable prediction, but assumed they were in error (and hence did not publish). What was the reason, and why could we not identify the cause of this problem? Is it because they are selected incorrectly? Do they require optimisation of their pacemaker settings? How do we then proceed to optimise these patients? Or is there a fundamental flaw in our understanding and expectation of CRT devices in heart failure patients? The hunt for suitable mechanical dyssynchrony marker to select patients for CRT implantation continued. Surely assessing many favourable markers head-to-head was the way forward?

Only a few people pushed back by arguing that mechanical dyssynchrony was never the principal basis of the endpoint trials. Instead, because it is clinically obvious that electrical dyssynchrony surely can only be the mediate harm through mechanical dyssynchrony, and the CARE-HF pivotal endpoint trial had a mechanical dyssynchrony avenue of entry (albeit used by only a minority of patients), most accepted without hesitation that mechanical dyssynchrony ought to be worth assessing, if only we could fathom how, as part of selection of patients for CRT.

1.3 The PROSPECT Crisis

The PROSPECT study was a disaster for those in favour of mechanical dyssynchrony and people were forced to confront serious gaps within the reasoning (Chung et al. 2008a). The arguments included statements that those centres involved in PROSPECT were not skilled or able to perform these measurements (even though these were the leading centres!), and that there were different machines used across different centres.

However, these alone could not explain the abject failure of prediction by mechanical dyssynchrony.

Looking back through the literature, despite the dominant reputation of a large number reporting that mechanical dyssynchrony controlled response almost exclusively (Bax et al. 2003a; Bleeker et al. 2007b; Kapetanakis et al. 2011), there was a smaller group of less highly cited papers contradicting these findings (Diab et al. 2011; Marcus et al. 2005). In the early phase of my thesis I worked with others to systematically review these to understand why this contradiction is in the literature.

1.4 The next step…

Many research centres are now focussing work on trying to reduce mechanical dyssynchrony by optimising the settings of the devices (Bax et al. 2003a; Chung et al. 2008a; Duvall et al. 2010; Inoue et al. 2005a; Jansen et al. 2006a; Sawhney et al. 2004; Scharf et al. 2005; Thomas et al. 2009; Turcott et al. 2010; Zhang et al. 2008; Zuber et al. 2008). This materialised due to a sizable proportion of patients not responding to CRT implantation (Reuter et al. 2002). The question arose…. Why do these patients not respond and benefit as expected? How do we handle this, as the device is already implanted? The immediate attractive answer once the device is implanted is to attempt to improve the settings and potentially improve the function of the device for that particular patient. Hence the consideration of optimising the settings: firstly altering the timings of the atrioventricular (AV) delay and secondly the ventriculo-ventricular (VV) delay. Shortening a prolonged AV delay by atrially sensed RV pacing has already been shown to be beneficial (Hochleitner et al. 1990), long before CRT devices were used; so AV adjustment was a potential avenue to consider in CRT also.

Optimising the intra-cardiac settings has potentially two mechanisms of action. First, changing the AV delay timings can result in improved filling of the ventricle if it permits long filling time while preventing atrial ejection being uninterrupted by ventricular systole (Ritter et al. 1999). Second, choosing the ideal VV delay setting should reco-ordinate the activation timings of the ventricle and improve mechanical dyssynchrony (Sogaard et al. 2002; Vanderheyden et al. 2005).

After implantation of a resynchronisation device, in typical UK practice relatively few patients undergo routine optimisation even though guidelines recommend that AV and VV delay should be optimised, and even though clinical trials have only demonstrated survival benefit of individually-optimised CRT (Carluccio et al. 2011; Cleland et al. 2005). Are we, as clinicians, right to cut corners from the trial-validated, guideline-mandated process? To address this, I examine the basic science of optimisation in Chapter 4.

Conflicting studies also emerged as to what was a reliable way to optimise intra cardiac timings of CRT devices successfully (Chung et al. 2008a; Duvall et al. 2010; Parreira et al. 2005; Thomas et al. 2009; Turcott et al. 2010; Vesely et al. 2008; Whinnett et al. 2006b).

1.5 The Problem

There are two very large unanswered practical questions with regards to CRT which are inter-related.

(1) Which patients are in need of resynchronisation and which are not i.e. what is a good measure of dyssynchrony out of the many that have been reported to be reliable predictors of benefit?

(2) Once implanted, how should we determine the best timings of the two ventricular leads?

Provided that CRT works by reducing mechanical dyssynchrony, there is a systematic solution to these questions. First we need to find a tool that can accurately, reliably and reproducibly measure small changes in mechanical dyssynchrony. We can then test this tool for its ability to identify patients who would benefit from CRT, by running a trial comparing CRT-on versus CRT-off, on a wide spectrum of values of this variable at baseline. Following implantation, we should then be able to optimise VV delay to minimise this marker of dyssynchrony.

1.6 The different families for measuring mechanical dyssynchrony

Measuring mechanical dyssynchrony has a wide variability and at present there is no single method that has been shown to be used in multiple centres which can reliably measure it, hence it is unclear as to what to do practically?

Echocardiography has been widely studied and in Chapter's 3 and 4 I have reviewed current literature to identify those parameters which have been used for both selecting

and predicting response to CRT implantation, and those feasible for guiding optimisation.

For each of the imaging techniques, measurements may be made within the left ventricle, or as a comparison between the right and left ventricle.

1.6.1 Measurements focusing between the right and left ventricle

Doppler timings for guiding selection for CRT device implantation have included the difference between ejection of the right and left ventricle, namely the inter-ventricular ejection delay using pulmonary Doppler flow and aortic Doppler flow measurements.

1.6.2 Measurements focusing within the left ventricle

Tissue Doppler imaging conventionally is used for dyssynchrony assessment by looking at the left ventricular walls and establishing the timings for contraction of the different walls. Most commonly quoted in the literature is the 2 segment, 6 segment of 12 segment model (Bax et al. 2003a; Diab et al. 2011; Notabartolo et al. 2004; van Bommel et al. 2010b; Vesely et al. 2008; Yu et al. 2004a).

3 dimensional echocardiography is used for dyssynchrony assessment in some centres (Kapetanakis et al. 2005) and divides the left ventricle into multiple segments. Dyssynchrony assessment is performed on each of these segments reaching their minimal volume.

Doppler timings are recommending in some local guidelines (Lane et al. 2004) and can be used as guide to selection for implantation. However, the aortic pre-ejection time is only half of the story, because it addresses the left ventricle only.

1.6.3 Final pathway echocardiographic measurements for optimisation

Stroke volume can be assessed using Doppler and has been used in some centres to guide optimisation (Gorcsan et al. 2004b). This is the LVOT VTI measurement which is the resulting cardiac flow that is ejected and often used as a surrogate for cardiac output.

1.7 How do we evaluate a proposed marker for assessing dyssynchrony and guiding optimisation?

Any marker proposed for this purpose must meet a series of criteria

1) At bare minimum, the method proposed must be able to reliably detect changes which occur when intra-cardiac timings are changed experimentally within an individual with everything else kept constant.

2) The measured outcome or marker should be independent from that which is used to optimise, so that individual samples of random error are not used to validate themselves.

These are cheap and quick to perform on any technique which is proposed as a method for optimisation, and in fact should be mandatory before clinical studies and trials are performed as these will be doomed unless they are satisfied.

1.8 Aims of this thesis

Within this thesis I shall probe the plausibility of the "dyssynchrony-response" hypothesis, and examine the basic science for developing echocardiographic protocols for CRT optimisation. My individual aims are as follows:

1.8.1 To evaluate and understand the current literature on cardiac resynchronisation therapy.

I critically review the world literature on CRT and methods of predicting those who may benefit from CRT implantation. I cover large and small studies to gain a full overview of the work that have been performed. I review our local data of echocardiographic variability of ejection fraction, a marker often used as an endpoint. I provide mathematical algorithms to understand the degree of correlation possible between measurements and outcomes, so that claims in the literature may be viewed in the context of actual possible findings.

1.8.2 How can we ensure protocols for optimisation are not a waste of time, but in fact reliable?

I perform a mathematical simulation to understand the effect of signal and noise on the reliability and reproducibility of optimisation methods. I evaluate the literature to determine how this relates to the methods employed currently for optimisation purposes and will suggest ways we can use protocols that are robust for clinical use.

1.8.3 Are echocardiographic markers plausible candidates for measuring dyssynchrony that can be measured by CRT?

Many echocardiographic markers have been proposed as potential criteria to select patients for CRT implantation. I aim to test a series of these markers for their ability

to detect changes in mechanical dyssynchrony for experimental manipulation of dyssynchrony i.e. alteration of VV delay.

I assess the clinical implications of their various reproducibility values, not only for optimisation of CRT devices but also for selection of patients for CRT implantation.

1.8.4 Assess the effect of changing atrioventricular delay on invasive flow, invasive pressure and non-invasive pressure.

My thesis also addresses the interplay between blood flow and arterial blood pressure in the immediate aftermath of changes in AV delay

I aim to identify the beat-by-beat changes which occur when AV delay is changed from a favourable to a non favourable AV delay setting. With high temporal precision I measure invasive and non-invasive pressure and invasive flow, to assess whether improvements in pressure and cardiac output persist following improvements in AV delay.

2 Methodology

2.1 Patient recruitment

Patients were recruited from the Imperial College NHS Healthcare trust if they met the selection criteria, mentioned in individual chapters. They were given the opportunity to discuss the study and provided with a letter and an information sheet (Appendix 1 and 2). Patients who agreed to participate in the study signed consent forms (Appendix 3). Patients were free to withdraw from the study at any time.

2.2 Equipment

2.2.1 Echocardiography

Echocardiography was performed in all patients using the Philips IE33 machine (Figure 2-1). This machine was chosen as it is able to perform both the advanced 3 Dimensional (3D) echocardiography and tissue Doppler imaging (TDI) required for some sections of the study in addition to the more routine 2D spectral Doppler necessary for the other sections of the study.

The IE33 enabled me to acquire images which were compatible with the 3D analysis software Tomtec (4D analysis CAP, Tomtec GmBH, Unterschlessheim, Germany) and the tissue Doppler analysis software, Philips Qlab 8.1with cardiac motion analysis (CMQ). Echocardiographic imaging was performed in all the study patients in the outpatient setting, which meant that this machine was therefore an ideal choice given its features and locality for this use.

2.2.2 Transducers

Both transducers which were used in this study have Purewave Crystal technology that is able to transfer energy with greater precision and efficiency and hence superior overall quality (Yu et al. 2009) compared to previous technology. A watery based gel is applied to all transducers to improve contact impedance between the probe and patients skin.

2.2.2.1 Transthoracic Probe 2D

2D studies and Doppler images were acquired using the Philips S5-1 sector ultrasound transducer (Figure 2-1). This is a phased array probe with frequencies between 1 and 5MHz. Phased array probes can be swept through a wide volume without physically needing to turn the probe. This is possible because of the multiple elements making up the probe and which can be pulsed at different times.

2.2.2.2 Transthoracic probe 3D

3D echocardiography was performed using a X3-1 matrix array transducer (Figure 2-1) which is fully compatible with the IE33 machine and enables triggered full volume images required for the real time 3D images. Therefore, full volume pyramid shaped datasets are acquired and can be manipulated. Geometric assumptions can be corrected for. This probe operates at frequencies between 1 and 3MHz. A minimum frame rate of 20Hz has been used throughout this study.



Philips IE33 echocardiography machine



Philips S5-1 transducer



Philips X3-1 transducer

Figure 2-1: Photograph of the Philips IE33 machine used for data acquisition (top panel), with the 2D probe (middle panel) and the 3D probe (bottom panel).

2.3 Echocardiographic measurements used throughout the studies for optimisation processes.

All of the echocardiographic measures that I critically analysed in my studies have been used in previously published research into optimisation, and have been proposed as valuable tools for guiding the optimisation process both in the research arena and in routine clinical practice.

A clear and visible ECG trace was recorded with all echocardiographic images to allow timings to be calculated with precision and reference to the QRS complex.

2.3.1 Doppler assessment of VV delay

Pulsed wave Doppler was used in both (a) ejection flow dyssynchrony and (b) velocity time integral measurements. Both of these measures have been used in published studies of echocardiographic optimisation of biventricular pacemakers (Bertini et al. 2008; Duvall et al. 2010; Fischer et al. 2009; Lane et al. 2004; Parreira et al. 2005; Turcott et al. 2010; van Geldorp et al. 2011; Vanderheyden et al. 2005).

2.3.1.1 Ejection flow dyssynchrony – acquisition

Ejection flow dyssynchrony was measured in 2 ways. Firstly aortic pre-ejection time (APET), calculated from QRS onset to commencement of flow from left ventricular outflow tract (Figure 2-2) and, secondly, interventricular mechanical delay (IVMD) calculated as the difference between onset of LV ejection and RV ejection (Lane et al. 2004). Both have been demonstrated to be suggestive of intraventricular and interventricular delay respectively (Cazeau et al. 2003).

Firstly we acquired pulsed wave (PW) Doppler images in the left ventricular outflow. Images were taken from the apical 5-chamber view with the pulse wave Doppler

cursor at the left ventricular outflow tract approximately 1 cm below the aortic annulus with the PW-Doppler (Lopez-Candales & Edelman 2011). For the right ventricular outflow tract, the PW Doppler was suitably placed approximately 1cm distal to the pulmonary valve (Lopez-Candales & Edelman 2011).

Eight beats were acquired at each VV delay namely RV excitation first at 40ms, 20 ms, LV excitation first at 20ms, 40ms and 60ms, and simultaneous ventricular excitation at VV 0ms. All acquisitions were made at end-expiration to minimise respiratory variation.

During spectral Doppler imaging the position of the transducer needs to allow the blood flow to be parallel to it (and hence parallel to the sound waves). Any different inclination of the probe will result in inaccurate blood velocity to be measured unless corrections are made. Therefore I took great care to ensure not only that the angle was correct, but also that the subsequent images also have the same careful consideration when acquiring Doppler on subsequent beats / days.

2.3.1.2 Ejection flow dyssynchrony – analysis

The analysis was performed by me, a single experienced operator, offline which ensured that there was no haste in the measurements with the patients present but instead time and a calm environment to analyse the data ensured an accurate and precise analysis. All images were stored in DICOM format to a specialised echocardiography image storage and analysis system, Medcon UK (Mc Kesson, CA). Offline analysis was performed here.

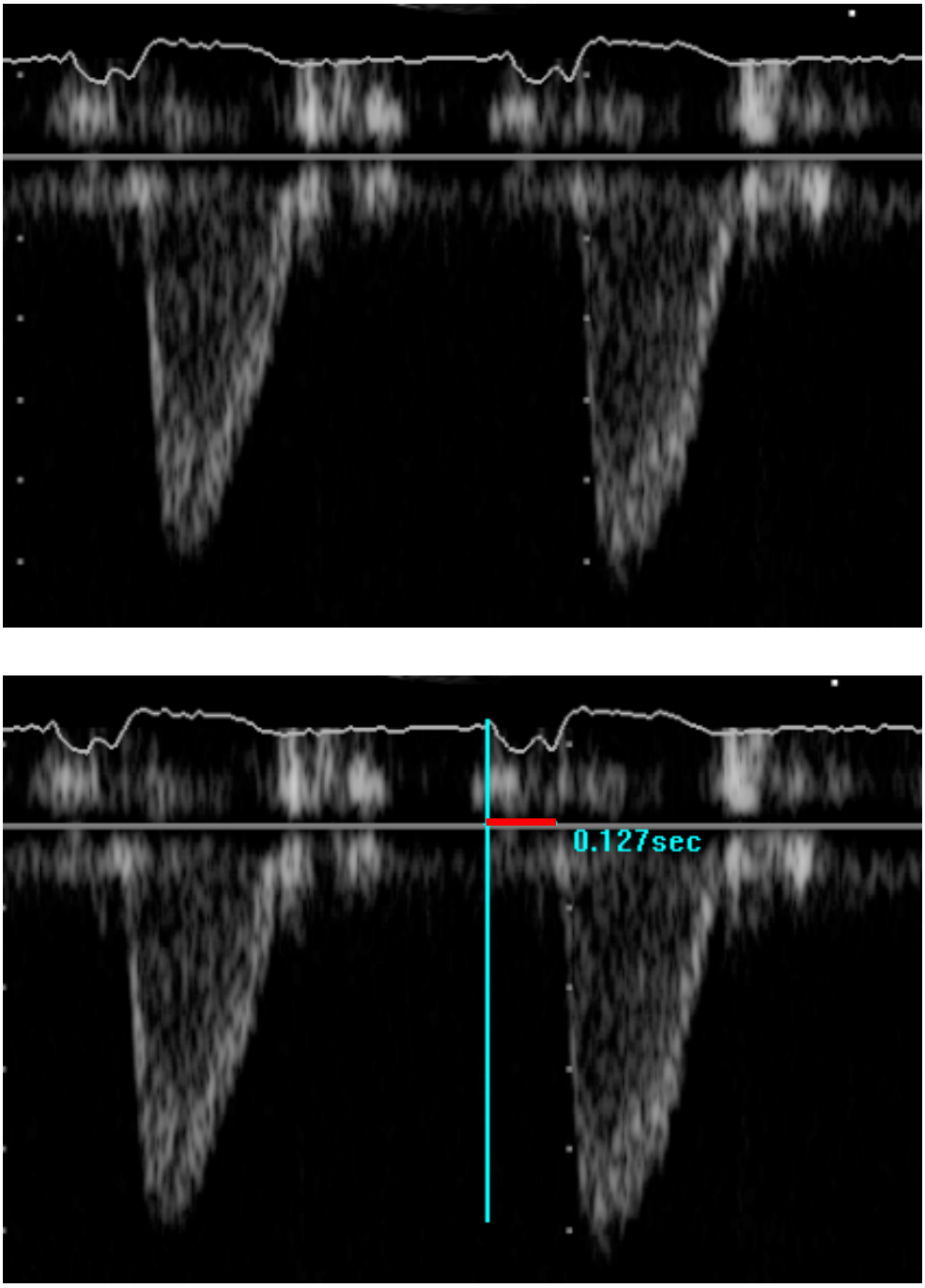


Figure 2-2 Measuring aortic pre-ejection time.
In the top panel I have shown the image that is acquired with the QRS clearly seen and the pulsed wave Doppler trace shown below it. The lower panel demonstrates the measurement taken from the onset of the QRS (the blue vertical line) and the timings (the red horizontal line). This was performed for the pulmonary flow in a similar manner.

2.3.1.3 Velocity-time Integral – acquisition

Images were taken from the left ventricular outflow tract as described above. A clear pulsed wave Doppler image was acquired as shown in Figure 2-3 and again 8 beats were acquired at each VV delay at end expiration.

2.3.1.4 Velocity-time Integral – analysis

Offline analysis was manually traced and resulted in digitised velocity time integral (VTI) tracings. These VTI values were measured for each patient, at each VV delay, at eight beats per VV delay.

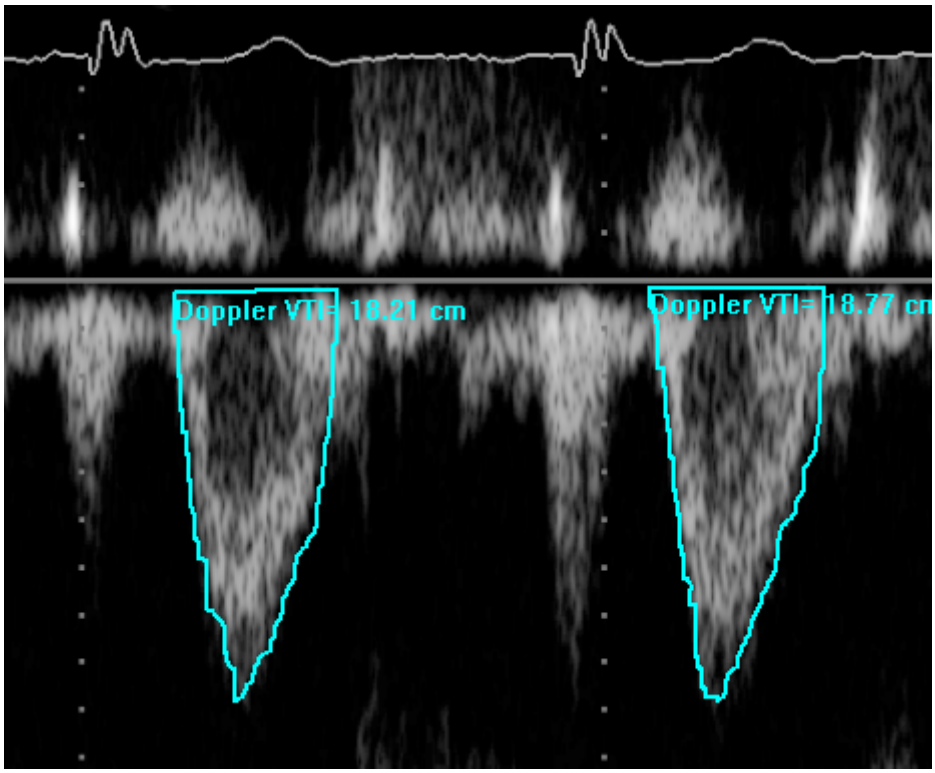


Figure 2-3 An example of a velocity time integral tracing of a pulsed wave Doppler image.

2.4 Segmental assessment of VV delay optimisation

Routinely the left ventricle is divided into a 16 or 17 segment model (Flachskampf & Daniel 2010; Liodakis et al. 2009) for other uses such as reporting stress echo in addition to dyssynchrony assessment, and in a similar fashion for 3 dimensional echo this process has been adopted (Kapetanakis et al. 2005).

For TDI imaging we are able to look at individual walls and adopt either a 2, 6 or 12 segment model (Chung et al. 2008a; Yu et al. 2004a) this can be used either for intra ventricular dyssynchrony assessment (most commonly) or interventricular dyssynchrony. In some cases these are combined to give an overall value to consider dyssynchrony (Lane et al. 2004).

2.4.1 3D echocardiography

The development of real time 3D echocardiography enabled LV assessment to be performed in yet another way (Corsi et al. 2005; Kuhl et al. 2004; Soliman et al. 2007a). Real time 3D echocardiography has the potential advantage of eliminating any geometric assumptions that may occur when taking measurements of the left ventricle. As such good results had been obtained in some centres; it has been assumed that other uses for this modality may also be beneficial.

Real time 3 dimensional echocardiography considers all 16 segments of the ventricle separately (Figure 2-4) and therefore also enables the operators to assess the intraventricular delay of the LV by calculating the time to minimal volume of each segment as a percentage of the cardiac cycle.

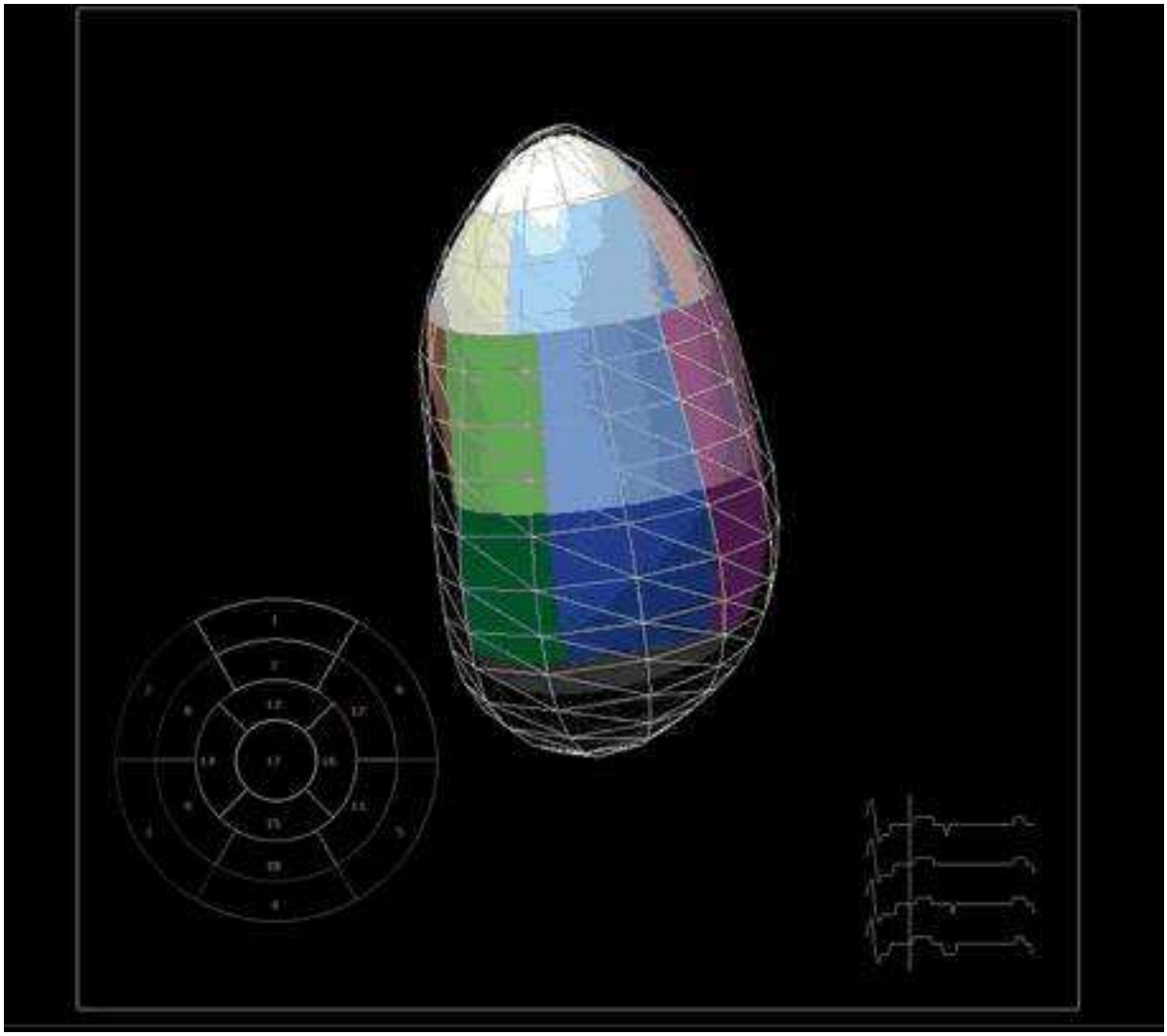


Figure 2-4 Segmental division of the left ventricle.
In the bottom left corner we can see the ventricle divided up into 16 segments (plus apex), each of which corresponds to a different coloured section of the ventricle.

The systolic dyssynchrony index (SDI) (Gimenes et al. 2008; Kapetanakis et al. 2011; Monaghan 2006) is a value which is the standard deviation of the time to minimal volume of all 16 segments of the left ventricle. A lower value correlates to less dyssynchrony. In a synchronous ventricle, we would get a value of approximately 3% to 5%. As the ventricle becomes more dyssynchronous this will increase. Values above 10% are said to be predictive of response to biventricular pacing (Kapetanakis et al. 2011).

Changing VV delay towards an optimal setting, should therefore lead to a lower SDI value, and therefore should theoretically be a potential tool for guiding optimisation of VV delay. The use of 3D assessment of dyssynchrony has been used in some centres for predicting response to biventricular pacing (Kapetanakis et al. 2005), but it has not been fully assessed for VV optimisation

2.4.1.1 3D echocardiography- acquisition

Echocardiography was performed by myself using the X3-1 matrix array transducer. Full-volume datasets were acquired over 4 cardiac cycles using this matrix array transducer, each recorded over 4 consecutive beats. Patients were in the left lateral decubitus position and images were recorded from the apical 4 chamber view with the left ventricle clearly identified. Acquisitions had a minimum frame rate of 20 frame/s in keeping with ASE guidelines (Chung et al. 2008b; Gorcsan et al. 2004b), with the patient breath holding at end expiration. For each VV delay four replicate measurements were taken.

2.4.1.2 3D echocardiography- analysis

Measurements of 3DE volumes were performed off-line using semi-automated border detection software (4D analysis CAP, Tomtec GmBH, Unterschlessheim, Germany), a method that has been well validated both for measurements of ejection fraction (Jenkins et al. 2009; Kuhl et al. 2004), and also for calculating the systolic dyssynchrony index (SDI), a method which attempts to predict response to CRT (Kapetanakis et al. 2011; Liodakis et al. 2009).

The apical 4 chamber is used as the reference plane, with the apical 2 and 3 chamber views derived automatically from a 60 degree rotation between planes, using manual adjustments as required. I chose the frames for EDV and ESV measurement in accordance with American Society of Echocardiography guidelines. EDV measurements were taken at the frame following mitral valve closure and ESV were measured on the image with the smallest left ventricular cavity. Initial contours were set by tracing the endocardial borders end-diastolic and -systolic images in the apical views. I traced the contours using the automatic border detection (Figure 2-5). Further manual editing was performed to ensure the smoothest delineation in the border between diastole and systole. A 3D mesh and the segmental model of the 3D shell are created. From this data, the software created time-volume curves and determines the time to minimum volume for each of the 16 segments. The standard deviation of these timings is then calculated automatically by the software which gives the systolic dyssynchrony index.

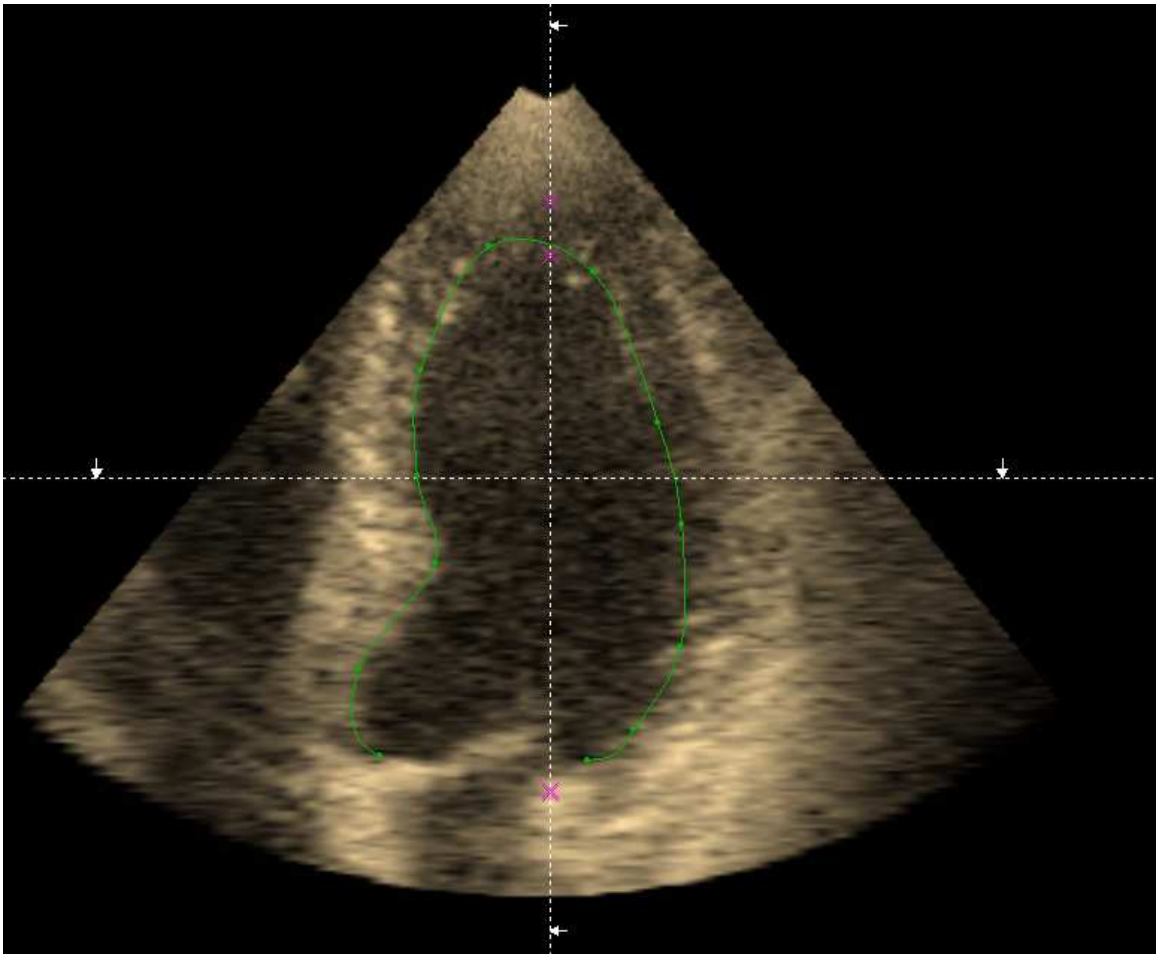


Figure 2-5 An example of drawing around the endocardium which is shown here for the 4 chamber view. This is repeated for 2 chamber and 3 chamber, which then results in the 3D reconstructed ventricle.

Time-volume curves are produced (Figure 2-6) and the systolic dyssynchrony index was calculated with the analysis performed by myself for each VV delay. Four replicate measurements were taken at each VV delay – therefore 4 optimisation processes were performed. In Figure 2-7 I have drawn schematic diagrams of time volume curves of a synchronous ventricle along side a dyssynchronous ventricle so that it is clearly shown how the systolic dyssynchrony index is calculated.

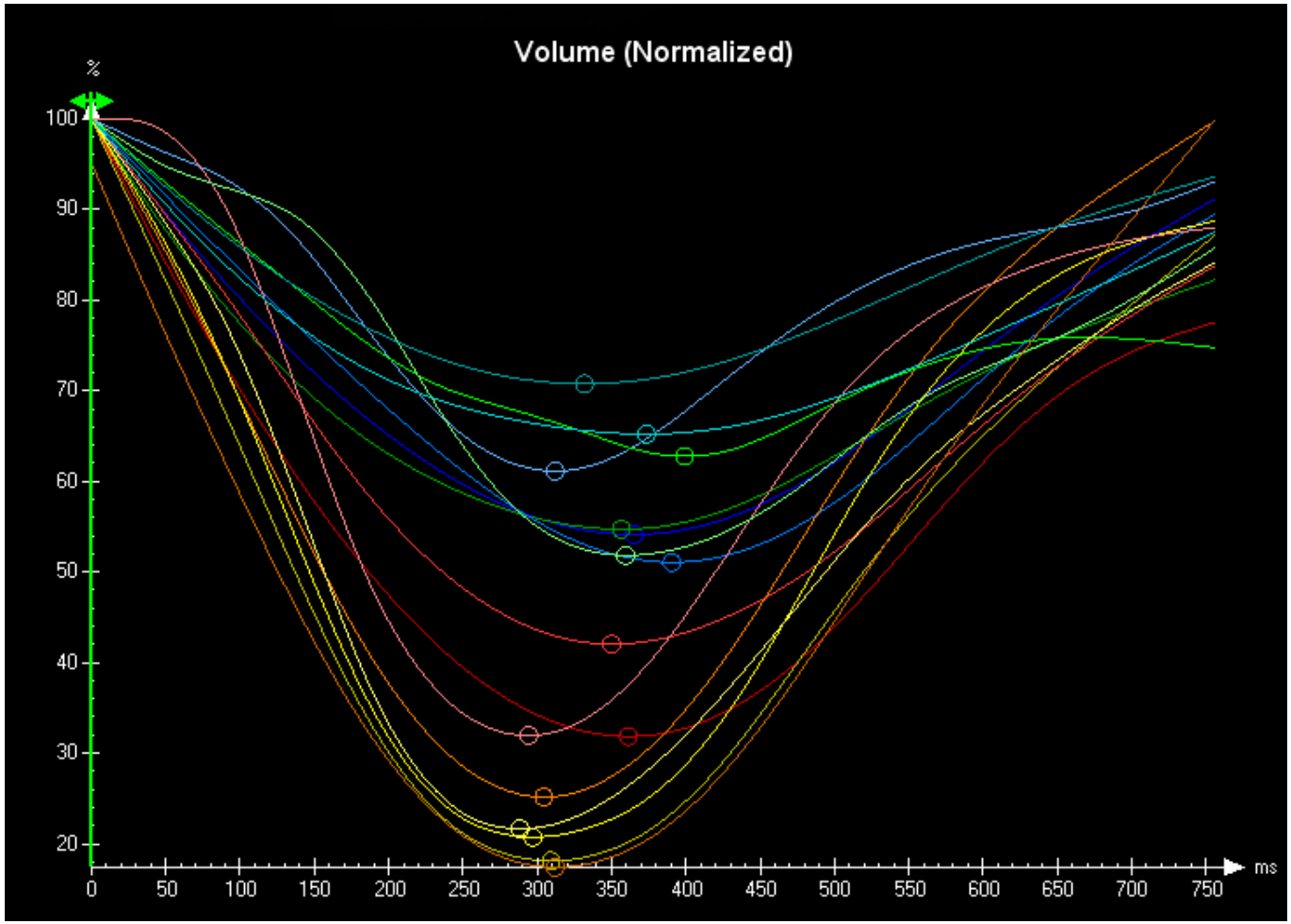


Figure 2-6 The dyssynchrony curves of a normal synchronous ventricle as seen when analysis is performed. All segments reach the time to minimum volume at a similar time. An SDI value is then calculated.

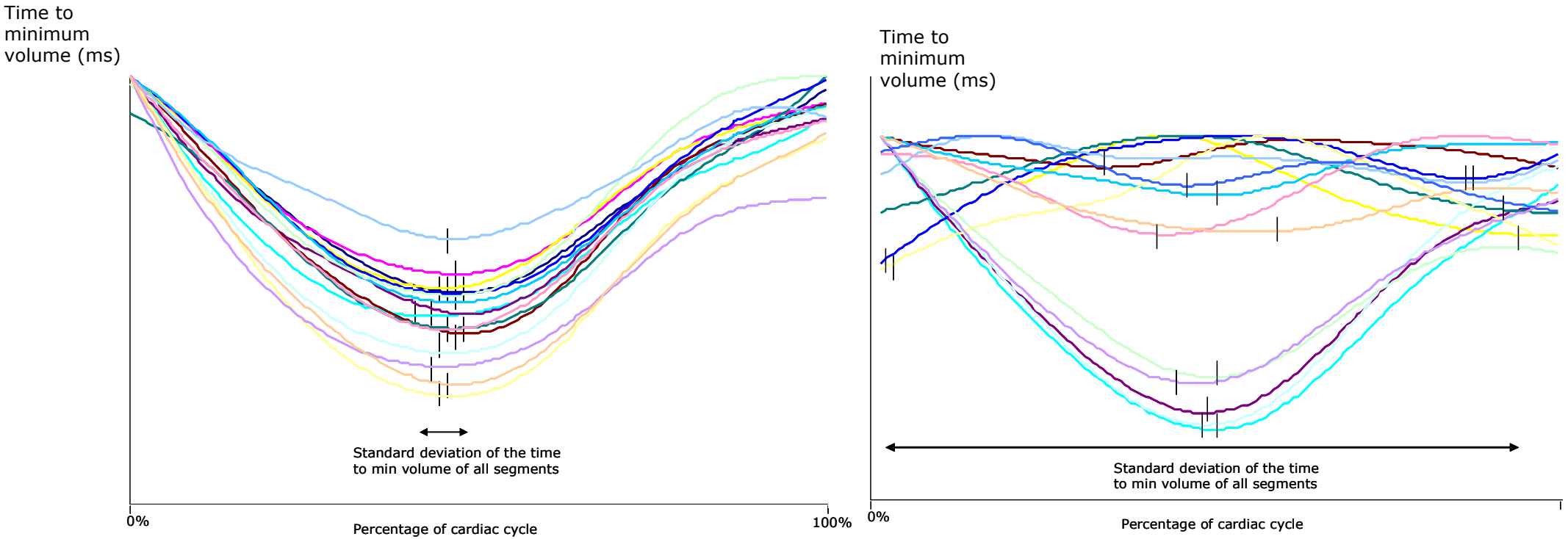


Figure 2-7 Schematic diagram showing the segmental analysis with each coloured line representing a segment of the ventricle.
Each 16 segments have a black line that denotes the time to minimal regional volume. The SDI calculation is the standard deviation of all of these time points. On the left there is a synchronous ventricle, and hence would calculate a low SDI. On the right there is a dyssynchronous ventricle with a high standard deviation and hence high SDI.

2.4.2 Tissue Doppler Imaging- technology

For tissue Doppler imaging, the colour Doppler frame rates were at least 90 frames/s in all subjects, with pulse repetition frequencies of 500 Hz to 1 kHz, resulting in aliasing velocities of 16 to 32 cm/s.

LV dyssynchrony using TDI has been proposed using several walls of the ventricle and several methods of calculating the maximum level of dyssynchrony (Chung et al. 2008a; Yu et al. 2002). Since none of the methods currently available have been deemed to be able to recommend improvement in patient selection for CRT beyond current guidelines (Chung et al. 2008a), we tested and calculated various possible combinations including 2 segment, 6 segment and 12 segment analysis (Sogaard et al. 2002; Yu et al. 2005). This therefore incorporated all walls and different levels within the ventricle.

2.4.2.1 Tissue Doppler Imaging-acquisition

Studies were performed using the IE33 and images were obtained using a 3.5MHz transducer. Apical images of the left ventricle including the level of the mitral annulus were acquired with a minimum frame rate of 90 Hz (Gorcsan et al. 2004b). Images were acquired with LV cavity positioned in the centre of the sector showing clear myocardial definition. 4 chamber, 3 chamber and 2 chamber views of the left ventricle were acquired with the patient holding their breath at end-expiration. A minimum of 4 cycles were acquired, triggered to the QRS complex and saved in cineloop format, allowing 4 separate analyses.

2.4.2.2 Tissue Doppler Imaging-analysis

We calculated offline (a) the time from onset of QRS to onset of S wave and (b) the time from onset of QRS to the peak of S wave for 2 segment, 6 segment and 12

segment tissue Doppler measurements. To obtain time–velocity curves, a sample volume was placed within the segments at the region of interest (Figure 2-8). 2 segments involved the basal septum and basal lateral walls. 6 segments involved the basal septum, lateral, anterior, inferior, posterior and anteroseptal walls. 12 segments consisted of the 6 of the 6 segment model and their mid level counterpart. All measurements were taken at rest at end expiration.

To assess LV dyssynchrony, the interval to-peak systolic velocities was obtained by placing sample volumes in desired walls. The optima was defined by, firstly, the absolute range (maximum minus minimum) of times between segments of the left ventricle (Bleeker et al. 2007a; van Bommel et al. 2010b) and, secondly, the standard deviation of these dispersion times using 2-segment, 6-segment and 12-segment models. For the 2 segment model a cut-off of >60 ms has previously been quoted as a marker of dyssynchrony (Bax et al. 2003a). For a 12-segment model often the Yu index is championed (Yu et al. 2002; Yu et al. 2005) with a cut off >32ms as the standard deviation between the 12 segments as the threshold for dyssynchrony. I considered that the more synchronous the ventricle, the smaller the value for the delay between the walls. The greater the dyssynchrony, the greater the delay between the walls.

Colour coded TDI was chose in preference over pulsed wave TDI since you can get several segments analysed in the same heart beat and analyse offline (Bleeker et al. 2007c). We used Philips Qlab 7.0 3DQ and 3DQ advanced software the QLAB Strain quantification plug-in which allowed segmental analysis of the tissue Doppler traces.

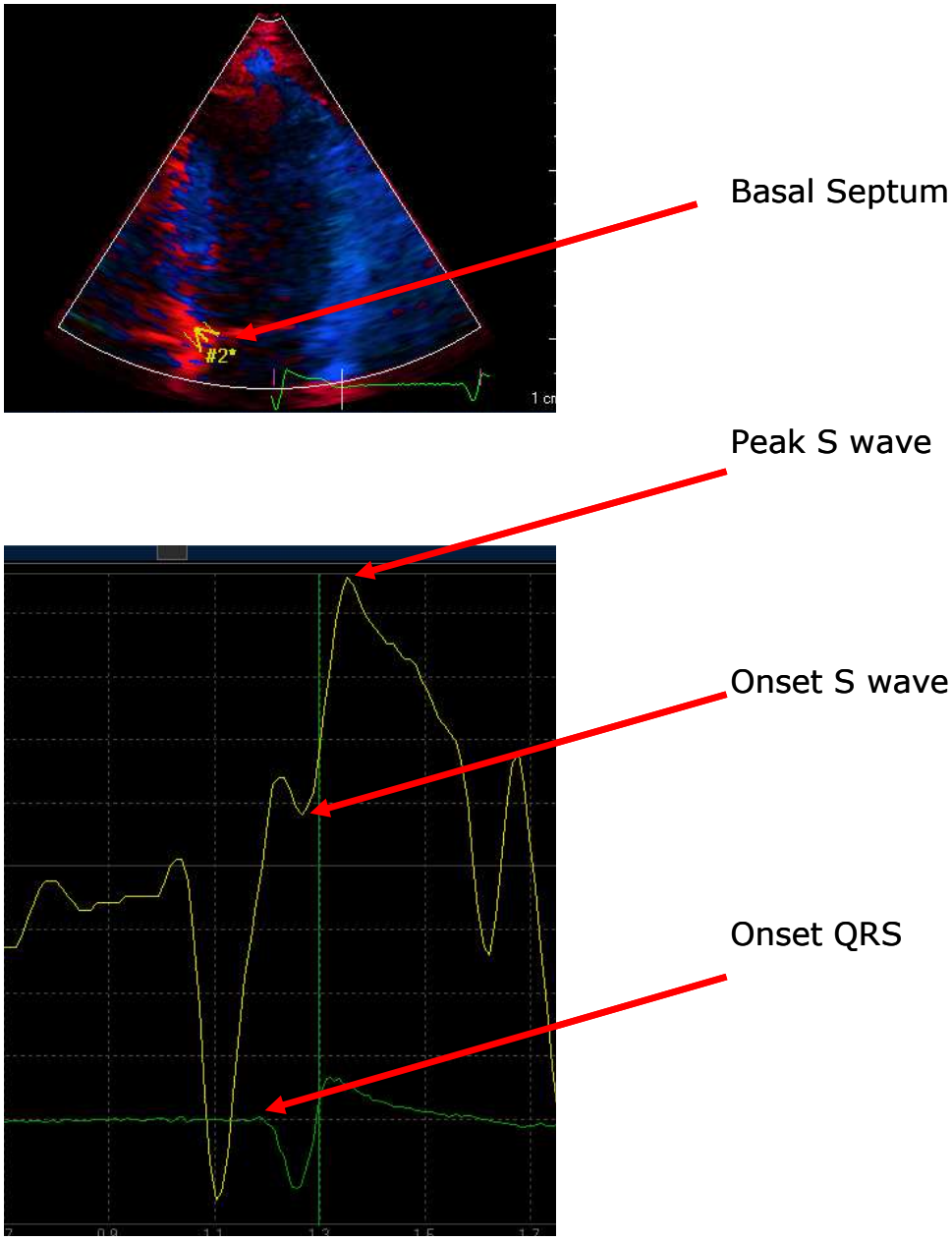


Figure 2-8 An example of a TDI trace where the region of interest is the basal septum which is where the cursor is positioned in the top panel.
The subsequent tissue Doppler trace is seen in the bottom panel. Measurements are taken from (a) the time from onset of QRS to onset of S wave and (b) the time from onset of QRS to the peak of S wave.

2.5 QRS duration as a marker for optimisation of CRT

QRS duration is used as a guide to help select patients for CRT implantation (Barnett et al. 2007). Therefore the question could be raised, that if it is a useful and recommended method for selection, could it also be used for optimisation? (Tamborero et al. 2011; Vidal et al. 2007). To add further weight to this we know that a longer QRS duration is associated with increased mortality (Shamim et al. 1999) and furthermore, even within the normal reference ranges it is associated with a greater LV mass, LVEDV and LVESV(Stewart et al. 2011; Yerra et al. 2006)

2.6 Pacemaker programmers

Patients had either Medtronic or Boston Scientific biventricular pacemakers /ICD's implanted, therefore we used both these pacemaker programmers to manipulate both the AV and VV delay settings in order to perform these studies (**Figure 2-9**).

In the non-invasive studies we used both programmers in the standard way to enable connection between the patient and the programmer.

The patient's pre study settings were recorded at the beginning of the study. We then changed the atrioventricular (AV) and ventriculo-ventricular (VV) delays as per each study protocol, documented in each chapter. We were also able to change the atrial rate of the pacemaker. At the end of each study visit the patients initial pacemaker setting were re-program



Figure 2-9 A photograph of the Boston Guidant pacemaker used in the study (left panel) and a close up of the programmer screen (right panel) as programmed during one of my experiments.

2.7 Invasive study

Temporary biventricular pacing was performed in eligible subjects. A quadripolar electrode catheter (Josephson Curve, Bard Viking) was placed in the right atrium (usually the right atrial appendage) and a pentapole electrode catheter was placed at the right ventricular apex. An AL1 and/or a channel sheath was used to gain access to the coronary sinus and an ATW wire was positioned in a lateral or posterior-lateral coronary sinus branch for LV temporary pacing(Lane et al. 2008). LV capture was verified using a 12 lead ECG.

In this study a CRT pacemaker (Medtronic InSynch III 8042) was connected extracorporally and transitions made via a standard pacemaker programmer.

2.7.1 Invasive study - acquisition

The protocol consisted of a series of transitions from AV delay 40ms to 120 ms. Measurements were recorded for at least 100 beats before and 100 beats after the change in AV delay. To minimise the effect of random variation, triplicate runs of the experimental protocol were conducted so that each patient's dataset was composed of an average of 3 runs, aligned by beat and registered to the time point of transition from 40 ms to 120 ms. Heart rate was fixed at 100bpm in all patients by atrially pacing. Stable pacing and sensing for all 3 pacing wires was monitored throughout the protocol in all patients.

2.7.2 Invasive pressure and flow measurements

Aortic pressure was measured using a fluid filled catheter positioned in the aorta approximately 5 cm from the aortic valve. The catheter pressure was initially normalised to the pressure wire signal with the pressure wire positioned at the tip of

the catheter. The aortic flow was measured with a flow wire (Volcano FloWire 1400) also approximately 5 cm from the aortic valve within the aorta.

2.7.3 Non-invasive blood pressure measurements - Finometer technology

Beat-by-beat blood pressure was recorded using the Finometer system (Finapres Medical Systems, Amsterdam, Holland). The blood pressure is measured based on the arterial-volume clamp method (Imholz et al. 1998; Penaz 1973), after a cuff is placed in the middle finger.

This is a form of plethysmograph which uses light to non-invasively detect changes in microvascular blood volume (Dorlas & Nijboer 1985; Turcott & Pavek 2008). The main components within the cuff are an inflatable air bladder and a plethysmograph with an infra red light source and light detector. The front-end unit is connected to the air bladder by a hose and to the light source and detector by a cable (Bogert & van Lieshout 2005). This front-end unit is then connected to the main unit and pump, Figure 2-10.





Figure 2-10 A Finometer unit (top panel) which has the connections at the front as described: the main unit is strapped to the back of the patients arm with the cuff around the index finger (bottom panel).

By applying a finger cuff and monitoring pressure it is able to detect the change in volume as the amount of pressure requires will vary. This is then digitised when there is a change in pressure which occurs i.e. when AV or VV delay settings are altered for example. Transient changes in both volume and pressure (which are proportional) will allow optima to be identified. This has been investigated by Whinnett et. al who have identified that the Finometer, an advanced plethysmograph, can indeed be used to optimise changes in both AV and VV delay (Whinnett et al. 2006a). Maximising blood pressure is the ultimate aim given low SBP readings are a predictor of mortality in the heart failure patient population (Abraham et al. 2008; Adams, Jr. et al. 2005).

The pressure is recorded based on the amount of pressure required to keep a constant arterial diameter. Hence there is a pulsatile feeling on the patient's middle finger although if the trace is not good, then the other fingers can be used. There is a balloon which is within the finger cuff that inflates and deflates in order to maintain a constant pressure. Prior to commencing the recordings, there is a calibration process that occurs. The amount of cuff pressure required to counteract the intra-arterial pressure is what gives the indirect measure of the change in intra-arterial pressure (Wesseling et al. 1985).

The Finometer has been well validated for a method of monitoring blood pressure (Schutte et al. 2004) and is widely recommended as a valid technology for the measurement of changes in arterial blood pressure (Kurki et al., 1987; Smith et al., 1985; Wesseling et al., 1985). It has also been used for recording changes during optimisation protocols in the research setting (Whinnett et al. 2006a).

2.7.4 Surface ECG

A three lead ECG recording was continuously measured for all of the patients undergoing invasive testing (**Figure 2-11**). The leads were positioned (1) on the right side of the chest just below the right clavicle, (2) on the left side of the chest, just below the left clavicle and (3) on the lower chest, just above and left of the umbilicus (Pope 2002). Those undergoing non-invasive echocardiography without the invasive protocol had the ECG recorded on the echocardiography machine only.

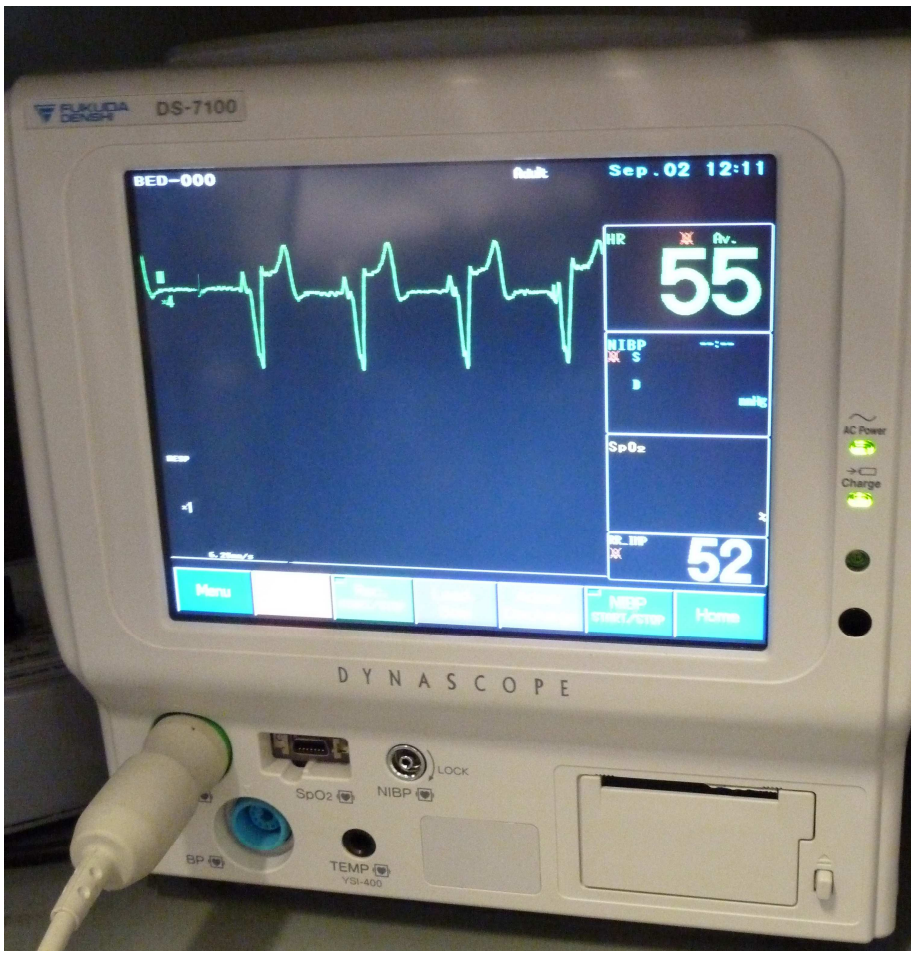


Figure 2-11 An example of the ECG acquisition which occurs simultaneously with the invasive measurements

2.7.5 Acquisition system

Analogue output feeds of all the signals were taken via a National instruments DAQ-Card AI-16E-4 (National Instruments, Austin, TX) and acquired in digital form using Labview (National Instruments, Austin, TX), Figure 2-12. All the raw data was stored in text files for use in the subsequent offline analysis.

2.7.6 Analysis software

The data for non-invasive blood pressure, invasive blood pressure and invasive flow was collected as mentioned and analysed with custom software based ion the Matlab platform (MathWorks, Natick, MA).

This software allows automated analysis, therefore many more replicates can be analysed in a shorter space of time. This improves the signal-to-noise ratio. Automation also removes human error which can occur if measuring traces and recordings by manual methods.

The Matlab platform was the platform upon which the simulation for AV and VV optimisation was performed. This is described in detail in Chapter 4.

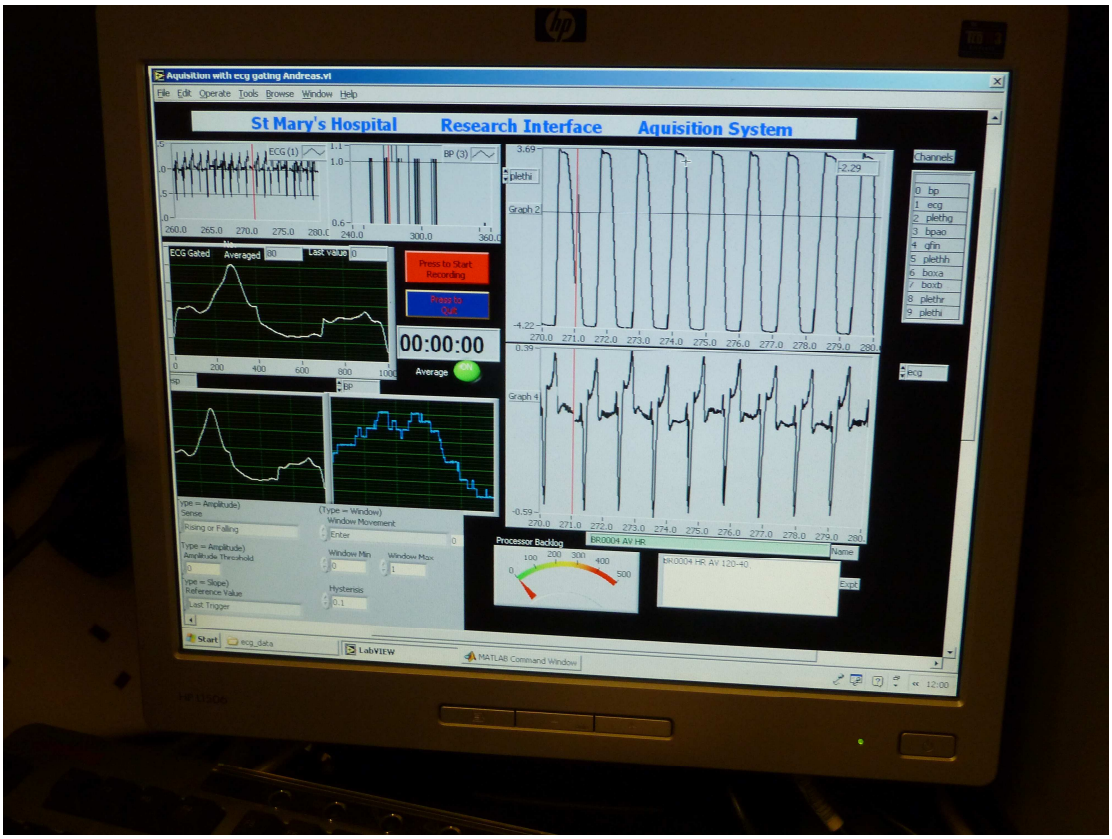


Figure 2-12 A snap shot of the data acquisition system on Labview where all data is recorded

2.8 Exclusion of ectopic beats

Throughout all the studies, I wanted to maximise the precision of all measurements to ensure that we are accurately and fairly assessing each of the measurement parameters. Ventricular ectopic beats were excluded and the beat following this ectopic was also excluded. This is because the ectopic beat had a cardiac output which was not representative of a normal beat i.e. too little or too much cardiac output as shown by echocardiography and the subsequent beat had the compensatory additional or lesser blood volume within it therefore a greater output. This can be seen by the echocardiographic images below which demonstrate this point, Figure 2-13.

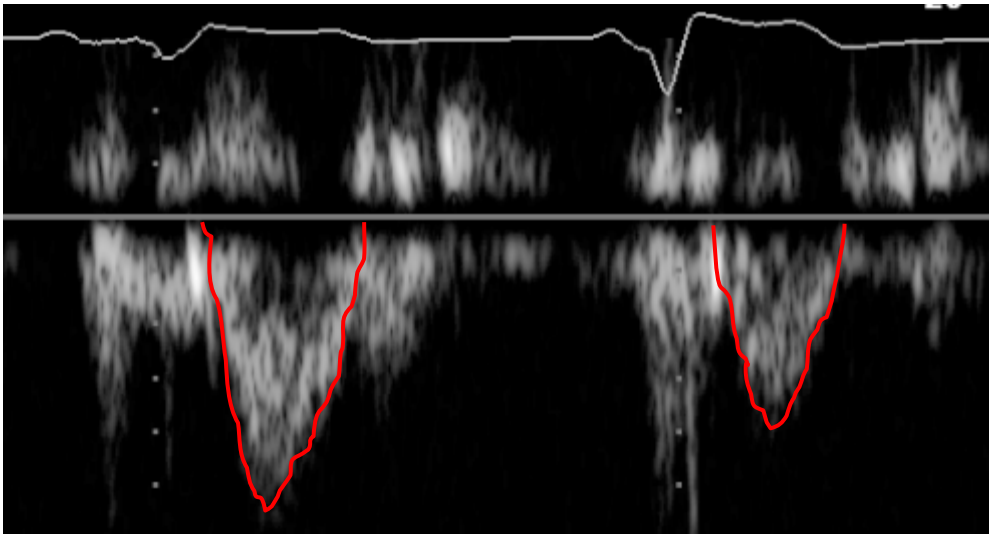


Figure 2-13 showing ectopic and a normal beat with VTI measured. We can see the final beat in the frame (on the right) is a normal beat following the ectopic. The difference in VTI is visibly noticeable.

This is also reflected in the end product of output i.e. the blood pressure which is being measured (Figure 2-14) and hence these beats in the Finometer non-invasive and aortic invasive blood pressure tracings should be excluded in a similar fashion.

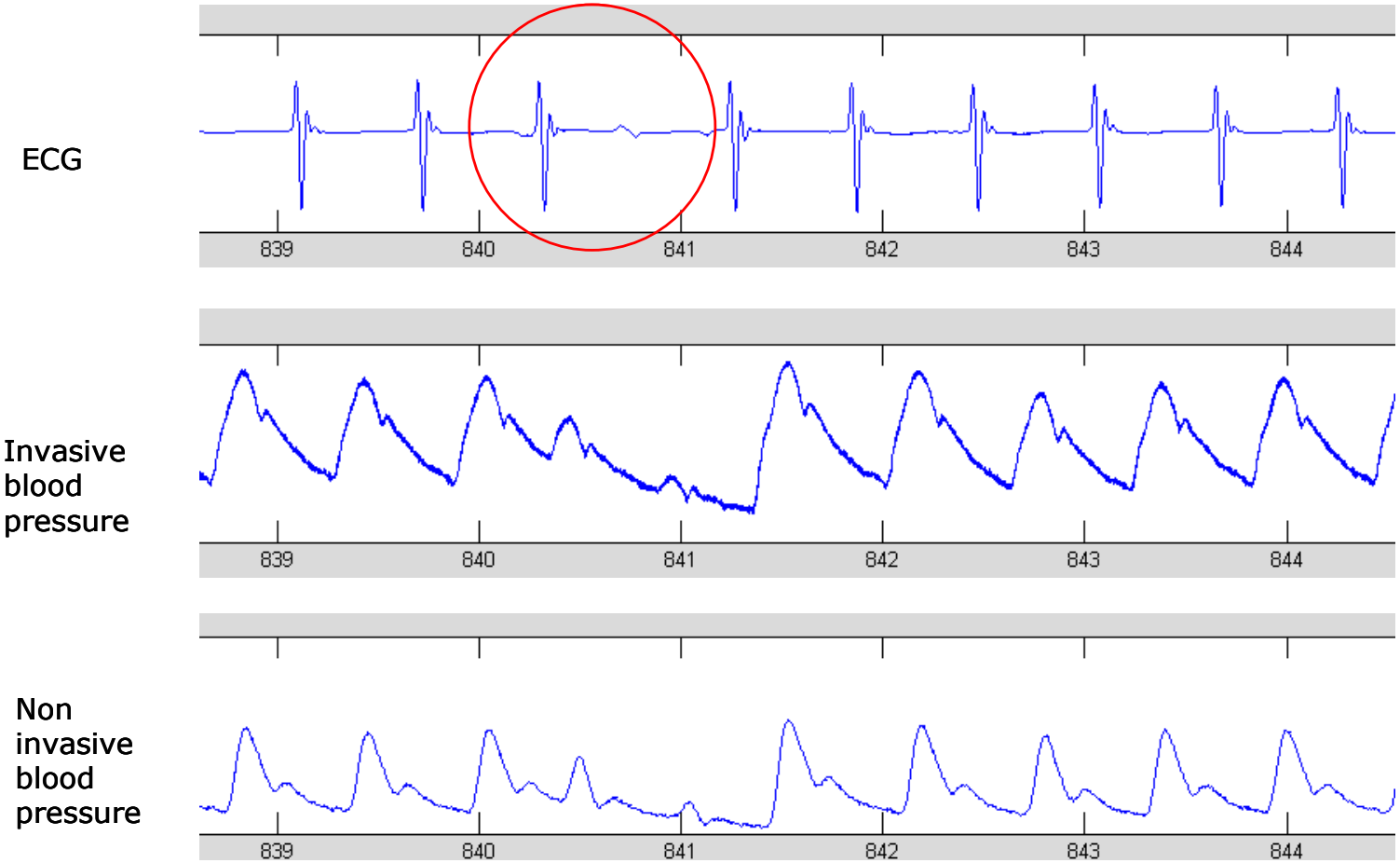


Figure 2-14 An example of an ectopic beat whilst recording ECG, invasive blood pressure and non invasive blood pressure.

3D echocardiography measured the dyssynchrony index by superimposing 4 beats to create a full volume image which is only reliable if there are at least 4 beats of similar RR lengths. Therefore ectopic beats will lead to “stitching artefacts” and these therefore needed to be excluded. AF patients and patients with arrhythmias could not therefore be included in the study as the technology here does not allow accurate recordings and readings which are truly representative of the hypothesis we are trying to demonstrate.

During invasive data collection there were times when the patient would move due to discomfort of being on the catheter laboratory table for a length of time. Therefore, on the matlab traces we can see that these are artefacts and not true recordings. Hence I have excluded these anomalies from the analysis to ensure that what is calculated and measured is a true reflection of the intra cardiac timing manipulations. I am aware that the beat-to-beat variability that is present is inherent and cannot be eliminated, but by ensuring we have a dataset which is firstly free of artefact i.e. either movement or equipment failure but still includes all other data, we can then get a true understanding of physiological changes and answers to the hypothesis we have queried.

2.9 Methods of Analysis – Acquiring data

2.9.1 Single measurements

When performing each of the studies, individual measurements were taken at each AV or VV delay. This was then repeated multiple times when using advanced echo modalities of 3D echocardiography (Chapter 5), TDI (Chapters 5 and 6), PW Doppler measurements (Chapters 5 and 7) and invasive measurements (Chapter 8).

These single measurements were analysed to assess agreement between replicate measurements within the same modality and between the different modalities. I then analysed this data to calculate the agreement, intra class correlation and scatter between replicate optimisations processes.

2.9.2 Averages of multiple measures

A major component of my study hinged on the concept of improving the signal-to-noise of the measurements to ensure optimisation is performed in the best way possible rather than randomly performed and hoping to improve patient outcomes according to chance. I have shown in Figure 2-15 in this schematic diagram how I have measured beats and averaged them.

A well know phenomenon to reduce the noise component is to average multiple values and hence improve the signal-to-noise quality (Le Cam L 1986; Pabari et al. 2011). This is the rationale for the many images I have acquired and analysed in this study. The theory behind this is explained in Chapter 4, and will demonstrate the effect on actual measurements in Chapters 5, 6 and 7.

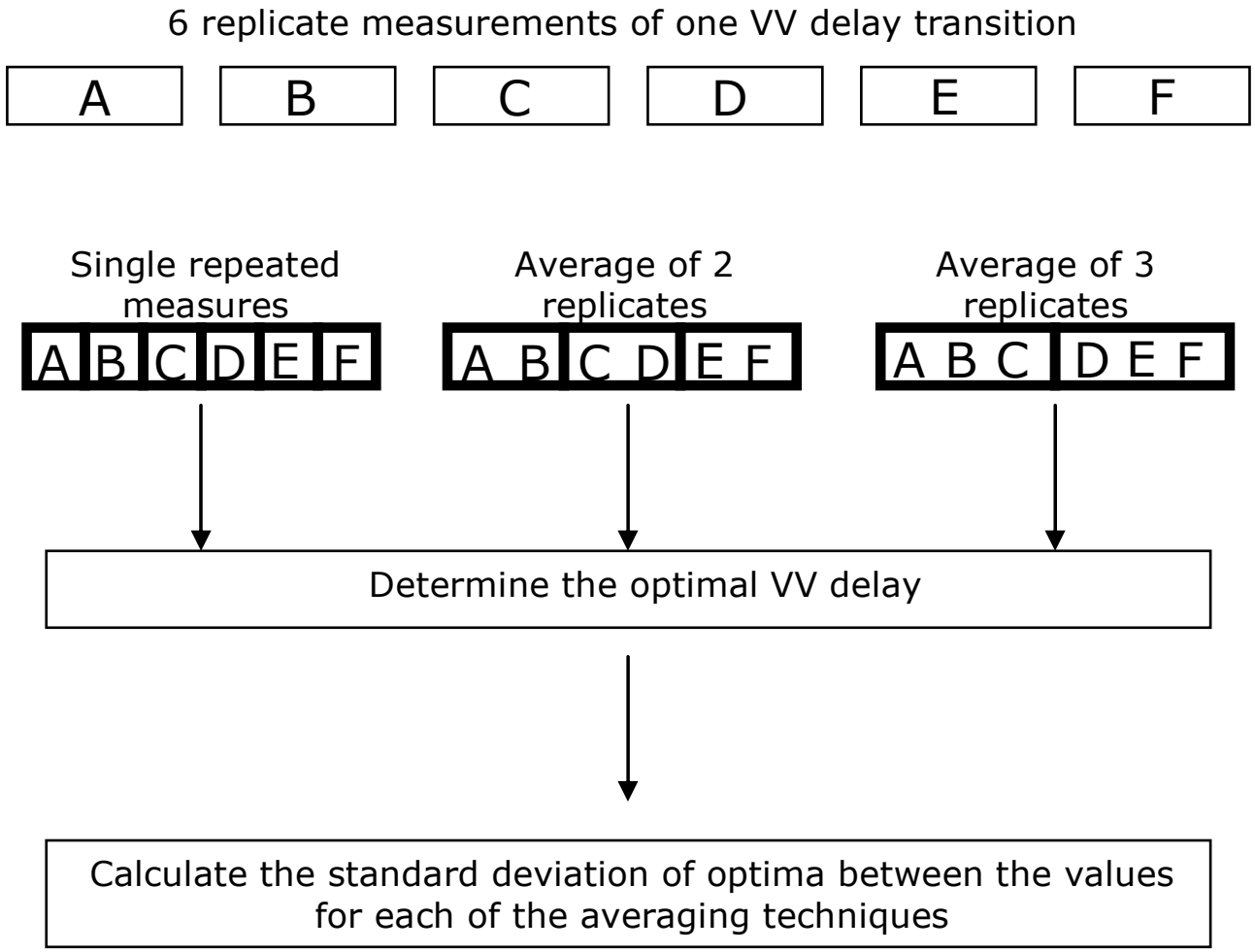


Figure 2-15 Schematic diagram showing how we measured single beats, average-of-two and averages-of-three to ensure each beat was independent. We then calculated the standard deviation from the optimum derived.

2.10 Calculation of Intraclass Correlation Coefficient

The intraclass correlation coefficient (ICC) quantifies the extent to which measurements differ between settings because of genuine difference between settings versus random biological variation. I used the ICC to quantify how well each modality could detect a genuine improvement arising from a change in VV delay, as distinct from background spontaneous beat-to-beat variability. An ICC close to 1 indicates that changes in settings make a relatively large difference to the measurement in comparison to random beat-to-beat variability. An ICC close to 0 indicates that random beat-to-beat differences are relatively large in comparison to the genuine effect of changing settings, Figure 2-16.

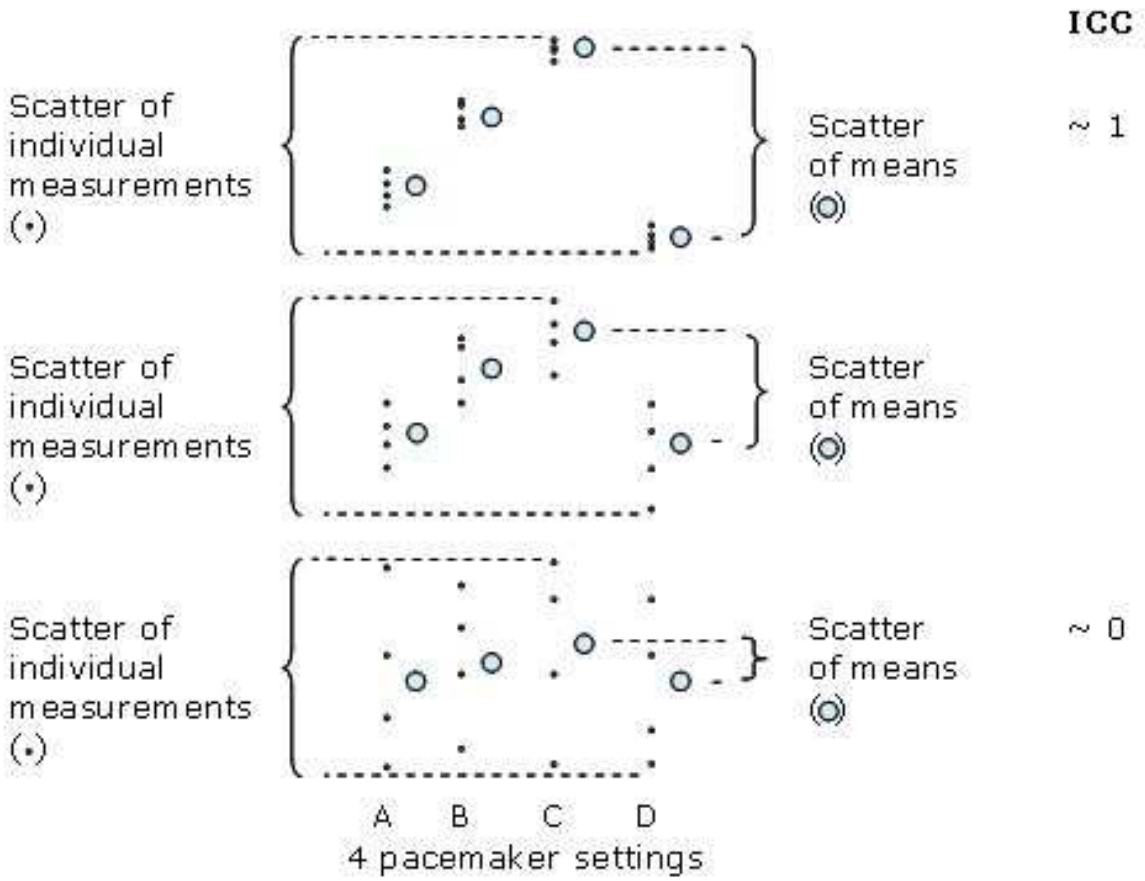


Figure 2-16 Intuitive description of intraclass correlation coefficient, as applied to dyssynchrony measures.
Replicate measures at the same setting are likely to have slightly different values (small dots). Averaging the measures from the setting gives the means (large blue dots). If the scatter between the means is of similar size to the scatter between all raw individual measurements (top panel), then reproducibility is very good i.e. ICC = 1. If the scatter between the means is much smaller than the scatter between all raw individual measurements, then reproducibility is poor i.e. ICC = 0 (bottom panel).

3 Limits to predictability of response to biventricular pacing from dyssynchrony indices, and systematic analysis of impact of study design on findings

3.1 Abstract

3.1.1 Background

Can markers of mechanical dyssynchrony strongly determine ventricular remodelling by biventricular pacing (cardiac resynchronisation therapy, CRT)? The aim of this chapter is to quantify the fundamental limits on the consistently-observable coefficient of determination (R^2), arising from variability between repeat echocardiographic measurements and between successive mechanical dyssynchrony measurements. It examines the literature and induces a small retrospective study in the test-retest distributions of changes in Ejection Fraction (EF) over time, when no intervention has occurred.

3.1.2 Method and Results

First, I quantified the mathematical depression of observable R^2 between dyssynchrony criteria and response arising from spontaneous biological variability of response markers over time, and test-retest variability of dyssynchrony measurements.

Second, I compared actual published R^2 values, between externally-monitored randomised controlled trials (EMRCTs) and highly-skilled-single-centre studies (HSSCSs).

Inherent variability in dyssynchrony markers and measures of ventricular remodelling such as ΔLVEF, ΔESV or ΔEDV means that even if a perfect dyssynchrony predictor of response is found (underlying $R^2=1$), observable R^2 between dyssynchrony and ΔLVEF cannot sustainably exceed 0.33 (for ΔESV, ≤0.34; for ΔEDV, ≤0.27). Interestingly, many R^2 values published in HSSCS exceed these mathematical limits whereas none in externally-monitored studies did so. Overall HSSCSs seem to

overstate R^2 by >5 to 20-fold (p=0.002), whether response is ΔLVEF, ΔESV or ΔEDV.

Finally, I assessed the spread of repeated ΔEF in a real-life population of general cardiology patients undergoing follow-up in a tertiary centre. We found 1298 patients reporting the distribution of ΔEF on measurements 3 to 12 months apart within a 5 year period. The standard deviation of ΔEF for these patients was 18.6%. This is wider than the scatter seen in the control arms of the RCT's.

3.1.3 Conclusions

Even in an elaborately conducted prospective study, it is not realistic to expect to find a significant correlation between ΔEF achieved by CRT and outcomes (or baseline dyssynchrony). It is unlikely to happen however skilled the operator, whatever the method of measuring EF and even if an alternative parameter to EF is used.

Even more unrealistic is to hope that analysis of routine clinical measurements will show responses that can be predicted. Overall, our findings suggest HSSCSs often overstated the predictability of response by mechanical dyssynchrony markers. In many cases they report values exceeding the mathematical limits imposed by spontaneous variability. Even if they continue to arise in the literature, markedly positive results will fail to be replicated in independent, blinded hands with neutral external monitoring. These results favour testing of mechanical dyssynchrony markers using scientific quality standards as strict as externally monitored RCTs, and suggest that prediction will be poor.

3.2 Background

There are questions raised regarding the ability and the reliability of dyssynchrony markers to predict response to biventricular pacemakers. In this study I have collaborated with my colleague Dr S. Nijjer, to perform a systematic review of the literature to assess the degree to which dyssynchrony markers can predict response to biventricular pacemakers. We evaluated both externally-monitored randomised controlled trials (EMRCTs) and highly-skilled-single-centre studies (HSSCSs).

I then reviewed our local datasets of echocardiograms performed in patients to identify and assess commonly measured echocardiographic parameters, such as ejection fraction to assess their test- retest reproducibility.

Biventricular pacing is thought to deliver its benefits in heart failure through resynchronisation of dyssynchronous cardiac mechanical function, hence the term "cardiac resynchronisation therapy" (CRT) (Abraham et al. 2002; Bristow et al. 2004; Cazeau et al. 2001; Cleland et al. 2005; Gras et al. 2002; Young et al. 2003) The correlation coefficient (r) and its square, the coefficient of determination (R^2), are commonly used to quantify the strength of the association between variables such as a baseline parameter and an outcome measure; for example, baseline mechanical dyssynchrony and the degree of reverse remodelling after biventricular pacing. The R^2 statistic is important because it indicates how much of the variability of future outcomes between patients can be correctly predicted from baseline information. R^2 is always positive, with values close to 1 indicating that all the differences in outcomes between patients are fully predictable from the baseline marker, and values close to 0 indicating no predictive value. Some studies (Bax et al. 2003a; Bleeker et al. 2007a; Yu et al. 2004c) demonstrate high R^2 values between baseline mechanical

dyssynchrony and echocardiographic outcome measures whereas other (Chung et al. 2008a; Marcus et al. 2005; Miyazaki et al. 2010) show much weaker relationships.

The failure of dyssynchrony markers to consistently predict response to biventricular pacing is a cause of concern. Unsurprisingly, mechanical dyssynchrony is not included in guidelines which instead rely on electrical dyssynchrony as denoted by wide QRS (Hawkins et al. 2009). Nonetheless, the quest for new and more robust mechanical dyssynchrony markers or more complex selection algorithms continues, in the hope of giving higher R^2 values.

In this study I test whether the current processes, of examining patient sets for variables that predicted outcomes well, is a rational strategy to solve to the problem of unreliable prediction of response. In particular, if a marker (or combination index) is found with excellent prediction of response, how high can its R^2 be when independently re-tested beyond the initial cohort or, when in real clinical practice? The answer cannot be $R^2=1$ because of unavoidable variability or uncertainty in measurements. But the precise impact of this variability has never been addressed. Both mechanical dyssynchrony markers (Palmieri et al. 2010; Vesely et al. 2008) and commonly used outcome markers of reverse remodelling have variability when re-tested in the same patient as opposed to re-analysis of a stored image. The clinically relevant measure of variability of any diagnostic test must capture all sources of variability. The complete test must be performed afresh, conducted and analysed with blinding to the first result.

I calculate the limit on R^2 by using randomised controlled studies and test-retest studies to determine for the first time the impact of these variability's on R^2. In this study I calculated the R^2 that one should realistically expect and provide a simple,

usable tool for clinicians and researchers to understand to what extent low test-retest reproducibility, and spontaneous changes in markers of LV function, attenuate associations in studies of biventricular pacing response.

Having reviewed the world literature, I also wanted to test whether data in my local hospital showed the same properties of surprisingly wide test retest reproducibility of repeated echocardiographic EF assessments. I found that in my hospital reproducibility was very wide meaning that there was certainly no chance of defining echocardiographic responders in any meaningful way, any therefore all claims of others saying that they have done so should be viewed sceptically.

This evaluation is fundamental in understanding and further unravelling to puzzle of dyssynchrony and response to cardiac resynchronisation therapy.

3.3 Methods

3.3.1 Quantitative separation of device-mediated, versus spontaneous, changes in LVEF

Patients undergoing biventricular pacing have two drivers that affect the pre-to-post change in the chosen echocardiographic outcome measure (e.g. change in ejection fraction, ΔLVEF). Firstly, inherent phenomena will contribute to individual patients' ΔLVEF even without device implantation. This includes true biological variability and measurement error intrinsic to the assessment technique. The size of this effect can be assessed by size of the variance (the square of standard deviation) in the control population in randomised controlled trials assessing biventricular pacing. Secondly, the device itself will impose an effect on outcome measure, over and above the inherent variation. Individual patients have differing responses to the device, some improving and some worsening. Therefore, the variance of the ΔLVEF should be larger in the device population of a randomised trial (Figure 3-1). If the variance of ΔLVEF is the same as in the control population, then the device had a perfectly uniform effect on that chosen marker across all patients.

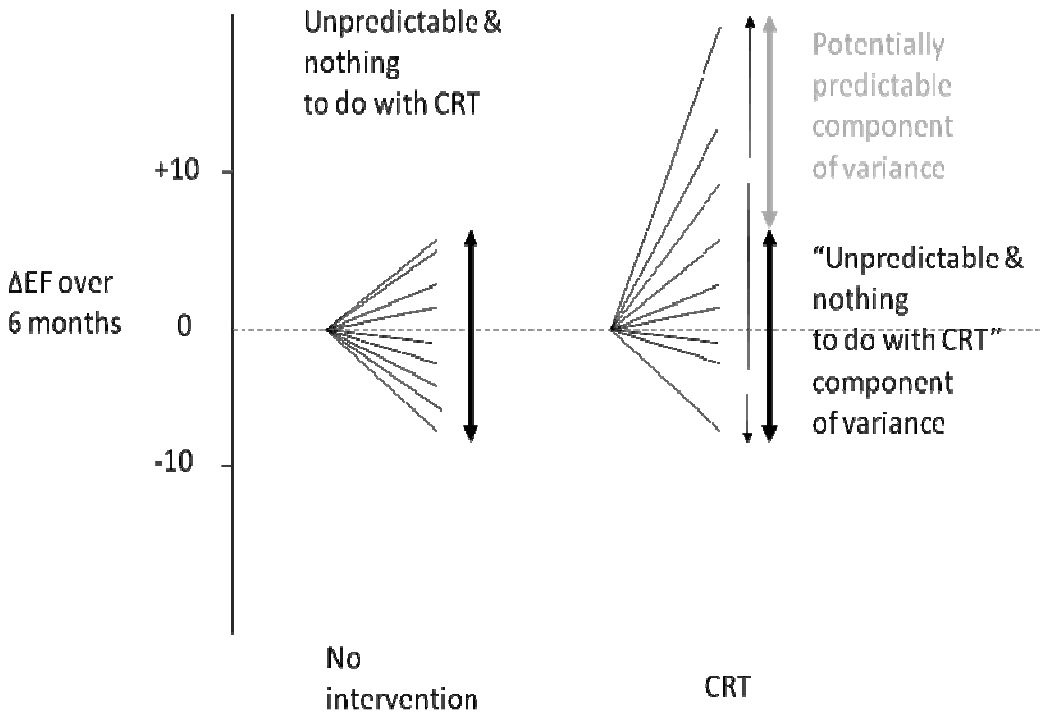


Figure 3-1 The variance of the intervention group is only partly due to the intervention.
Control populations in randomised controlled trials of biventricular pacing have changes in their outcome markers, such as LVEF, LVESV and LVEDV even without intervention. This change is measured by the variance (SD^2) of the change in LVEF (ΔLVEF) and represents inherent or unpredictable change (left panel). Those undergoing biventricular pacing will have a further change in LVEF over and above inherent changes and therefore have a wider variance in ΔLVEF (right panel). This is shown schematically as the grey section. Only this component is related to the device and therefore only this section of ΔLVEF is potentially predictable by any baseline marker.

Baseline dyssynchrony markers are typically only measured once but will also have an inherent variability within a given patient and across the population being studied. Only test-retest reproducibility studies can reveal the extent of this.

When correlating two variables such as mechanical dyssynchrony and echocardiographic outcome (*r*), or determining the predictive value of one on the other (R^2), then the variance of both combines in an additive fashion to depress the relationship (Equation 3-1). We have termed this the R^2 contraction factor. This is the maximal R^2 that can be calculated between those variables for that population. Below I have shown the derivation of the formula for contraction factor. Appendix 4 shows the derivation of the formula.

The contraction factor can be calculated easily if the standard deviation of the Δ in the outcome measure is known for both the control arms and device arms of a randomised control trial are known. It is not sufficient to know the distribution of the initial and final LVEFs. Rather, the distribution *of the change*, i.e. the standard deviation of Δ, is needed. This can be used in the following calculation:

$$= 1 - \left(\frac{\sigma^2(\Delta control_arm)}{\sigma^2(\Delta intervention_arm)}\right)$$

$$= 1 - \left(\frac{\sigma(\Delta control_arm)}{\sigma\ (\Delta intervention_arm)}\right)^2$$

Equation 3-1 R^2 contraction factor caused by inherent variability in LVEF, where σ is standard deviation and σ^2 is variance.

This formula can also be used for the baseline mechanical dyssynchrony measure.

The two contraction factors are then multiplied to determine the combined contraction factor. The observed R^2 is then calculated as shown in **Equation 3-2**.

$$\text{Observed } R^2 = \text{Underlying } R^2 \times \underbrace{R^2 \text{ contraction factor imposed by } \textit{Dyssynchrony Marker} \times R^2 \text{ contraction factor imposed by } \textit{Response Marker}}_{\text{Combined } R^2 \text{ Contraction Factor}}$$

Equation 3-2 The combined contraction factor which takes into consideration the R^2 contraction factor due to the dyssynchrony marker and the response marker

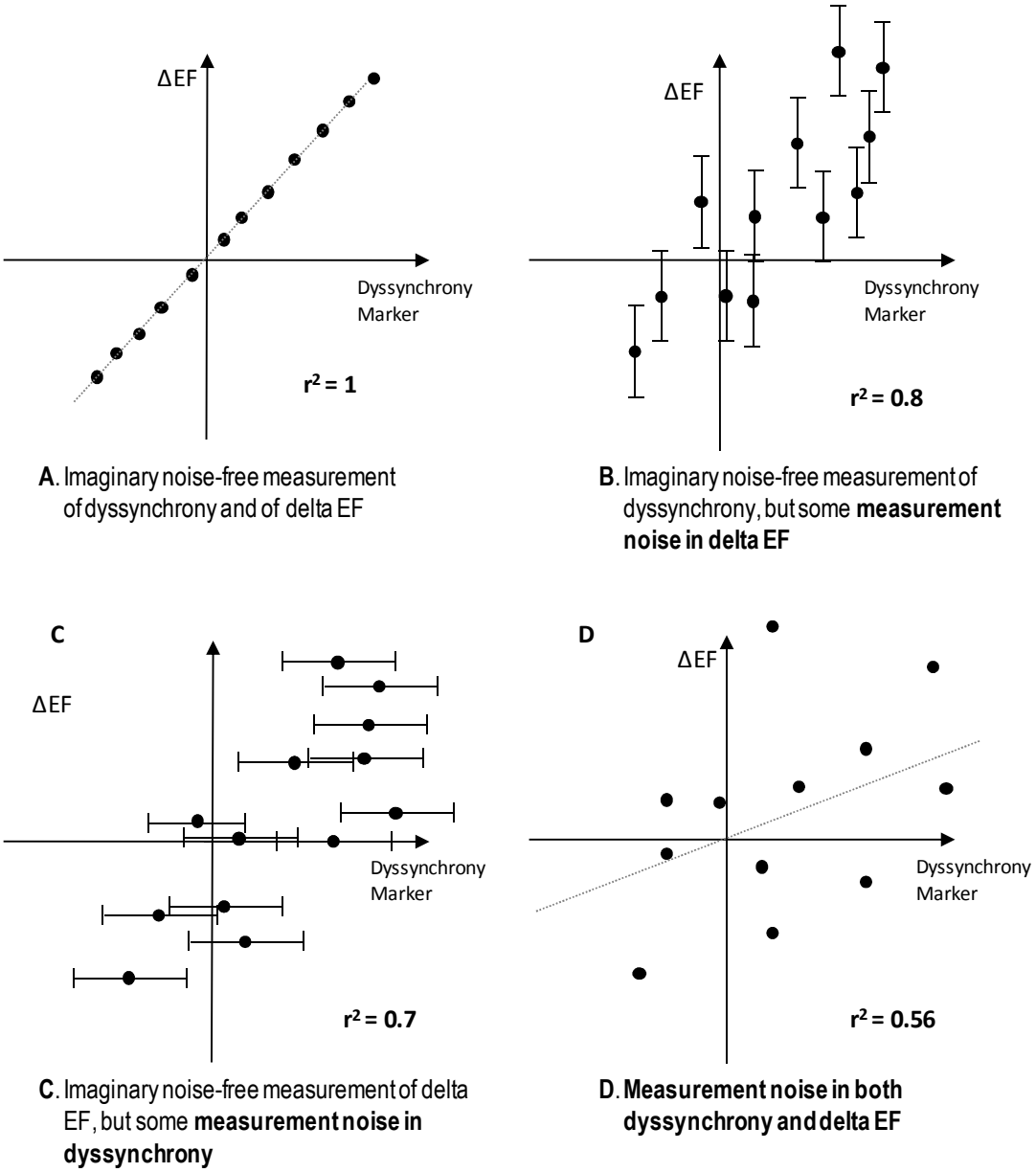


Figure 3-2 How inherent variability in two measures reduces the maximum achievable R^2 between them.
Imagine a dyssynchrony marker that can perfectly predict response, as long as measurement noise could somehow be eliminated (panel A). In practice there is natural variability in measurements of EF (panel B) and of the dyssynchrony marker (panel C). These noise properties combine together multiplicatively to depress the actually-observable R^2 value (panel D). In this example case, it is mathematically impossible for R^2 value over 0.56 to be observed sustainably.

The following provides a worked example. In the MIRACLE-ICD II trial (St John Sutton et al. 2003), in the control arm ΔLVEF has a standard deviation of 6.2. In the biventricular pacing arm, ΔLVEF has a standard deviation of 8. The SD of ΔLVEF is larger in the device arm, because additional change is occurring over and above the inherent variability in LVEF. The effect of spontaneous, inherent variation in LVEF can be calculated by measuring the R^2 contraction factor as $1-(6.2/8)^2 = 0.40$. That is, in the MIRACLE-ICD population, if one attempted to assess any baseline marker as a predictor of ΔLVEF, the maximum R^2 that could be determined would be 0.40.

If a mechanical dyssynchrony marker was used to predict the ΔLVEF, then we must assess the contraction factor applied by the variability in the mechanical dyssynchrony marker and multiply together with the 0.40 already calculated. For example, if the mechanical dyssynchrony marker has a contraction factor of 0.50, then the maximum observable R^2 between ΔLVEF and the mechanical dyssynchrony marker will be $0.40 \times 0.50 = 0.20$. Therefore, regardless of the true underlying relationship between dyssynchrony and the ΔLVEF, the maximum R^2 that can be measured will be 0.20 with these example figures.

This calculation is distinct from the proportion of the mean effect that is caused by the device: that proportion is simply 1 minus the ratio of the mean effect in the control arm divided by the mean effect in the device arm. This may be as high as 100%, even if the proportion of between-patient variability caused by different responses to the device is far lower. The two are independent phenomena.

3.3.2 Data extraction from published studies

A systematic review of studies assessing the response to biventricular pacing was performed using EMBASE and Medline databases (Figure 3-3). The terms 'cardiac resynchronisation therapy', 'biventricular pacing' and 'dyssynchrony' were used and abstracted reviewed for relevance. Further papers were identified through review of citations including the extensive review by Hawkins *et al*, 2009 (Hawkins et al. 2009).

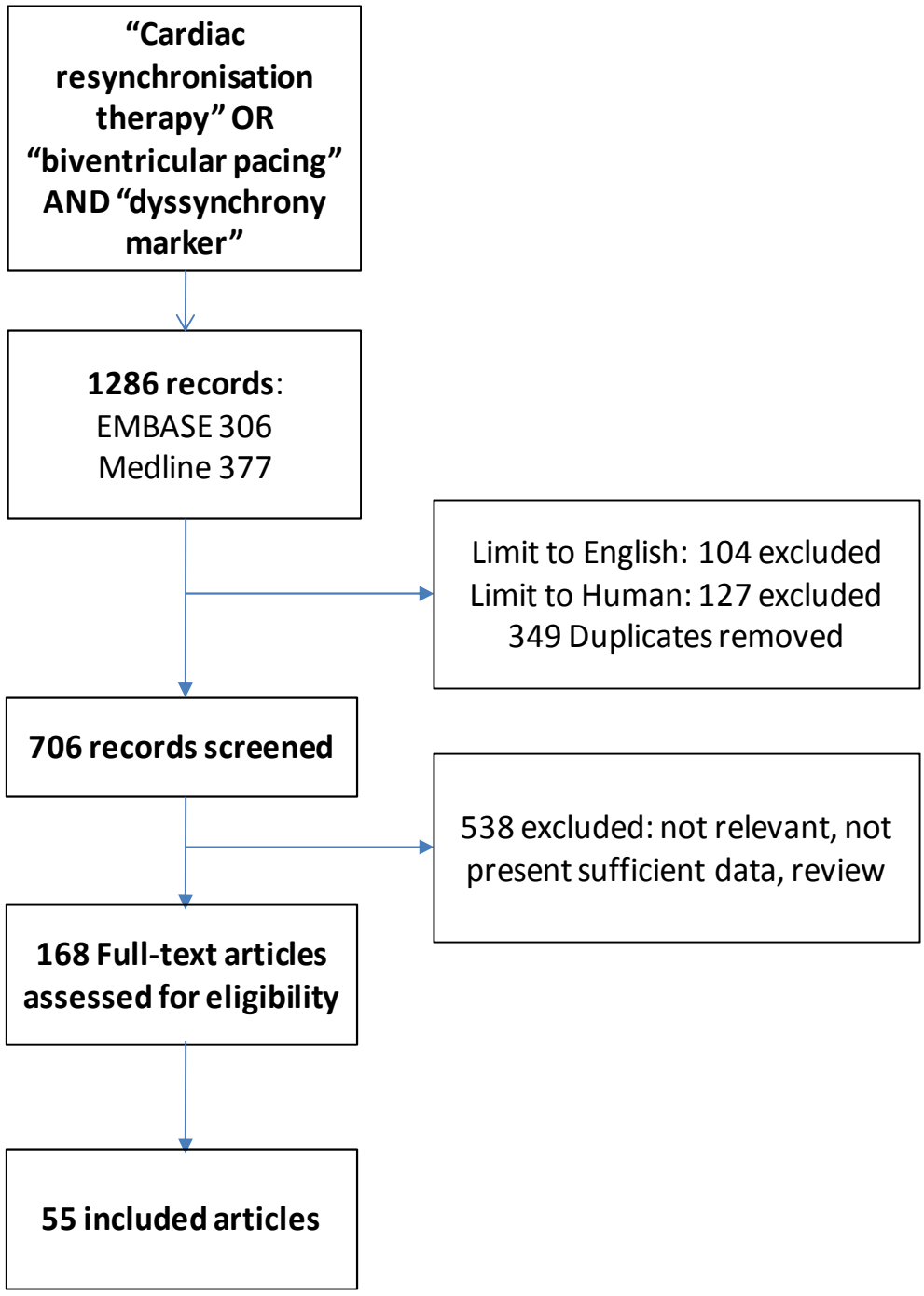


Figure 3-3 Systematic search strategy

All published studies that assessed mechanical dyssynchrony markers against the change in LVEF (ΔLVEF), ΔLVESV and ΔLVEDV were analysed in detail and data extracted (Abraham et al. 2004; Bank et al. 2009; Bax et al. 2003a; Bleeker et al. 2007a; Bleeker et al. 2007b; Bordachar et al. 2010; Conca et al. 2009; De Boeck et al. 2008; Delgado et al. 2008; Deplagne et al. 2009; Diaz-Infante et al. 2007; Duncan et al. 2006; Faletra et al. 2009; Gorcsan et al. 2004a; Gorcsan et al. 2007; Kaufman et al. 2010; Liodakis et al. 2009; Marcus et al. 2005; Marsan et al. 2008; Mele et al. 2006; Norisada et al. 2010; Notabartolo et al. 2004; Park et al. 2010; Penicka et al. 2004; Pitzalis et al. 2002; Pitzalis et al. 2005; Porciani et al. 2006a; Sassone et al. 2007; Soliman et al. 2007b; Soliman et al. 2009; St John Sutton et al. 2003; Suffoletto et al. 2006; van Bommel et al. 2010b; Van de Veire et al. 2007; Wang et al. 2010; Yu et al. 2003; Yu et al. 2004a; Yu et al. 2005; Yu et al. 2007). We extracted R^2 or calculated it from the correlation coefficient between the mechanical dyssynchrony marker and outcome measure, using the published data in tabular, text or graphical form. The weighted averages of the R^2 were calculated using the size of the study.

The landmark EM-RCTs of biventricular pacing were also assessed, specifically to compare the spread (standard deviation) of ΔLVEF, ΔLVESV and ΔLVEDV in the control and intervention arms (Abraham et al. 2004; Beshai et al. 2007; Cappola et al. 2006; Cleland et al. 2008; Cleland et al. 2005; Foley et al. 2011; Ghio et al. 2006; Ghio et al. 2009; Linde et al. 2008; Lubitz et al. 2010; Moss et al. 2009b; Solomon et al. 2010; St John Sutton et al. 2003; Sutton et al. 2006b; Wikstrom et al. 2009). The spread of the Δ in the control arm shows the intrinsically unpredictable element of the spread of Δ in the intervention arm.

3.3.3 Data from my hospital's database assessing ejection fraction at repeated studies

1298 patients who attended a general cardiology clinic at a teaching hospital in London who had more than one echocardiogram were included. All the echocardiograms were stored on Medcon systems with LV measurements and calculated ejection fractions at each attendance were compared. I looked to see how ΔEF changed with each successive measurement. Each measurement was compared to the first measurement and each preceding measurement.

Patients were included if they had repeat echocardiograms which enabled me to calculate the EF. This was not a study designed to perform the most robust and reproducible measurements using a study protocol, but rather a review of the echocardiograms performed in standard clinical practice. All patients had at least two echocardiograms performed. Some had further echocardiograms - up to 3, 4, 5 or 6 echocardiograms between July 2004 and June 2009.

To find the data, a search was performed on the Medcon database using the instructions shown in Figure 3-4.

3.3.4 Medcon search utility

Go to http://connserver/crn Log on using Username and Password

Click Query Studio (top left)

3.3.4.1 Part I: Create report layout

Under the 'Insert Data' tab, drag and drop fields that you want to be in your final report. This should include fields that you will be searching in, as well as patient identifiers and any other Echo data you may need.

To locate fields, go to 'Echo 2'

For demographic details, look under Common

For echo data, look under Echo

When you drag and drop them across to the main frame, a new column will form.

3.3.4.2 Part II: Create search

1. To search, click Edit Data in the top left
2. Click the column you wish to search in
3. Click Filter
4. Search criteria will come up at the bottom of the main frame.
5. You can add multiple search criteria.
6. The search will then run automatically.
7. When finished, click 'Manage File' and export to a CSV or XLS sheet.

Figure 3-4 Step by step method to search Medcon to perform searches for echocardiograms as performed in this study

3.3.5 Statistics

Values are shown as mean (95% confidence interval), except where otherwise indicated. Comparisons between classes of study were made using Student's unpaired t test and the Mann-Whitney U test. A p value of <0.05 was predefined as statistically significant. Stata/SE 10.0 for Windows (Stata Corp LP; College Station, Tex) were used to perform the statistical analysis.

3.4 Results

3.4.1 Reported R^2 for echocardiographic response in EMRCTs and HSSCs

52 reports were identified and assessed. The majority were retrospective cohort studies either with or without matched controls, performed in highly skilled single centres (HSSC) with specific interest in echocardiographic dyssynchrony markers and a track record of innovation in the field (Abraham et al. 2004; Bank et al. 2009; Bleeker et al. 2007b; Bordachar et al. 2010; Conca et al. 2009; De Boeck et al. 2008; Delgado et al. 2008; Deplagne et al. 2009; Diaz-Infante et al. 2007; Duncan et al. 2006; Faletra et al. 2009; Gorcsan et al. 2004a; Kaufman et al. 2010; Liodakis et al. 2009; Marcus et al. 2005; Marsan et al. 2008; Mele et al. 2006; Norisada et al. 2010; Notabartolo et al. 2004; Park et al. 2010; Penicka et al. 2004; Pitzalis et al. 2005; Porciani et al. 2006b; Sassone et al. 2007; Soliman et al. 2007b; Soliman et al. 2009; St John Sutton et al. 2003; Suffoletto et al. 2006; van Bommel et al. 2010b; Van de Veire et al. 2007; Wang et al. 2010; Yu et al. 2003; Yu et al. 2005). The R^2 reported in these studies between individual dyssynchrony markers and echocardiographic response to biventricular pacing (ΔLVEF, ΔLVESV or ΔLVEDV) are tabulated in Table 3-1.

Table 3-1 A comparison of the baseline dyssynchrony variables found to predict response in externally monitored randomised controlled trials with those found highly skilled single centre studies. Correlation coefficients and coefficient of determination (R^2) are presented. Abbreviations are those used in the original publication

Cardiac Response	Trial	Baseline Variable	n	Observed correlation coefficient	R^2	Trial	Baseline Variable	n	Observed correlation coefficient	R^2
	Highly skilled single centre studies					**Externally monitored Randomised Controlled Trials**				
ΔLVEF	Bax 2003	TDI Septal-lateral delay	25	0.47	0.22	CONTAK: Marcus 2005	SPWMD	79	-0.11	0.01
	Pitzalis 2005	SPWMD	51	0.69	0.48	MADIT-CRT: Pouleur 2011	Transverse Strain Dyssynchrony	761	-0.29	0.08
	Marsan 2008	Systolic dyssynchrony index	56	0.7	0.49					
		Weighted Average			**0.42**		**Weighted Average**			**0.07**
%ΔLVEF	Penicka 2004	Sum asynchrony	49	0.73	0.53					
	Mele 2006	SPWMD	37	0.07	0.005	*No EM-RCT reported this outcome*				
		TPS-SD	37	0.86	0.74					
		SPWTD	37	0.53	0.28					
		Weighted Average			**0.39**		**Weighted Average**			**-**
ΔLVEDV	Pitzalis 2002	SPWMD	20	-0.73	0.53	MIRACLE: Cappola 2006	Mitral regurgitation index	776	0.0022	0.000005
						MIRACLE: Cappola 2006	QRS Width	776	0.12	0.01
						CONTAK: Marcus 2005	SPWMD	79	-0.14	0.02
						MADIT-CRT: Pouleur 2011	Transverse strain dyssynchrony	761	0.25	0.06
		Weighted Average			**0.53**		**Weighted Average**			**0.02**
ΔLVESV	Pitzalis 2002	SPWMD	20	-0.7	0.49	CARE-HF: Ghio 2009	IVMD	735	Not published	-
	Yu 2003	Ts-SD	30	-0.76	0.58	CONTAK: Marcus 2005	SPWMD	79	-0.1	0.01
	Bax 2004	TDI Septal-lateral delay	80	0.84	0.71	REVERSE: Linde 2009	IVMD	487	Not published	-
	Yu 2004	TDI-Ts-SD	54	-0.74	0.55	MADIT-CRT: Pouleur 2011	Transverse strain Dys	761	0.25	0.06
		TDI-Ts-12	54	-0.6	0.36					
		SRI-Tsr-SD	54	-0.01	0.00					
		PSS-12	54	-0.28	0.08					
	Yu 2005	Ts-SD-12-ejection	56	-0.61	0.37					
		Ts-SD-6-ejection	56	-0.52	0.27					
		Ts-12-ejection	56	-0.6	0.36					
		Ts-6-ejection	56	-0.53	0.28					
	Porciani 2006	Ts-SD	59	-0.32	0.10					
		oExcT	59	-0.48	0.23					
	Yu 2006	TDI-SD-12	55	-0.76	0.58					
		TDI-SD-6	55	-0.63	0.40					
		Diff-12	55	-0.64	0.41					
		Diff-6	55	-0.62	0.38					
		Sep-Lat	55	-0.54	0.29					
		Sep-Post	55	-0.49	0.24					
	Marsan 2008	Systolic dyssynchrony	56	0.6	0.36					
	Bank 2009	IVCT	64	-0.24	0.06					
		TT_{S-L} Delay	64	-0.14	0.02					
		TVI_{S-L} Delay	64	-0.12	0.01					
		SD Rad-6	64	-0.2	0.04					
	Soliman 2009	SDI (3DE)	90	0.56	0.31					
	Van Bommel 2010	LV Dyssynchrony	361	0.38	0.14					
	Miyazaki 2010	SPWMD	117	0.32	0.10					
		S-L Delay	117	-0.05	0.00					
		Tv-SD	117	0.15	0.02					
		Tε-SD	117	0.38	0.14					
		Tε-dif	117	0.43	0.18					
		Weighted Average			**0.24**		**Weighted Average**			**0.05**
%ΔLVESV	Mele 2006	SPWMD	37	-0.42	0.18					
		TPS-SD	37	-0.01	0.0001	*No EM-RCT reported this outcome*				
		SPWTD	37	-0.73	0.53					
	Delgado 2008	AS-P (radial strain)	161	0.41	0.17					
		SDt_{65s}	161	0.26	0.07					
	Wang 2010	RDI-6 basal segments	30	0.42	0.18					
		RDI-6 mid-LV segments	30	0.75	0.56					
		RDI-12 (combination)	30	0.62	0.38					
		Weighted Average			**0.22**		**Weighted Average**			**-**
ΔNYHA	No HSSCS reported this outcome					MIRACLE: St John Sutton 20	IVMD	323	0.12	0.01
ΔQoL	No HSSCS reported this outcome					MIRACLE: St John Sutton 20	IVMD	323	0.19	0.04
							Weighted Average			**0.03**

Table 3-2 Calculation of the contraction factor for ΔLVEF, ΔLVESV and ΔLVEDV in externally monitored randomised controlled trials assessing biventricular pacing. The maximum possible R^2 for any parameter to the ΔEF in each study is presented for each study. Δ represents change. SDD – standard deviation of difference (Δ)

Response measure	Study		Breakdown of variability		Mandatory ceiling on R^2 value imposed solely by unpredictable variability in response measure	
		Total N	Unpredictable element of variability SDD of control arm	Total variability in intervention arm SDD of intervention arm	Calculation	Point estimate of maximal achievable R^2 value
ΔLVESV	MIRACLE ICD II	153	57	77	$1-(57/77)^2$	0.45
	CARE-HF	735	42.8	66.6	$1-(42.8/66.6)^2$	0.59
	MADIT-CRT	1366	16.3	31.2	$1-(16.3/31.2)^2$	0.73
	REVERSE	487	23.4	29.5	$1-(23.4/29.5)^2$	0.37
	RETHINQ	142	5.1	5.61	$1-(5.1/5.61)^2$	0.17
	RESPOND	55	44.1	46.9	$1-(44.1/46.9)^2$	0.12
	Variance Weighted Average					**0.50**
ΔLVEDV	MIRACLE ICD II	154	62	76	$1-(62/76)^2$	0.33
	CARE-HF	735	50.7	74.5	$1-(50.7/74.5)^2$	0.54
	MADIT-CRT	1366	14.4	33.2	$1-(14.4/33.2)^2$	0.81
	REVERSE	487	28	33.4	$1-(28/33.4)^2$	0.3
	RETHINQ	142	7.14	5.36	$1-(7.14/5.36)^2$	-0.77
	RESPOND	55	47.6	51.9	$1-(47.6/51.9)^2$	0.16
	Variance Weighted Average					**0.40**
ΔLVEF	CONTAK-CD	490	10.3	10.4	$1-(10.3/10.4)^2$	0.02
	MIRACLE ICD II	153	6.2	8	$1-(6.2/8)^2$	0.4
	CARE-HF	735	4.5	8.5	$1-(4.5/8.5)^2$	0.72
	MADIT-CRT	1366	3	5	$1-(3/5)^2$	0.64
	REVERSE	487	6.5	9.3	$1-(6.5/9.3)^2$	0.51
	RETHINQ	142	0.99	1.22	$1-(0.99/1.22)^2$	0.34
	RESPOND	55	11.7	15.7	$1-(11.7/15.7)^2$	0.44
	Variance Weighted Average					**0.48**

The ΔLVESV, ΔLVEDV, ΔLVEF values in RETHINQ had distinctive properties that make them potentially unsuitable for this analysis. First, the test-retest reproducibility (i.e. delta in the control arm) had a variance that was an order of magnitude smaller than that of all of the other studies. Second, the distributions of delta values for these echo markers where observers had a choice of beat to analyse, were skewed, with a relative curtailment of tail trending towards deterioration, and relative extension of the tail trending towards improvement. These phenomena were not present in the markers that were continuous variable where only a single measurement was realistically possible, such as exercise duration. Although the study carefully blinded assessors to allocation, the report did not state that it blinded them to prior measurements. Reluctance to state values far from prior measurements, together with the understandable preference in real-life clinical practice to prefer to document improvement, might have been the reason for these anomalies.

I calculated the contraction values for ΔLVESV, ΔLVEDV and ΔLVEF from the externally monitored RCT's and have shown the maximum R^2 values in Table 3-2. This formula can also be applied to non echocardiographic markers of response as shown in Table 3-3.

Table 3-3 R^2 contraction factor for non-echocardiographic markers of response to biventricular pacing. The maximal achievable R^2 column shows the maximum R^2 that any predictor could find when correlated against any of these outcome measures. Δ represents change. SDD - standard deviation of difference (Δ). Abbreviations: 6MWD - six minute walk distance; peak VO_2 - peak oxygen consumption; V_E/V_{CO2} - ratio of minute ventilation (V_E) and minute production of CO_2 (V_{CO2}), a measure of ventilatory response to exercise.

Response measure	**Study**	**Breakdown of variability**		**Mandatory ceiling on R^2 value imposed solely by unpredictable variability in response measure**	
		Unpredictable element of variability SDD of control arm	Total variability in intervention arm SDD of intervention arm	Calculation	Point estimate of maximal achievable R^2 Value
6MWD (m)	COMPANION	93	96	$1-(93/96)^2$	0.06
	CONTAK	103.8	104.8	$1-(103.8/104.8)^2$	0.02
	MIRACLE	98	109	$1-(98/109)^2$	0.19
Peak VO_2 (ml/kg/min)	CONTAK	4.3	4.4	$1-(4.3/4.4)^2$	0.05
	MIRACLE	3.2	3.2	$1-(3.2/3.2)^2$	0.00
Quality of Life Score	CONTAK	30	30.6	$1-(30/30.6)^2$	0.04
	MIRACLE	21.7	25.1	$1-(21.7/25.1)^2$	0.03
Quality of Life Improvement (%)	COMPANION	23	26	$1-(23/26)^2$	0.22
V_E/V_{CO2}	MIRACLE	5.2	6.2	$1-(5.2/6.2)^2$	0.30

3.5.1 Impact of R^2 contraction factor arising from the dyssynchrony variable

I assessed the published variability of mechanical dyssynchrony markers between repeated echocardiograms in the same patient (test-retest reproducibility, Table 3-4). The full variability is not reflected in simple re-measurements of identical previously-acquired images because this omits biological variability over time.

There are two studies reporting in detail the test-retest inherent variability of mechanical dyssynchrony on entirely separate echocardiograms (Table 3-4). In one study (Palmieri et al. 2010) within-patient variation and between-patient variation was small when performed in specialist hands. The data for SD within one patient is from their Table 3, and the SD across the population is from their Table 1. From these, the R^2 contraction factor due to the mechanical dyssynchrony marker can be calculated as $1-(SD_{\text{within patient}} / SD_{\text{between patient}})^2$. The other study (Vesely et al. 2008) assessed test-retest reliability of TDI mechanical dyssynchrony markers and presents the R^2 contraction factor directly. Where test-retest variability of a baseline variable is known in the form of a correlation coefficient r, that coefficient is a suitable estimate of the R^2 contraction factor if that variable is used as a predictor of a response. In (Vesely et al. 2008), r between test and retest of 2-segment dyssynchrony is 0.26 for reader A, and 0.43 for reader B, giving an averaged test-retest r value of 0.35 (Table 3-4).

Table 3-4 Test-retest variability of dyssynchrony markers within individuals, in populations who are candidates for biventricular pacemaker implantation. Most studies did not report test-retest reproducibility.

			Manatory Ceiling on R2 arising from variability in Dyssynchrony Marker		
Dyssynchrony marker	**Study**	**Within Patient SD**	**Between Patient SD**	**Estimated limit on R^2**	**R^2**
Inter ventricular mechanical delay					
Pulsed flow Doppler	Burri *et al* 2007	not reported	26	Incalculable	
	Duncan *et al* 2006	not reported	22	Incalculable	
	De Boeck *et al* 2008	not reported	25	Incalculable	
	Bordachar et al 2010	not reported	29	Incalculable	
Intra ventricular mechanical delay					
Pulsed tissue Doppler					
onset of systolic motion - 2 segment	Burri *et al* 2007	not reported	20	Incalculable	
onset of systolic motion - 4 segment	Penicka *et al* 2004	not reported	37	Incalculable	
onset of systolic motion - (S-L)	Bleeker *et al* 2007	not reported	49	Incalculable	
peak of systolic motion - 2 segment	Burri *et al* 2007	not reported	34	Incalculable	
onset of systolic motion	Penicka *et al* 2004	not reported	37	Incalculable	
onset of systolic motion	Burri *et al* 2007	not reported	29	Incalculable	
peak of systolic motion	Burri *et al* 2007	not reported	33	Incalculable	
peak of systolic motion - 4 basal segments	Bax 2004	not reported	49	Incalculable	
peak of systolic motion - 12 segments	Yu *et al* 2003	not reported	17	Incalculable	
onset of systolic motion - 2 basal segments (S-L)	Soliman *et al* 2007	not reported	48	Incalculable	
peak of systolic motion - 2 basal segments (S-L)	Soliman *et al* 2007	not reported	44	Incalculable	
onset of systolic motion - 2 basal segments (AL-IS)	Palmieri *et al* 2010	32	42	$1-(32/42)^2$	= 0.42
onset of systolic motion - 2 basal segments (A-I)	Palmieri *et al* 2010	22	27	$1-(22/27)^2$	= 0.34
onset of ejection	Foley *et al* 2011	29	76	$1-(29/76)^2$	=0.85
Colour Tissue Doppler					
peak of systolic motion - 2 basal segments (S-L)	Van Bommel et al 2010	not reported	51	Incalculable	
peak of systolic motion - 2 basal segments (S-L)	Conca *et al* 2009	not reported	50	Incalculable	
peak of systolic motion - 2 basal segments (S-L)	Yu *et al* 2007	not reported	54	Incalculable	
peak of systolic motion - 2 basal segments (S-L)	Faletra *et al* 2009	not reported	75	Incalculable	
peak of systolic motion - 2 opposite segments	Yu *et al* 2007	not reported	32	Incalculable	
peak of systolic motion - 2 segment	Shanks *et al* 2010	not reported	42	Incalculable	
peak of systolic motion - 2 segment	Conca *et al* 2009	not reported	50	Incalculable	
peak of systolic motion - 2 segment	De Boeck *et al* 2008	not reported	57	Incalculable	
peak of systolic motion - 2 segment	Vesley *et al* 2008	not reported	not reported	-	0.35
peak of systolic motion - 4 segment	Vesley *et al* 2008	not reported	not reported	-	0.46
peak of systolic motion - 6 segment	Notabartolo *et al* 2004	not reported	137	Incalculable	
peak of systolic motion - 6 segment	Conca *et al* 2009	not reported	41	Incalculable	
peak of systolic motion - 12 segment	De Boeck *et al* 2008	not reported	16	Incalculable	
peak of systolic motion - 12 segment	Deplagne et al 2009	not reported	9	Incalculable	
peak of systolic motion - 12 segments	Yu *et al* 2007	not reported	39	Incalculable	
peak of systolic motion - 12 segments SD	Conca *et al* 2009	not reported	15	Incalculable	
peak of systolic motion - 12 segments SD	Yu *et al* 2004	not reported	13	Incalculable	
peak of systolic motion - 12 segments SD	Faletra *et al* 2009	not reported	19	Incalculable	
peak of systolic motion - 12 segments SD	Yu *et al* 2007	not reported	15	Incalculable	
peak of systolic motion - 12 segments SD	Van de Veire *et al* 2007	not reported	16	Incalculable	
peak of systolic motion - 2 basal segments (AL-IS)	Palmieri *et al* 2010	23	81	$1-(23/81)^2$	= 0.92
peak of systolic motion - 2 basal segments (A-I)	Palmieri *et al* 2010	29	105	$1-(29/105)^2$	= 0.92
peak of systolic motion - 12 segments SD	Palmieri *et al* 2010	10	22	$1-(10/22)^2$	= 0.79
peak of disastolic motion - 2 segment	Shanks *et al* 2010	not reported	49	Incalculable	
Tissue synchronization imaging	Gorcsan *et al* 2004	not reported	158	Incalculable	
	Conca *et al* 2009	not reported	15	Incalculable	
	Faletra *et al* 2009	not reported	35	Incalculable	
M-Mode					
Septal posterior wall motion delay (SPWMD)	Bleeker *et al* 2007	not reported	119	Incalculable	
	Pitzalis *et al* 2002	not reported	92	Incalculable	
	Pitzalis *et al* 2005	not reported	96	Incalculable	
	Diaz-Infante *et al* 2007	not reported	113	Incalculable	
	Sassone *et al* 2007	not reported	46	Incalculable	
	Foley *et al* 2011	91.7	99.8	$1-(91.7/99.8)^2$	= 0.16
Lateral wall postsystolic displacement (LWPSD)	Sassone *et al* 2007	not reported	24	Incalculable	
3D Systolic dyssynchony index	Marsan et al 2008	not reported	0	Incalculable	
	Faletra et al 2009	not reported	3	Incalculable	
	Conca *et al* 2009	not reported	1	Incalculable	
	Deplagne *et al* 2009	not reported	0	Incalculable	
	Liodakis *et al* 2010	not reported	0	Incalculable	
	Soliman *et al* 2009	not reported	0	Incalculable	
Transverse strain					
time to peak transverse strain - 12 segments SD	Pouleur et al 2011	64	63	$1-(91.7/99.8)^2$	= - 0.03
LV pre-ejection period	Duncan *et al* 2006	not reported	16	Incalculable	
	De Boeck *et al* 2008	not reported	26	Incalculable	
	Bordeachar et al 2010	not reported	41	Incalculable	
RV pre-ejection period	Duncan *et al* 2006	not reported	10	Incalculable	
Combined inter and intra ventricular dyssynchrony					
Pulsed tissue Doppler	Penicka *et al* 2004	not reported	65	Incalculable	

3.6 True limits to correlation between baseline and delta response: Combined contraction factor from both sources of variability

Published studies provide only limited information to fully determine the combined contraction factor (calculated from the product of contraction factors imposed by the outcome marker and the mechanical dyssynchrony marker predictor). We therefore provide a theoretical contraction factor based on the data available, in Table 3-5.

Table 3-5 The true limit to the observed R^2 for the correlation between ΔLVEF and dyssynchrony markers is a product of their two R^2 contraction factors. Shaded rows represent the R^2 contraction factor arising from variability in ΔLVEF (range 0.34-0.44). Shaded columns represent the estimated range of R^2 contraction factor arising from variability in the dyssynchrony markers for which data is available (0.35-0.90). The resulting ceiling on observable R^2 values (total shaded region) is modest. Arrows represent values quoted in published studies

R^2 contraction factor imposed by variability in ΔLVEF	R^2 contraction factor imposed by variability in dyysynchrony marker								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0.1	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.2	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18
0.3	0.03	0.06	0.09	0.12	0.15	0.18	0.21	0.24	0.27
0.4	0.04	0.08	0.12	0.16	0.20	0.24	0.28	0.32	0.36
0.5	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
0.6	0.06	0.12	0.18	0.24	0.30	0.36	0.42	0.48	0.54
0.7	0.07	0.14	0.21	0.28	0.35	0.42	0.49	0.56	0.63
0.8	0.08	0.16	0.24	0.32	0.40	0.48	0.56	0.64	0.72
0.9	0.09	0.18	0.27	0.36	0.45	0.54	0.63	0.72	0.81
1	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90

3.6.1 Echocardiographic data from my local hospital database

All 1298 patients had undergone at least 2 scans. The distribution of the first scan across all 1298 patients had mean 60.6 percentage units and SD 19.0. For the second scan, mean was 60.5 and SD 17.5. There was no significant trend in the mean from the first to the second, $p = 0.923$.

Of these, 337 patients had undergone a third scan, as shown in Table 3-6.This subset too, showed no significant difference in means between the first and last scans $p = 0.800$. Likewise, in the progressively smaller subsets of patients who had 4, 5 and 6 scans, there was no significant difference between first and last scans ($p > 0.05$ for all comparisons, Table 3-7).

This dataset showed that the passage of time was not causing a progressive increment or decrement in average EF, neither in the full cohort nor in any of the sub cohorts who had had a larger number of scans.

Table 3-6 The distribution of EF for patients who had 2, 3, 4, 5 or 6 scans performed.

		All patients, 2 scans		Patients with 3 scans		Patient with 4 scans		Patients with 5 scans		Patients with 6 scans	
LVEF (%)	*n*	1298		337		136		45		16	
		mean	SD	**mean**	SD	**mean**	SD	**mean**	SD	**mean**	SD
Scan 1		**60.6**	19.0	**63.6**	16.5	**63.8**	15.7	**64.5**	15.5	**64.0**	17.7
Scan 2		**60.5**	17.5	**63.0**	16.1	**63.1**	16.0	**64.0**	16.0	**66.0**	20.0
Scan 3				**62.2**	19.2	**64.1**	16.2	**65.5**	16.8	**64.6**	19.9
Scan 4						**62.4**	15.3	**64.6**	12.1	**63.9**	13.7
Scan 5								**65.3**	13.5	**67.9**	9.7
Scan 6										**64.6**	16.3

Table 3-7 The scatter (SD) between the first scan and each subsequent scan. This shows that the scatter does not change significantly whether the first scan is compared to the 2nd or to the 6th, hence the passage of time is not a factor. Instead it is the variability of measurement which is causing the wide scatter (SD).

Δ LVEF	All patients, 2 scans 1298		Patients with 3 scans 337		Patient with 4 scans 136		Patients with 5 scans 45		Patients with 6 scans 16	
	mean	**SD**	mean	**SD**	mean	**SD**	mean	**SD**	mean	**SD**
Δ (Scan 1 → Scan 2)	0.0	**18.6**	-0.5	**18.0**	-0.7	**17.0**	-0.5	**16.1**	1.9	**19.0**
Δ (Scan 1 → Scan 3)			-1.4	**20.2**	0.3	**15.6**	1.0	**16.4**	0.5	**20.7**
Δ (Scan 1 → Scan 4)					-1.4	**16.0**	0.1	**16.0**	-0.1	**18.7**
Δ (Scan 1 → Scan 5)							0.9	**18.3**	3.8	**16.4**
Δ (Scan 1 → Scan 6)									0.6	**16.7**

The change in EF, i.e. ΔEF, between immediately successive scans therefore had a value of approximately zero. Of interest was its standard deviation across the patient population, which is an index of reproducibility of the variable. As shown in Table 3-8, this SD of ΔEF had values between 15.7 and 21.1 in all the reasonably sized cohorts of patients. There was no tendency for the SD of ΔEF to become narrower later in any sequence of scans. This supports the concept that this variability is random. I have not calculated the F test p value for thee comparisons of the standard deviations because the interlinking of the successive SD's would break Fisher's original assumption.

Table 3-8 The scatter (SD) of ΔEF between immediately successive scans is shown below

Δ LVEF	All patients, 2 scans 1298		Patients with 3 scans 337		Patient with 4 scans 136		Patients with 5 scans 45		Patients with 6 scans 16	
	mean	**SD**	mean	**SD**	mean	**SD**	mean	**SD**	mean	**SD**
Δ (Scan 1 → Scan 2)	0.0	**18.6**	-0.5	**18.0**	-0.7	**17.0**	-0.5	**16.1**	1.9	**19.0**
Δ (Scan 2 → Scan 3)			-0.9	**21.1**	1.0	**16.9**	1.5	**15.8**	-1.4	**12.9**
Δ (Scan 3 → Scan 4)					-1.7	**16.3**	-0.9	**15.7**	-0.6	**15.3**
Δ (Scan 4 → Scan 5)							0.8	**15.8**	3.9	**12.8**
Δ (Scan 5 → Scan 6)									-3.2	**16.3**

3.7 Discussion

In this study I present a method for calculating the ceiling on the achievable R^2 values between any baseline echocardiographic marker of electromechanical dyssynchrony and subsequent response to biventricular pacing. It may be a surprise how low this ceiling is for both dyssynchrony markers in widespread use, and echocardiographic measures of long term response. This mandatory ceiling applies even if a new mechanical dyssynchrony marker is found that perfectly encapsulates all aspects of dyssynchrony status. The sustainable R^2 will never exceed this ceiling. Even supplanting individual variables with a combination index will not evade this limit because it arises from spontaneous variability inherent in individual patients and in the measurement processes.

Reports of new dyssynchrony markers and combination indexes with strong predictive powers may continue to arise, but this analysis shows with certainty that when tested in independent environments the sustainable R^2 will be small. A different approach to improving our selection for CRT may be required.

3.7.1 The fundamental ceiling on observable R^2

Regression dilution is a well-recognised concept in statistics but less so in medical literature (Hutcheon et al. 2010). With any outcome variable and any predictor variable, the greater the test-retest variability in the predictor and/or the response — whether due to biological factors, measurement error or random noise — the more the observed R^2 between predictor and response must be diluted towards zero except in exceptionally rare chance circumstances. Even if the underlying 'true' relationship (i.e. between imaginary versions of the variables in which the random variability had been stripped away) is a perfect correlation of 1.00, this can never be observed in

practice unless the variability's are vanishingly small (Francis et al. 1999; Hutcheon et al. 2010).

Studies of prediction of response have assumed that the variability of predictors and of outcome measures was small and statistically unimportant. This study demonstrates that this not true, and that therefore our community's expectation of predictability of CRT response from current study designs is set far too high.

3.7.2 Minimising contraction in R^2 values

Critics of the PROSPECT study suggested that poor echocardiographic technique was responsible for poor predictive ability of mechanical dyssynchrony markers. The present study shows that variability is important, but highlights that all sources of test-retest variability conspire to limit predictability of response. Data on test-retest variability between separate echocardiographic sessions are surprisingly sparse, but are fundamental to knowing the maximum credible limit of the ability of a variable (or index constructed from such variables) to predict response. Our data from the echocardiograms analysed, showed results for ejection fraction reproducibility in keeping with findings in 1997 (Otterstad et al. 1997).

A vital step before embarking on a clinical study of the predictive value of a mechanical dyssynchrony marker should be assessment of its test-retest reproducibility in formal, blinded conditions. Both test and re-test acquisitions should be performed on separate days at the very least and preferably with longer intervals between measurements.

Ideally this test-retest reproducibility should be conducted by a group independent of the originators of the marker because even small biases destroy the scientific value of the result. Many candidate indices will fail at this stage, allowing the effort and

expense of trials to be expended fully on a smaller set of indices which have passed basic scientific testing.

It is not sufficient to analyse intra-observer and inter-observer ability to reproduce measurements from the same acquired images (Zhang et al. 2011), because this omits biological and other temporal sources of variability such as subtle differences in (for example) probe position. Such image re-analysis is inexpensive and needs no additional patient effort, and may be a useful early preliminary step to identify a few highly unsuitable modalities where observers cannot agree on interpretation of even identical digital images. Conducting entirely separate scans is necessary to capture the true test-retest reproducibility that is relevant to limiting predictability of response.

The test-retest reproducibility of both the baseline mechanical dyssynchrony markers and the response variables should then be evaluated in the context of Table 3-4. Unless they are much narrower than the population distribution of these variables, the methods should be refined before any major study is initiated. Effort expended on maximising the ratio of signal (between-patient variability) to noise (within-patient variability), is indispensable to improving prediction of response.

Spontaneous variability of measurements across time in my hospital was measured. The data from my centre shows variability no better that that of formal reproducibility. My study is retrospective and the operators concerned were not acquiring data with reproducibility in mind, but in a way this makes it a more favourable representation of clinical practice.

The fact that successive pairs of several EF's showed no greater scatter is not surprising. What is surprising is that the scatter between two echoes many sessions apart was same as those sequentially. This proves that this is not due to evolution of

biological processes which would have taken much longer to take effect but instead due to measurement variability.

3.7.3 Opportunities and limits of replicate averaging

One may not always have the luxury of starting with a better marker. If a marker has substantial noise variability, contraction of R^2 by noise can be minimised by making multiple measurements of each variable and then using an average of these as the patient's final value. However, although noise would be reduced (Pabari et al. 2011), signal would not be increased. Thus sustainable R^2 value will not rise above the underlying value which indicates how much the underlying (between patient) variability in response is genuinely due to underlying (between patient) variability in the marker. For markers with no relationship to response, no degree of replication could make their R^2 value with response high, sustainably. Increased replication can only improve R^2 contraction due to noise.

Simply increasing the number of patients recruited will not raise the ceiling on R^2, rather it will enforce the same mathematical ceiling all the more firmly, since there is less scope for fluke high values.

Often when making multiple replicates for averaging, it is conventional to attempt to make each replicate as similar to each other as possible. However, it is actually preferable for measurements to be made without reference to each other. Thus the random disturbances in the replicates are statistically independent of each other. Fundamentally, cardiovascular variables may fluctuate over periods of hours or days in a manner that is not captured in recordings a few seconds apart but which becomes important when measurements days or weeks apart are compared. This means that if, say, 10 measurements are to be averaged, the average might be more repeatable if the

measurements were done on 10 separate days than on 10 successive beats. This extreme example illustrates the principle that replicate measures which fail to fully sample the intrinsic variability give less helpful averages than those that do. In reality, resources preclude multiple days of measurements, but one should adopt the principle by giving as much opportunity as possible for different noise elements in each of the multiple measurements: using the full range of echocardiographs, sonographers, times of day, prior rest or activity, etc. Equally, intolerably elaborate markers of mechanical dyssynchrony or response that cannot be replicated under the pressures of clinical practice will not be applicable to clinical practice or the generality of patients.

The final advantage of replicate measurement, if conducted "another day, other hands, other eyes", it can expose irreproducibility of some markers allowing their early dismissal such that efforts can be focused on the few with adequate properties.

Whilst readily available and convenient, repeated echocardiography has significant variability. Otterstad et al (1997) showed that repeated echocardiograms a week apart in post infarct patients without change in therapy, had a coefficient of variation of 7-19%. Serial echocardiographic recordings and LV measurements are influenced by small differences in angulation and placement of the transducer. Loading conditions, adrenergic drive and the post-absorption state after a meal also interact. Obesity and COPD, common in the heart failure population also limit image quality. Adjustment needs to be made when considering apparent changes in an echocardiographic variable considering ascertainment error.

This data demonstrates that the variability between 2 echocardiograms, measuring ejection fraction, is between 15% and 20%. This is above the 10% or 15% threshold for detecting response which is the cut-off in many studies (Deplagne et al. 2009).

At least a 10% error should be expected if looking for a one-way change. Thus, setting a 15% change as cut-off for response, this may be too close to change occurring due to error and natural variation. Furthermore, it is known that whilst both very low and high LVEFs can be accurately determined, echocardiography is less accurate in determine LVEFs between 30-50% - the range expected in this patient population (McGowan & Cleland 2003).

Dyssynchronous ventricles are also harder to accurately measure the LVEF. Ghio et al (2009) has suggested that dysynchronous ventricles appear worse than they are – that is the measured LVEF is lower than what is generally being produced. Thus CRT may just allow measurement of this true LVEF, and thus show an apparent improvement.

3.7.4 Why some HSSCSs report higher R^2 than EMRCTs (and higher than mathematically sustainable limits)

Several factors may have contributed to HSSCSs reporting significantly higher R^2 values than the sustainable values found in EMRCTs (Table 3-2).

(a) Chance association: the apparently high R^2 values found may represent statistical chance, which have been noticed and published with preferential enthusiasm. This could occur as submission bias from research groups and/or acceptance bias from journals.

(b) Russian-doll publication: Successive HSSCSs publications from the same site may have overlapping patient cohorts. Patients might understandably be

added to a growing database, from which publications naturally arise. High R^2 occurring by chance in early cohorts would repeatedly contribute to subsequent publications.

(c) Preferential recruitment of patients: Selection of extra patients who have unusually severe or mild mechanical dyssynchrony, or who have unusually large changes in the response variable will have significant impact on the R^2.

(d) Lack of blinding: The R^2 between mechanical dyssynchrony markers and response markers can only reliably inform real-life prospective clinical practice if each measurement is performed by observers blinded to the other relevant measurements in that patient. For example, mechanical dyssynchrony should be measured without knowledge of the LVEF, and vice versa. Concealment of ECG (which shows biventricular pacing spikes) is essential during analysis if unbiased ΔLVEF is sought. The majority of the EMRCTs report some degree of blinding; almost all of the HSSCSs did not (OR 81, 95% CI 6.3-1046.0, $p<0.01$ for response markers; OR 37.5, CI 3.5-399.4, $p<0.01$ for dyssynchrony markers).

(e) Selective inclusion or exclusion of particular patients: HSSCSs may receive unusual referral patterns distorting the distribution of dyssynchrony markers away from the pattern typically seen by future clinical practice and in EMRCTs. Finally, HSSCSs, if done without the advantage of formal, sequentially-numbered, prospective enrolment of patients may end up analysing an incomplete subset of the population at that centre: patients with notably strong concordance between physiological expectation and clinical

response are especially unlikely to be forgotten, but their preferential recollection would persistently bias R^2 upwards.

Table 3-9 Differences in study design between highly skilled single centre studies (HSSCS) and externally monitored randomised control trials (EMRCTs).

Study	EF /LVESV / LVEDV measurements stated to be blinded	Dyssynchrony measurement stated to be blinded	Study measurements only made after formal enrolment
Externally monitored randomised controlled trials			
CONTAK (Marcus 2005)	Yes	Yes	Yes
MIRACLE (Sutton 2003)	Yes	Yes	Yes
CARE-HF (Ghio 2009)	Yes	Yes	Yes
REVERSE (Linde 2009)	Yes	Yes	Yes
MADIT-CRT (Solomon 2010)	Yes	Yes	Yes
MADIT-CRT (Pouleur 2011)	Yes	Yes	Yes
RETHINQ	No	No	Yes
RESPOND	Yes	Yes	Yes
	86%	**86%**	**100%**
Highly skilled single centre studies			
Pitzalis *et al* 2002	No	No	Yes
Bax *et al* 2003	No	Yes	Unknown
Yu *et al* 2003 (AJC)	No	No	Yes
Gorcsan *et al* 2004	No	No	Yes
Notabartolo *et al* 2004	No	No	Yes
Penicka *et al* 2004	Yes	Yes	Unknown
Yu *et al* 2004	No	No	Unknown
Pitzalis *et al* 2005	No	No	Yes
Yu *et al* 2005 (JACC)	No	No	Yes
Mele *et al* 2006	No	No	Unknown
Porciani *et al* 2006	No	No	Unknown
Suffoletto *et al* 2006	No	No	Yes
Yu *et al* 2006	No	No	Yes
Gorcsan *et al* 2007	No	No	Yes
Soliman *et al* 2007	No	No	Unknown
Yu *et al* 2007	No	No	Yes
Delgado *et al* 2008	No	No	Unknown
Jansen *et al* 2008	No	No	Yes
Masan *et al* 2008	No	No	Yes
Van de Veire *et al* 2008	No	No	Unknown
Deplagne *et al* 2009	Yes	Yes	Unknown
Soliman *et al* 2009	No	No	Yes
Park *et al* 2010	No	No	Yes
Bordachar *et al* 2010	No	No	Unknown
Kaufmann *et al* 2010	No	No	Unknown
Norisada *et al* 2010	No	No	Unknown
Van Bommel *et al* 2010	No	No	Unknown
Wang *et al* 2010	No	No	Unknown
Shanks *et al* 2010	No	Yes	Unknown
	7%	**14%**	**≥50%**

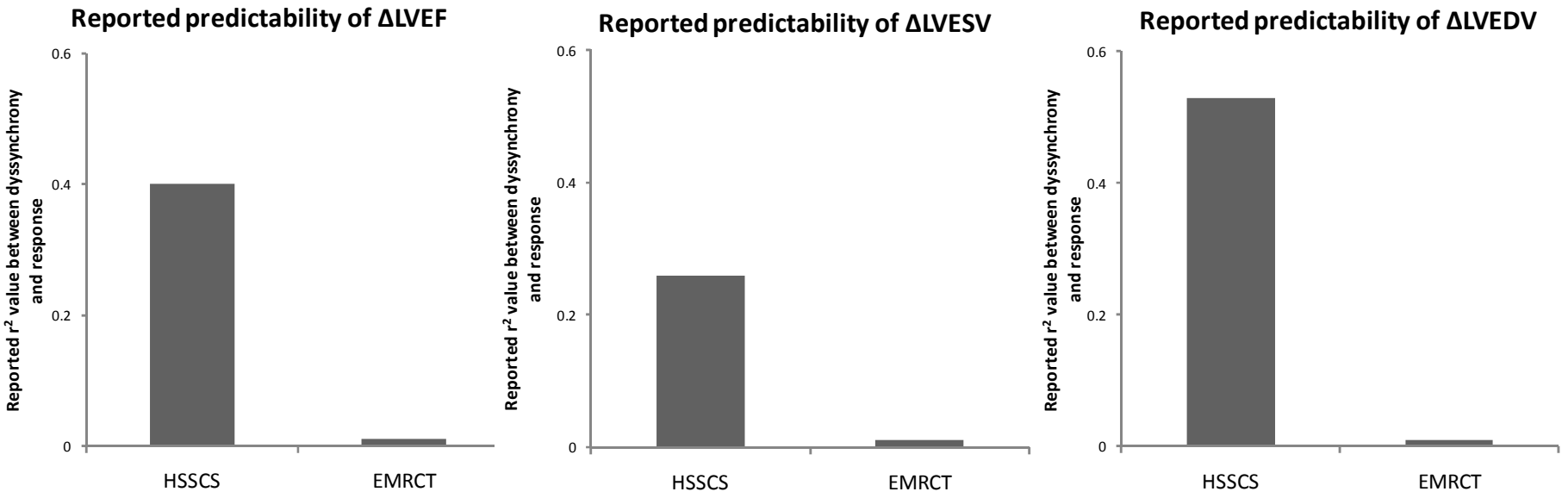


Figure 3-5 Comparison of highly skilled single centre studies (HSSCSs) and externally monitored randomised controlled trials (EMRCTs) for reported R^2 values between baseline dyssynchrony markers and echocardiographic markers of response to biventricular pacing

3.7.5 Clinical Implications

Clinicians are skilled in handling biological variability. It may be difficult for clinicians to believe that variability of a size readily accommodated in clinical practice can have such serious implications for prediction of response. But clinicians appropriately merge technical knowledge with human skill, for example one might enthuse patients who show even a small positive ΔLVEF (saying it is a small improvement, giving them the benefit of placebo) while refraining from dismaying patients who show a small negative ΔLVEF (saying it is essentially unchanged, protecting them from nocebo). Statistics of prediction of response cannot, however, apply this doublethink. Variability imposes a remarkably low ceiling on *sustainable* predictability of response which our discipline's present clinical research strategy is unlikely to dissolve.

The R^2 contraction factor arising from variability in ΔLVEF had a weighted average of 0.48. Even if there was a dyssynchrony marker with a very high R^2 contraction factor from its own variability, the resulting ceiling on the total R^2 contraction factor is at most modest. This means that no matter what one tries to use to predict response to biventricular pacing, the predictive value of that method would be at most modest, and in danger of being almost no clinical value. The predicament would persist for any response marker with wide blinded test-retest variability.

While individual dyssynchrony markers continue to have their own proponents, other workers are seeking combined indices to give better prediction of response. Our study shows that failure of prediction will occur even with a perfect index of propensity to benefit which might include *all* aspects of dyssynchrony and *all* other aspects of clinical status, assembled perfectly into a single ideal score. Failure of prediction is

guaranteed due to spontaneous variability in the outcome measure and test-retest uncertainty in the predictor value itself.

Even if all within-patient variability could be eliminated in a dyssynchrony marker (and in the response variable) this would not automatically make it possible to predict response reliably. Any given marker (or combination predictor variable) may or may not contain the information necessary to predict benefit. This paper does not quantify this particular restraint on prediction because currently the information is not yet available. This paper does show that the only reliable source of such information is carefully-blinded studies such as EMRCTs.

Our current research efforts and resources expended in indentifying strong predictors of response in longitudinal studies, would be better expended first screening predictors (single or conjoint) for test-retest reproducibility and outcome markers to eliminate from consideration those with substantial spontaneous variability with time. If this “first stage” takes forever, then at least we will not have wasted effort on premature reports of strong predictions of response which will be universally found not to stand the test of time.

3.8 Limitations

This study is limited by the nature of published data available for analysis. Not all studies present sufficient information to determine the distribution of the change induced by biventricular pacing. Few studies present test-retest reproducibility of either predictors or outcome measures. I have used rigorously performed EM-RCTs which have control populations not undergoing biventricular pacing to ensure I have accurately determined the inherent variability of response markers over the time periods over which response to implantation is normally calculated.

3.9 Conclusions

This study shows that no externally-monitored trial will have reliable prediction (e.g. R^2>0.5) by any current baseline marker of mechanical dyssynchrony of any current marker of response to biventricular pacing, across a representative range of patients, with measurements carried out in the currently conventional manner. This can be deduced confidently from the inherent variability in these measurements, which mandatorily contract the R^2 value.

It may by these reasons be time to critically reassess the HSSCS literature on echocardiographic prediction of response to mechanical dyssynchrony. The overstatement of relationship strength, of the order of 5 to 20-fold, and indeed the large proportion of the studies showing relationship strengths exceeding the mathematically achievable limits indicates this approach of HSSCSs to identifying predictive markers of response is highly unsatisfactory.

It would be preferable to see a suspension of the natural, but misguided, competition to describe high correlations between baseline markers of mechanical dyssynchrony and response, until two conditions are met:

(1) Techniques for measuring mechanical dyssynchrony become available which have high test-retest reproducibility in the hands of multiple centres beyond their prime developers;

(2) Techniques for measuring "response" are available which consistently show minimal between-individual scatter in patients who undergo no intervention, over the time periods similar to those over which biventricular pacing response is typically measured, and with analyses blinded and externally monitored.

The latter is essentially biologically impossible suggesting the current approach should be set aside. Fascination with, and uncritical acceptance of, near-perfect prediction of individual-patient ΔLVEF from baseline characteristics has evolved uniquely in the environment of cardiac dyssynchrony. Lack of a solid ceiling on what can be claimed, and perhaps perception by authors and journals that the competition is for highest (and not the most reliable) value, may have conspired to yield the present glut of excessive assertions.

4 The Puzzle of Optimisation: When is an optimisation not an optimisation?

4.1 Abstract

4.1.1 Background

Impact of variability in the measured parameter is rarely considered in designing clinical protocols for optimisation of atrioventricular (AV) or interventricular (VV) delay of cardiac resynchronisation therapy (CRT). In this article I approach this question quantitatively using mathematical simulation in which the true optimum is known, and examine practical implications using some real measurements.

4.1.2 Method and Results

I calculated the performance of any optimisation process that selects the pacing setting which maximises an underlying signal, such as flow or pressure, in the presence of overlying random variability (noise).

If signal and noise are of equal size, for a 5-choice optimisation (60,100,140,180,220 ms), replicate AV delay optima are rarely identical but rather scattered with a standard deviation of 45 ms.

This scatter was overwhelmingly determined ($\rho = -0.975$, $p < 0.001$) by Information Content, $\frac{Signal}{Signal + Noise}$, an expression of signal-to-noise ratio. Averaging multiple replicates improves information content.

In real clinical data, at resting heart rate information content is often only 0.2–0.3; elevated pacing rates can raise information content above 0.5. Low information content (e.g. < 0.5) causes gross overestimation of optimisation-induced increment in VTI, high false positive appearance of change in optimum between visits and very wide confidence intervals of individual patient optimum.

4.1.3 Conclusions

AV and VV optimisation by selecting the setting showing maximum cardiac function can only be accurate if information content is high. Simple steps to reduce noise such as averaging multiple replicates, or to increase signal such as increasing heart rate, can improve information content, and therefore viability, of any optimisation process.

4.2 Background

After implantation of a resynchronisation device (biventricular pacemaker or defibrillator) not all patients undergo optimisation even though guidelines recommend that AV and VV delay should be optimised, and even though clinical trials have only demonstrated survival benefit of individually-optimised CRT. Are clinicians right to cut corners from the trial-validated, guideline-mandated process? To answer this, the basic science of optimisation needs to be examined.

For optimisation of atrioventricular (AV) delay, commonly a range of AV settings is tested, while monitoring a marker of cardiac function such as echocardiographic velocity-time integral (Scharf et al. 2005; Thomas et al. 2009; Valzania et al. 2008) VTI, a surrogate of stroke volume (Barold et al. 2008) or left ventricular dP/dt (Gold et al. 2007; Kass et al. 1999) The pacemaker setting that gives the best cardiac function is then defined as the optimum. A similar process can also be carried out for the delay between activation of left and right ventricular leads (VV delay).

However, every measurement has uncertainties, which might conceal the true optimum. This uncertainty in our measurement of VTI (or of any other marker for monitoring cardiac function (De Boeck et al. 2008; Turcott et al. 2010; Zhang et al. 2008) arises from numerous factors including natural biological variability (Turcott et al. 2010). Therefore, repeating the "optimisation protocol" often provides different optima, as shown in Table 4-1.

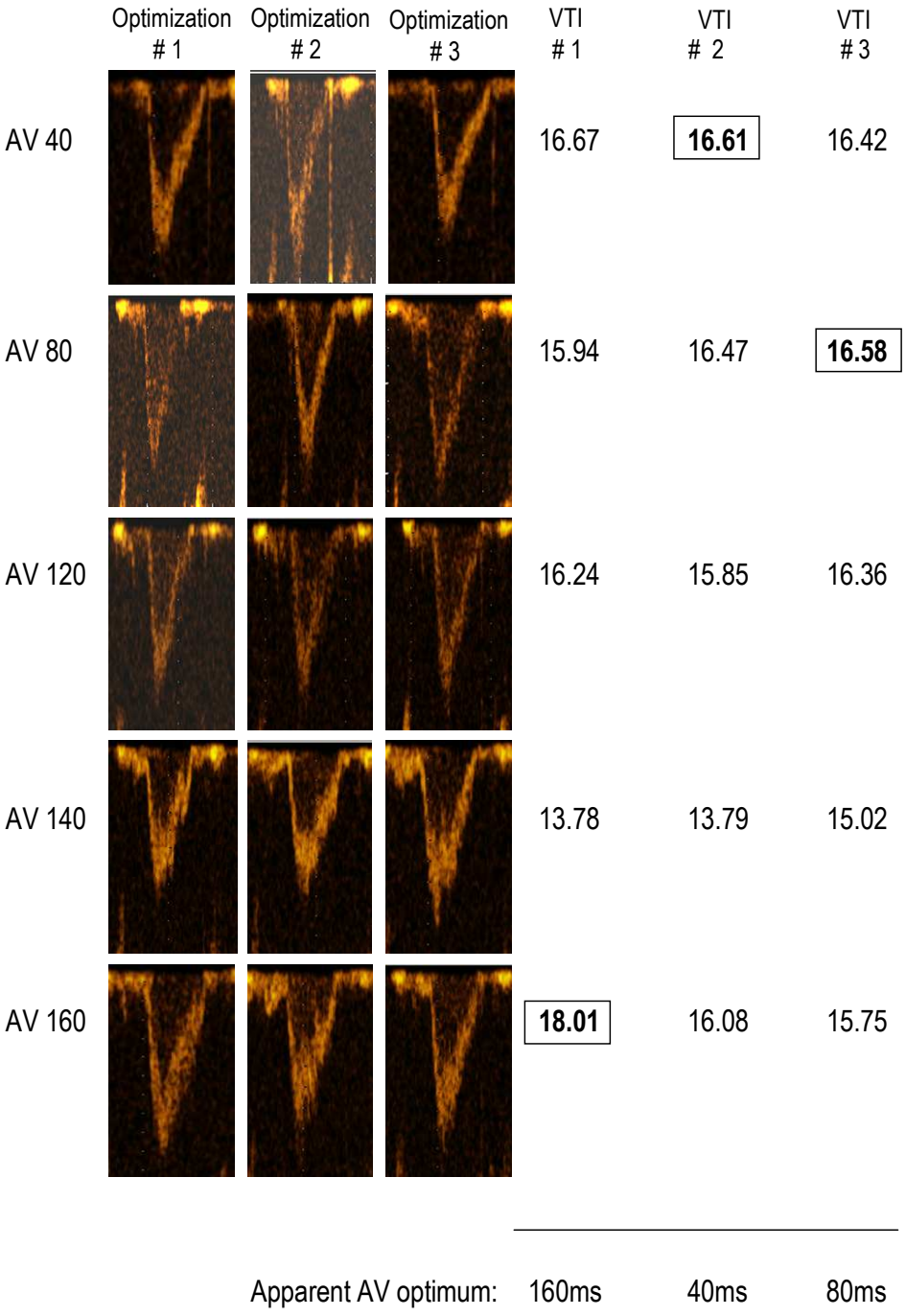


Figure 4-1: An example of clinical data from typical patient undergoing three separate Doppler optimisation processes (#1, #2 and #3) a few minutes apart, using one heartbeat of velocity-time integral as the measurement to be maximised. In this patient, the small differences in velocity-time integral between the three optimisation processes are enough to cause different AV delay settings to be identified as apparently optimal on the three occasions.

There are several clinically important questions. First, if the optimum is not necessarily the 'true' underlying optimum, can we at least express its precision, for example as a 95% confidence interval?

Second, can we trust the measured increase in VTI as a good estimate of the 'true' average underlying increase in VTI?

Third, if optimisations 6 months later show that many patients' optima have changed, would this imply that patients require more frequent re-optimisation (Barold et al. 2008; Porciani et al. 2006a; Zhang et al. 2008)?

Finally, how can the precision of the optimisation protocol be maximised?

It would be difficult and contentious to attempt to answer these questions by doing clinical studies. This is partly because in clinical practice it is normally assumed that the apparent optimum is indeed the true optimum (or at least the nearest of the tested settings to the true optimum). Persons other than the operator conducting the optimisation process itself rarely entertain the possibility that spontaneous variability of the monitored measurement during the optimisation procedure arising from beat-to-beat variability and inherent measurement uncertainty has caused the optimum to be misidentified. Confidence intervals are not reported for individual clinical patients' optima (Anselmino et al. 2009; Gold et al. 2007; Scharf et al. 2005; Valzania et al. 2008).

In this study, therefore, I created mathematical simulation having properties exactly like real-life studies, but in which I could truly know the underlying optimum, despite the presence of overlying noise. To understand the realistic balance between underlying optima and overlying noise we looked at published studies of optimisation.

4.2.1 Information content

A convenient way of quantifying in real-life optimisations the relative contributions of underlying true signal information versus overlying random noise (illustrated in Table 4-2) is using "information content". Signal, in this context, is the genuine underlying between-setting difference in VTI, which for computational convenience can be expressed as a variance (average of the squared deviate between the underlying value of each setting and the mean of all settings). Noise, correspondingly, is the unwanted variability that occurs when measures are repeated at the same setting. This too can be expressed as a variance (average of the squared deviate between individual replicate measurements at a setting and the underlying value of that setting). The advantage of using variances is that their sum is the total observed variance. The variance observed over a series of settings can be decomposed into the variance arising from the genuine between-setting differences (signal magnitude) and the remainder which is noise variance. The proportion of the total variance which is signal, can be called "information content", Equation 4-1.

Equation 4-1

$$Information\ Content = \frac{Signal\ Variance}{Signal\ Variance + Noise\ Variance}$$

The reasons to use information content rather than simply signal-to-noise ratio are three-fold. First, the information content conveniently varies between 0 and 1, rather than extending to infinity. Second, it is symmetrical: noise content is 1 minus information content, which makes it clear that there are two contributors to observed differences between settings. Third, it is numerically identical to the intraclass correlation coefficient, a simple index of reproducibility used in biological research.

4.3 Published data

Information content can be calculated in any study for which both the overall variability and the noise variability are available. I present, in Table 4-1, information content for three detailed physiological studies conducted in research environment where special attention was given to accuracy (Auricchio et al. 1999a; van Geldorp et al. 2011; Whinnett et al. 2006a). For each row of this table, I calculated for each patient the signal size (expressed as a variance) and the noise size (expressed as a variance), and displayed the average values across all patients. Each study had measurements at more than one heart rate, or via more than one monitoring technique, and so had more than one row. Where raw data of multiple replicates was available to us (Whinnett et al. 2006a), noise variance was quantified directly. Where data of only a single replicate was available (van Geldorp et al. 2011), noise variance was defined as the dispersion (expressed as a variance) of raw data away from a best-fit regression parabola between the observed measurements and the AV delay. Where noise variance was published graphically (Auricchio et al. 1999a), it was read off the graph. Signal variance was defined as the total observed variance of that patient minus noise variance. Because the protocols differed between studies, this table should not be used to compare optimisation technologies, but rather just to obtain an idea of the realistic range of information content achievable. It should be remembered that these were conducted in ideal research environments when there was effectively no time pressure. Routine clinical practice, because of time pressure, typically falls short of such ideal protocols that might require as many as 1500 beats to be acquired and analysed (Auricchio et al. 1999b).

Table 4-1 : Information content calculated from published studies of optimisation (Auricchio et al. 1999b; van Geldorp et al. 2011; Whinnett et al. 2006a)**.**

Average heart rate (bpm)	Pacing parameter being optimised	Study	Definition of 1 replicate	Number of replicates (*R*)	Parameter used for optimisation	Number of patients with sufficient raw optimisation data	Signal size, expressed as a variance	The noise size expressed as a variance	Information content of parameter, when used as an *n*-replicate average (mean± sem)
resting	AV	Auricchio et al	Average of first 5 beats at an AV delay setting, minus average of previous 6 beats at reference setting (intrinsic	5	% change in aortic pulse pressure	4 (pts 18, 5, 17 and 13)	10.74 $\%^2$	45.76 $\%^2$	0.20 ± 0.15
resting	AV			5	% change in LV dP/dt	4 (pts 18, 5, 17 and 13)	18.21 $\%^2$	19.44 $\%^2$	0.32 ± 0.05
90	AV	van Geldorp et al	Average of first 8 beats at an AV delay setting, minus average of previous 10 beats at reference setting (intrinsic conduction)	1	Relative change in stroke volume by Nexfin	20	21.78 $\%^2$	8.70 $\%^2$	0.63 ± 0.05
90	AV			1	Relative change in stroke volume by echo	20	40.31 $\%^2$	23.84 $\%^2$	0.71 ± 0.05
90	AV	Whinnett et al	Average of first 10 beats at an AV delay setting, minus average of previous 10 beats at reference setting (AV 120ms)	8	Relative change in systolic blood pressure	10	29.06 $mmHg^2$	3.81 $mmHg^2$	0.51 ± 0.09
110	AV			8	Relative change in systolic blood pressure	10	35.58 $mmHg^2$	2.44 $mmHg^2$	0.86 ± 0.02
130	AV			8	Relative change in systolic blood pressure	10	55.07 $mmHg^2$	4.60 $mmHg^2$	0.90 ± 0.01

In this study I present a simple way to establish the impact of spontaneous beat-to-beat variability, by simulating an optimisation in which there is a known underlying optimum setting at which cardiac function is best, and alternative settings at which cardiac function decays away. In the simulation I can then superimpose random variability simulating clinical beat-to-beat variability (the “noise”). This simulation gives information whose applicability is completely general across any method of optimisation that is based on selecting the settings which gives the most favourable value of a cardiovascular measure.

I aimed to determine:

- how reliable optimisation is
- how one can quantify the confidence interval of any observed optimum
- whether one should trust an apparent increment in cardiac function
- whether the observation that optima change over time is a good reason to increase the frequency of repeat optimisation, and finally
- whether there are any straightforward steps we can take to improve the quality of the optimisation process.

4.4 Methods

4.4.1 Observed measurement = underlying signal + superimposed noise

I constructed a simulation to identify the impact of noise variance, which is the random variability occurring between one beat and another. This noise is superimposed on the signal, which is the "true" underlying effect of the pacemaker setting changes in real patient data. In clinical practice signal and noise cannot be separated in individual raw data points because each such observed measurement contains both contributions mixed together. (However, if replicate measurements are made, their inter-replicate variance can be subtracted from the total variance of the observed raw data to reveal the signal variance).

4.4.2 Simulation

In keeping with real patient data, the underlying signal in this model was constructed as an inverted parabola with its peak – the underlying optimum – at 140 ms. The vertical size of the parabola was scaled to have the desired signal magnitude. The magnitude was defined as the average of the squared deviation from the mean: this definition is computationally identical to that of variance. Separately, we programmed noise as normally-distributed random values with mean zero and variance as desired. The signal and noise were added together to create the simulated observations. This process was repeated separately for each simulated patient. For each analysis in this study, 1000 patients were simulated.

I tested signal and noise sizes over a wide range, but for clarity in this paper I have presented a limited number of values, ensuring that the full spectrum of relative sizes of signal magnitude and noise variance are encompassed.

4.4.3 Identification of optimum

I defined the optimal setting as the one which gave the highest measurement of cardiac function (Barold et al. 2008; Nishimura et al. 1995; Zhang et al. 2008). Because of the presence of noise, the measured value of this optimum may not be the same as the underlying optimum. The measured hemodynamic parameter is not specified, but it could represent VTI (Barold et al. 2008), blood pressure or dP/dt. The measurement is expressed without physical units, for simplicity and generality. Because signal and noise always will have the same unit, the choice of the unit has no impact on reliability of optimisation.

4.4.4 Confidence intervals of the optimum

I simulated repeat optimisations within the same individual, and collected the resulting optima in order to see how widely these optima were scattered. I defined the 95% confidence interval of a single optimisation as 1.96 × the standard deviation of this collection of observed optima. This is the confidence interval that would be appropriate to report for each patient's individual optimisation, although by this method it is of course necessary to carry out several optimisations per patient in order to calculate the confidence interval.

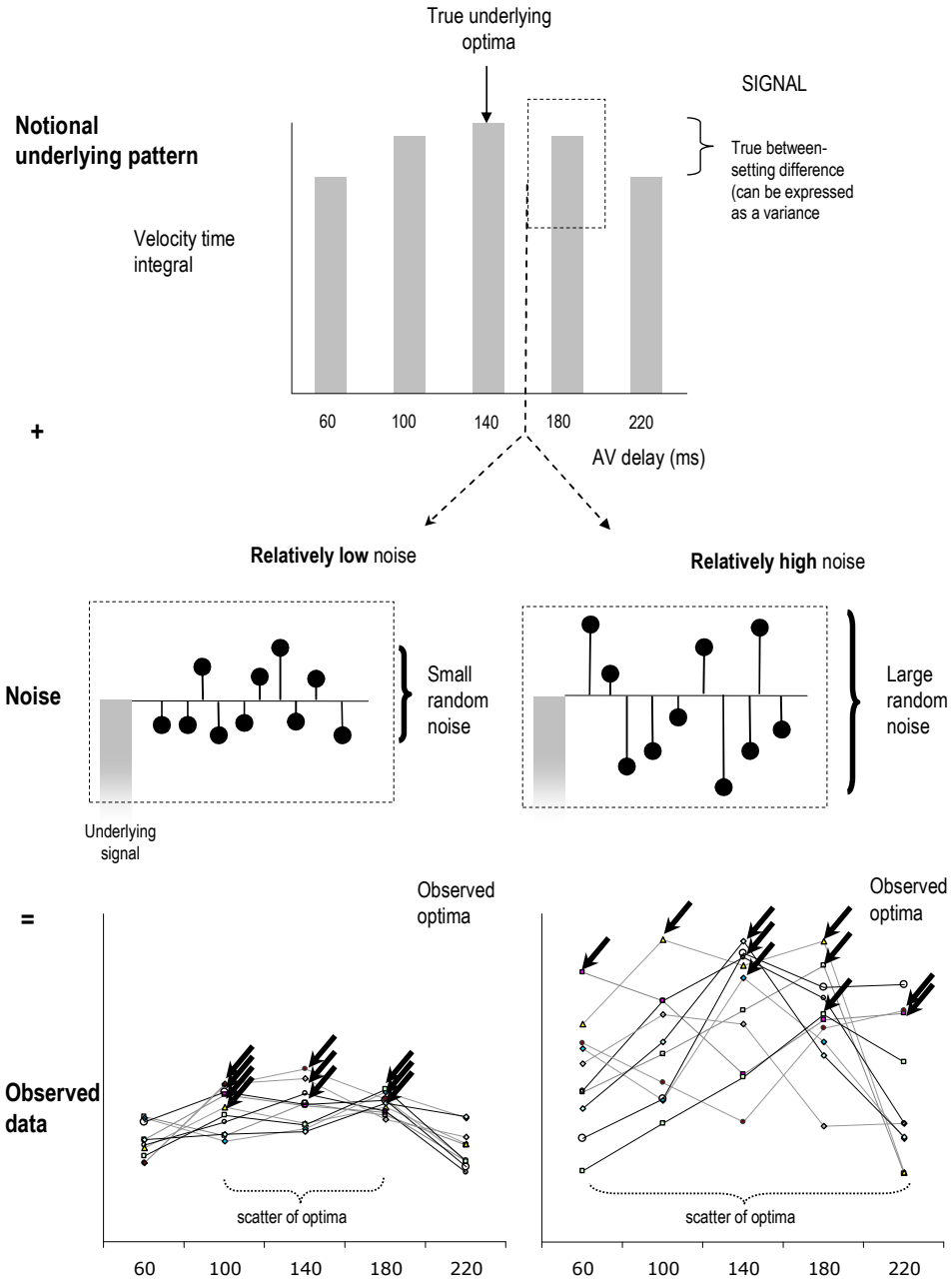


Figure 4-2 Observed measurements are composed of underlying true difference between settings ("signal", top panel) and beat-to-beat variability ("noise", middle panel) which may be small (left) or large (right) relative to the signal.

The relative sizes of underlying signal and overlying noise determine whether the observed measurements (bottom panel) reflect the underlying signal faithfully (left) or not (right). When the noise variability is relatively large (right), the observed optimum (arrow) is often not the true optimum (140 ms in all cases in this figure).

4.5 Results

4.5.1 Impact of information content on consistency of detecting optimum, using a single beat at each setting.

With signal and noise both configured to be the same size, the underlying curved shape of the signal was not always evident in the observed measurements (signal + noise). Nevertheless inevitably, in each run, one of the settings yielded the highest measurement and was duly selected as the observed "optimum". Since this was not always the true underlying optimum, the observed optima showed some scatter (as shown schematically in Figure 4-2).

For each combination of signal and noise size, I quantified the observed scatter of optimisation as the standard deviation of difference between the optima obtained on two successive optimisations of the same patient. I calculated the information content from the known sizes of signal and noise.

When signal and noise were equal, there was an optimisation scatter (standard deviation) of 45 ms. Making the signal magnitude small, made the scatter of the observed optimum wider. Making the signal larger, made the scatter of the observed optimum narrower (Figure 3, Spearman rank correlation coefficient $\rho = 0.973$, $p = 0.021$). When the noise was made smaller, the scatter of the observed optimum narrowed. When the noise was made larger, the scatter of the observed optimum widened (Figure 3, $\rho = 0.991$, $p = 0.0017$).

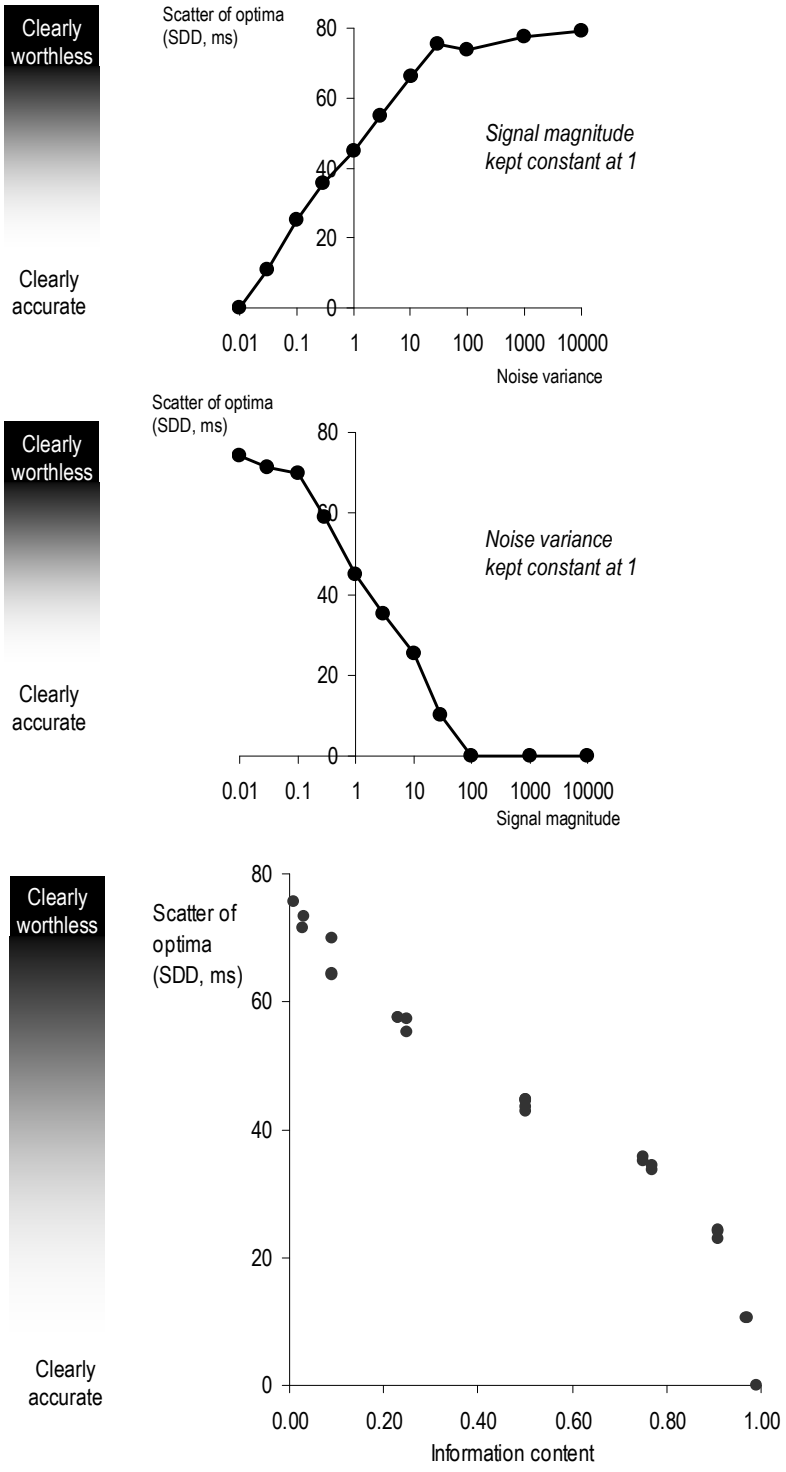


Figure 4-3 Scatter between successive optima increases when noise variance is increased (Top panel) and decreases when signal magnitude is increased (Middle panel). Information content, encompassing the relative sizes of signal and noise, has a powerful effect on the scatter between successive optima. The bottom panel shows the effect of information content (the proportion of variance that arises from signal) on the scatter between successive optima.

The information content was the overwhelming determinant of the scatter of optima ($\rho = 0.979$, $p<0.001$, Figure 4-3). In the worst case scenario, i.e. information content near zero, the scatter of optimisation was ~80 ms, the implied range, 60 – 220 ms, covers the full range of settings over which the simulations are performed.

This can be compared to the expected behaviour of an entirely worthless optimisation method which would be to use no physiological information but simply to select one of the settings (60, 100, 140, 180, 220 ms) at random and announce it to be the optimum. From first principles, the mean "optimum" expected from such an approach is 140 ms, and the expected variance (average square of deviate from that mean) is simply $(80^2+40^2+0^2+40^2+80^2)/5 = 3200$ ms^2, giving an expected optimisation scatter (SD of difference, SDD) of $\sqrt{2} \times \sqrt{3200} = 80$ ms. This forms an effective limit on how poorly reproducible any optimisation amongst these settings can be: SDD can never be more than 80 ms, for this range of tested settings.

Figure 4-3 shows that the information content needs to be rather high before the scatter of optimisation even comes close to values that clinicians may consider acceptable. Even to get the SDD of successive optima down to 25 ms, for example, we need information content of 0.91, i.e. signal-to-noise ratio of 10:1.

4.5.2 Size of confidence interval of the observed optimum

I calculated the size of the confidence interval of the observed optimum for a range of possible signal and noise size combinations (and therefore information content) as shown in Table 4-2.

4.6 Impact of averaging multiple replicates on reproducibility

I tested the impact of changing a clinic's optimisation policy to making, not just a single measurement at each pacemaker setting, but several raw replicates (3, 10, 30 or 100), with average of all those replicate raw measurements in that patient being plotted and used to select that patient's optimum setting. This process improved the fidelity with which the observed measurements reflected the underlying physiological value.

Effectively, the absolute impact of noise was reduced. For example, using averages of 3 replicate raw measurements reduced effective noise variance to one-third (Table 4-2). With this elevation of the signal-to-noise ratio, the shape of the underlying signal was more faithfully depicted in the observed measurements (Figure 4-4, left panels), and the true optimum more likely to be detected (Figure 4-4, right panels).

Table 4-2 Effect of signal and noise on the information content, and on widths of 95% confidence limits.

For simplicity the confidence intervals are shown centred on the "true" value. Greater information content gives narrower confidence intervals. The effect of averaging multiple replicate measurements is to reduce the effective noise and therefore narrow the confidence interval.

Properties of isolated measurements			Effective properties of averaged replicate measurements				
Signal magnitude	Noise variance	Information content	Number of replicate measurements	Effective noise variance	Effective information content	Scatter of optima, SDD (ms)	Size of 95% confidence interval of an observed optimum
1	1	0.50	1	1.00	0.50	45	± 62
1	1	0.50	3	0.33	0.75	37	± 51
1	1	0.50	10	0.10	0.91	24	± 33
1	1	0.50	30	0.03	0.97	12	± 16
1	1	0.50	100	0.01	0.99	0	± 0
1	10	0.09	1	1.00	0.50	66	± 91
1	10	0.09	3	0.33	0.75	58	± 80
1	10	0.09	10	0.10	0.91	42	± 58
1	10	0.09	30	0.03	0.97	35	± 49
1	10	0.09	100	0.01	0.99	26	± 36
1	100	0.01	1	1.00	0.50	74	± 103
1	100	0.01	3	0.33	0.75	74	± 102
1	100	0.01	10	0.10	0.91	68	± 94
1	100	0.01	30	0.03	0.97	57	± 79
1	100	0.01	100	0.01	0.99	43	± 60

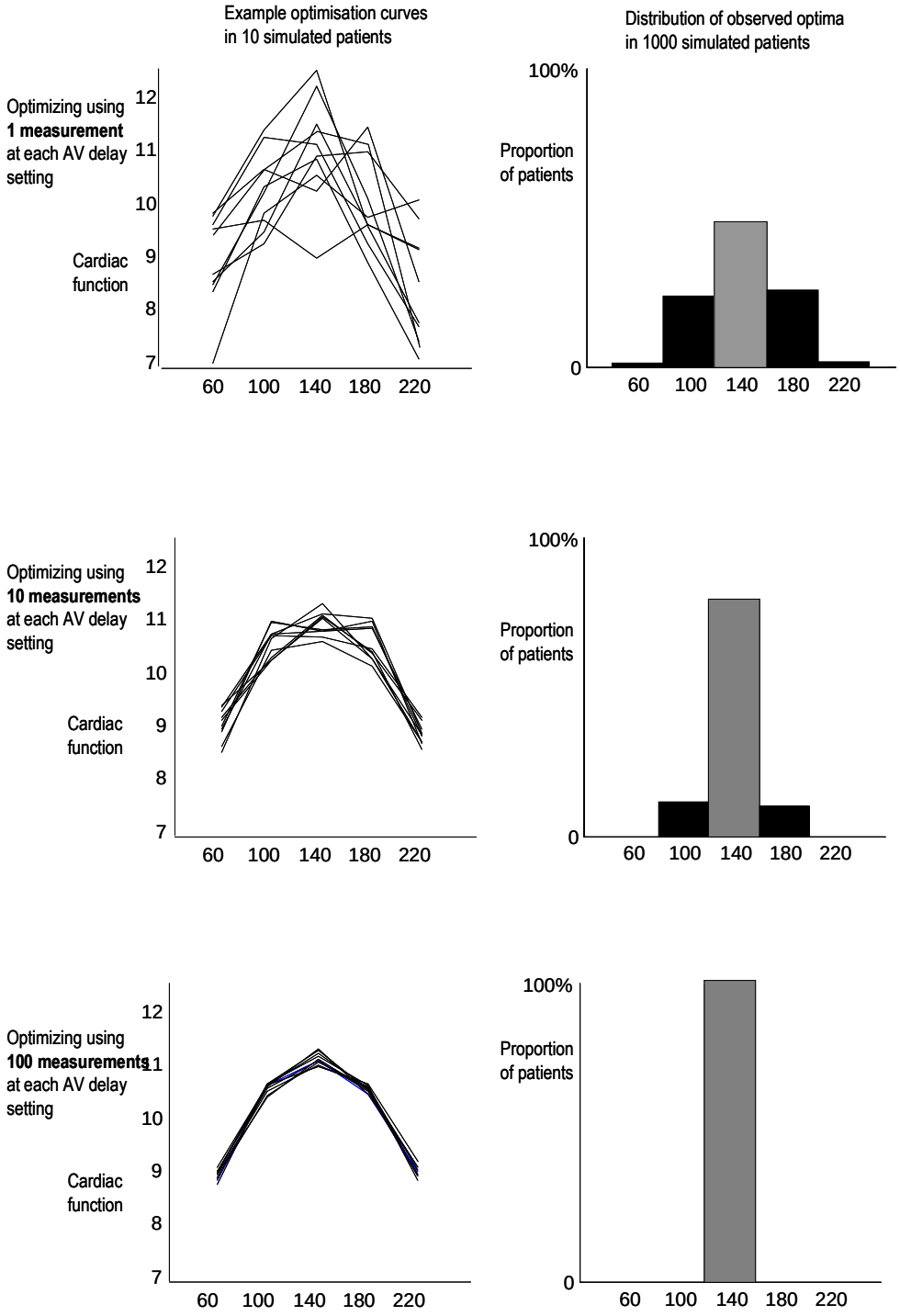


Figure 4-4 Impact of switching from single measurements to average of multiple measurements for the optimisation process.
For simplicity of presentation, the simulated patients all have the same underlying optimum (140 ms) and the signal magnitude and the noise variance for a single measure, is set to be 1. The top panel simulates optimisation with a single measurement made per setting per patient. The lower two panels simulate multiple measurements made for each setting in each patient, each patient's optimum being determined using the averages of that patient's replicate measurements. I display detailed optimisation curves in 10 example patients (left) and the overall distribution of the observed optimum in 1000 patients (right).

4.6.1 Apparent versus true size of improvement on optimisation

I measured the apparent size of the increase in the measured variable upon optimisation. To make the results easy to interpret, I simulated patients to arrive in the optimisation clinic with a reference setting of 100 ms and undergo an optimisation procedure. In each case the underlying true optimum is 140 ms, but because of noise variability, the setting selected as optimum may be this or another setting.

I calculated several aspects. First, the proportion of patients in whom the observed optimum was a correct reflection of the underlying true optimal AV delay.

Second I calculated by how much the observed optimum appeared to be better than the reference state. In reality I also knew how much the underlying optimum was better than the underlying reference state, and I reported this value too, for comparison. This enabled me to report the extent to which the apparent increase over- or under-estimated the underlying benefit. It was always an over-estimation, as shown in the column "Extent of Illusion" in Table 4-3. The size of this illusion was strongly determined by the information content, with lower information contents leading to larger illusory improvements ($\rho = -0.975$, $p < 0.001$).

Third, I calculated the observed difference between the "best" setting and "worst" setting. Because I knew the underlying difference between the true best and worst, I was able to report this too, for comparison. Again I was thereby able to calculate the illusory element.

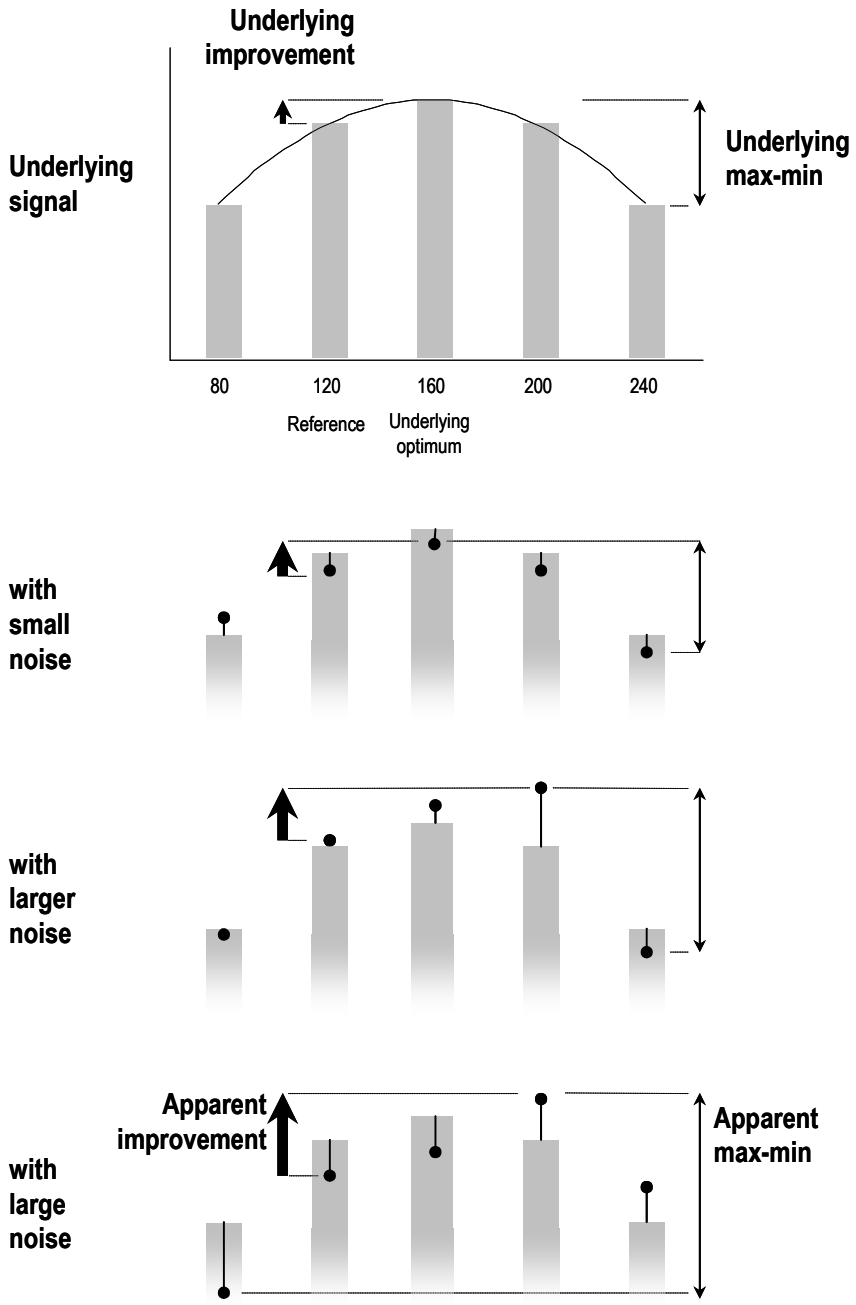


Figure 4-5 Relation between variability and apparent benefit of optimisation.

The presence of noise on its own does not consistently inflate the apparent difference between 2 predetermined settings, since the noise effect is sometimes positive and sometimes negative. However, it can inflate the difference between a predetermined setting and the apparently-best setting. When there is sufficient noise to cause a setting which is not the underlying optimum to appear to be the optimum (because it happens to have had a positive noise element) then we are effectively selecting the setting whose noise is most positive. This introduces a consistent, positive, bias whose size increases as the noise becomes more dominant because the most positive noise element is larger, and because in noisier environments there is a larger group of settings amongst which the apparent optimum might plausibly be drawn. As a result, larger noise consistently inflates the difference between the apparent optimum and any reference setting. Likewise the difference between maximum and minimum is also artefactually inflated by increases in noise.

Table 4-3 Impact of information content on the size of the apparent benefit of optimisation.

For each combination of signal and noise variance, we show the apparent benefit of optimisation (calculated from the measured data including noise) and the true benefit (calculated from the underlying benefit with no noise). The illusory element is also shown, defined as the degree to which the apparent measured increase overstates the true increase.

			Benefit from optimization: optimum minus worst			Benefit from optimization: optimum minus reference		
Signal magnitude	Noise variance	Information content	Apparent measured increase	True underlying increase	**Extent of illusion**	Apparent measured increase	True underlying increase	**Extent of illusion**
1	1	0.50	3.3	2.1	**+ 54%**	1.1	0.5	**+ 105%**
1	10	0.09	7.8	2.1	**+ 263%**	3.2	0.5	**+ 497%**
1	100	0.01	23.7	2.1	**+ 1008%**	11.2	0.5	**+ 2002%**
10	1	0.91	7.5	6.8	**+ 11%**	1.9	1.7	**+ 10%**
10	10	0.50	10.4	6.8	**+ 54%**	3.3	1.7	**+ 98%**
10	100	0.09	24.8	6.8	**+ 267%**	10.5	1.7	**+ 522%**
100	1	0.99	21.9	21.4	**+ 2%**	5.4	5.3	**+ 0%**
100	10	0.91	23.6	21.4	**+ 10%**	5.6	5.3	**+ 6%**
100	100	0.50	32.9	21.4	**+ 54%**	11.2	5.3	**+ 110%**

4.6.2 Apparent change in optimum over time

I simulated repeating the optimisation process after the passage of time, keeping the underlying optimum the same between sessions. I calculated whether the observed optima seemed to change between sessions, and by how much.

For each signal and noise combination, I observed the resulting distribution of differences between the optima found at the 1st and 2nd optimisation visits. Figure 4-6 shows these distributions which have information content of 0.91, 0.50 and 0.09 respectively. Since the true underlying optimum did not change between visits, all changes in observed optimum were false. The proportion of patients of patients giving this false apparent change in optimum is shown by dark shading.

Low information content was strongly linked to the likelihood of false-positive detection of change in optimum ($\rho = 0.975$, $p < 0.001$)

When signal and noise were of equal size (information content ~ 0.50) about two-thirds of the patients have spurious apparent changes in optimum between visits. Even when the signal-to-noise ratio was 10:1, giving an information content of 0.91, still one-third of patients had spurious apparent changes in optimum. Only when signal-to-noise ratio reached 30:1 (information content ~ 0.97), did the proportion of patients getting false-positive apparent change in optima fall to a clinically respectable 7% (top panel, Figure 4-6).

Information content	Proportion of patients whose optima **falsely** appeared to have changed	Proportion of patienmts whose optima are **correctly** identified as unchanged
0.99	0%	100%
0.97	7%	93%
0.91	31%	69%
0.75	56%	44%
0.50	65%	35%
0.25	75%	25%
0.23	73%	27%
0.09	78%	22%
0.03	80%	20%
0.01	79%	21%

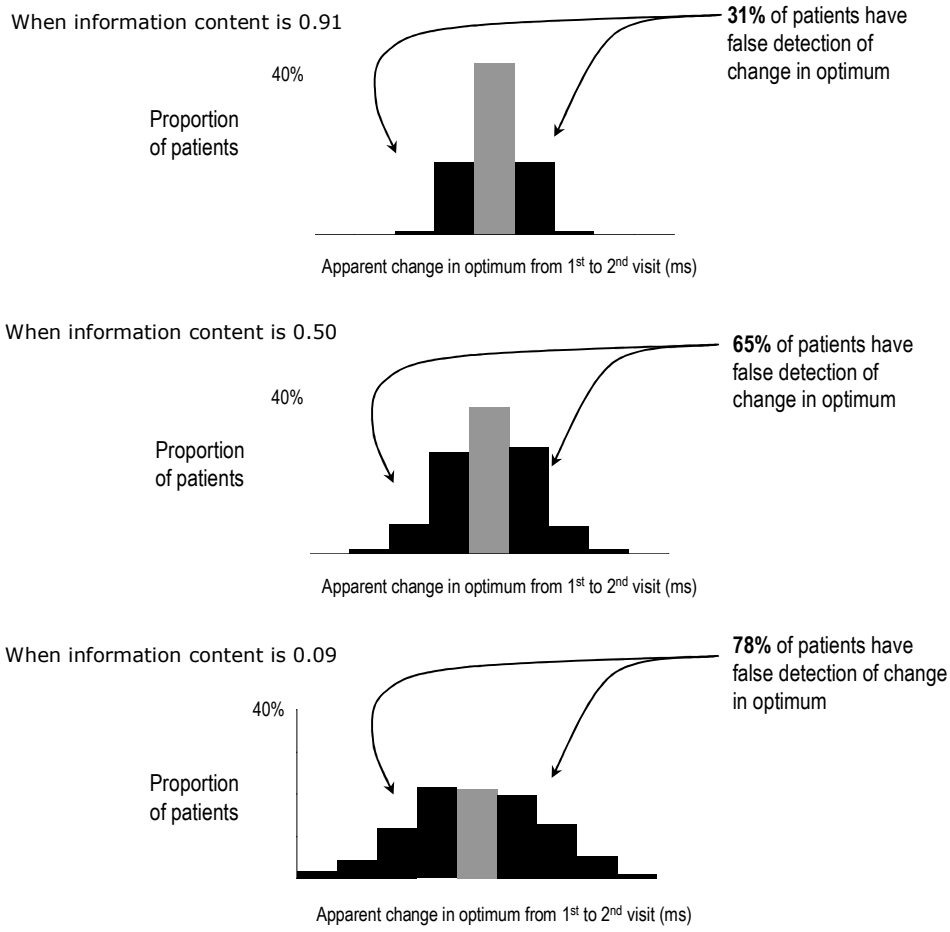


Figure 4-6 Impact of information content on the probability of falsely detecting a change in optimum.

A group of patients is simulated attending the optimisation clinic twice, with their underlying optima truly unchanged between visits. We calculate the apparent change in optimum between visit 1 and visit 2. The table shows the percentage who have a spurious apparent change depending on the information content. Graphically we can see that when information content is high (top panel) only 33% of patients have a spurious apparent change in optimum. When information content is slower (middle and bottom panels), the proportion of patient having a spurious apparent change in optimum becomes much higher. It should be noted that simply randomly choosing between 5 settings, gives an 80% (4/5) rate of spurious detection of change in optimum, which is no worse than the bottom panel.

4.7 Discussion

In this study I have shown that uncritically selecting the pacemaker setting which gives the best value of a monitored variable might be little better than random selection amongst a set of AV settings. These findings are generally applicable to any optimisation method that relies on testing a series of settings while monitoring some measure of cardiac function (such as echocardiographic velocity-time integral or pressure or any other cardiovascular marker) and then picking the setting that gives the highest measurement.

It is overwhelmingly important for signal-to-noise ratio (information content) to be high, otherwise a series of illusions automatically arise in any clinical data analysis.

4.7.1 Illusion 1: "We have selected the true underlying optima"

One tends to assume that the setting which gives the highest measurement is the best. However, our study shows that only a very small amount of variability is enough to seriously compromise this assumption because the true biological effect may also be very small. With signal and noise of equal size for example, in ~ 50% of cases (Figure 4-6) the optimum detected will not be the true optimum but an erroneous alternative.

The confidence interval of a clinical optimisation is never reported and (surprisingly) rarely asked for. A wide confidence interval will have immediate comprehensibility to any clinician reviewing the result. The simplest way to calculate the confidence interval of optimisation is to carry it out on several occasions (e.g. immediately, one after the other) and calculate the standard deviation. The 95% CI would be the mean ± 1.96×standard deviation. To make this reasonably valid, one would need to perform at least 3 or 4 optimisations. Of course this would be extremely time-consuming, and is therefore not realistic for routine clinical practice with current monitoring techniques.

Alternatively one can determine information content of the clinic's optimisation process in general. This could be calculated once and then applied to all similar patients without having to carry out multiple replicate optimisations in each new clinical patient. Fortunately, information content is easy to calculate: it is essentially the intraclass correlation coefficient. This can be calculated quickly for a representative group of patients by any laboratory. This is similar, in principle, to using concepts of statistical power analysis to routine clinical practice.

4.7.2 Illusion 2: "The optimisation increased flow (or pressure) by X and was therefore worthwhile"

It is tempting to average the apparent increments in velocity-time integral (or whatever measure was used for optimisation) achieved in an optimisation service, and believe that, first, (a) the process is almost always increasing stroke volume, (b) the size of the average increase in stroke volume is 'X' which sounds clinically worthwhile and (c) that since the increment in statistically significant it is not likely to be a chance finding.

This study reveals all three of these tempting conclusions to be wrong. First, the setting selected as apparently optimum will always have a higher measured cardiac function than the reference setting (except where the reference setting happens to be selected as the optimum). Even if an optimisation method was just roulette amongst n tested settings, then in (n-1)/n cases (i.e. almost always) it would be selecting an optimum different from reference. Therefore, the statement that stroke volume is higher on the optimal setting is meaningless.

Second, ironically, the worse the optimisation method, the larger the illusionary increase in stroke volume.

Third, unless carefully constructed (Turcott et al. 2010), the statistical test is assessing whether or not changes in stroke volume are randomly distributed (some positive, some negative) with a mean of zero. But each patient's increment will always be either positive or zero (never negative), so the average increment will always be statistically significantly positive unless the sample size is very small. Indeed, the worse the optimisation method, the more likely the apparent increment is to be statistically significantly positive.

4.7.3 Illusion 3: "The optimum has changed between X months and now"

A well established and indispensable optimisation clinic may start to consider how often these optimisations should be carried out (Porciani et al. 2006a; Zhang et al. 2008). Is the contrast between patients' optima on subsequent visits a useful guide? My analysis now shows that if a technology has poor information content (low signal-to-noise ratio) reproducibility will be poor. For example when signal and noise are approximately equal (information content = 0.5) at 6 months (or any other time) the optimum will falsely appear to have changed, purely through noise, 65% of the time (Figure 4-6). Ironically, the worse the optimisation process, the more the data will seem to encourage more frequent optimisations. The giveaway clue to this would be that however frequently we re-optimise, there would still be a similar proportion who would seem to need a change in setting.

4.7.4 Illusion 4: "We should not waste time making multiple replicate measurements at each setting in clinical practice"

In a busy clinical department, it may seem an unnecessary multiplication of work to make more than one measurement at each setting. Instead, it may seem rational to

concentrate on ensuring that each measurement is acquired and analysed properly by well-trained staff. Unfortunately, the reasons for beat-to-beat variability in measurement are many, and inadequate skill on the part of the sonographer or interpreter is typically not the dominant contribution. Rather, there is substantial beat-to-beat variability in transvalvular blood flow, ventricular volumes, arterial blood pressure and dP/dt. These variations may be due to respiration and numerous other less-easily monitored physiological processes that take place over periods of seconds and minutes. They will not disappear through wishful thinking alone. Instead, averaging multiple replicate acquisitions gives us a powerful method to reduce the effective noise. Effective noise (the variance of the averaged value from R replicate raw measurements) falls in direct proportion to 1/R, providing a simple way to improve the information content. Another strategy for blood pressure recordings, is to elevate heart rate, since this increases the size of the signal (Whinnett et al. 2006a).

4.7.5 Illusion 5: "We should optimise using whatever measurement method we are most familiar with"

Inter- and intra-observer variability may not be the dominant source of noise, rather there may well be genuine biological variation between beats. Even with excellent clinical acquisition and measurement technique, if the biological variability is large in comparison to the true signal between settings, information content will be low. One should quantify information content directly and not assume that the technique with which we are most familiar has a high information content.

4.7.6 Illusion 6: "Between separate beats, variability in my laboratory is only X%, therefore this measure is suitable for use in optimisation"

That X%, being the ratio between variability and mean measurement, is not the relevant ratio for quality of optimisation. Reliability of optimisation depends on the ratio between beat-to-beat variability (noise) and between-setting variability (signal). The ratio is much less favourable than X%. For example, a VTI measurement might have a mean value of 10 cm, and a standard deviation of 1 cm, giving a coefficient variance of 10%. However, the relevant signal is not 10 cm but the standard deviation between-settings which may only be (for example) 1 cm, In this case, the information content would be $\frac{1}{1+1}$ = 0.5. The naive figure of 10% variability, in isolation, is of no relevance.

4.8 A simple method of calculating information content of a cardiovascular measure used clinically for optimisation

Because this study was carried out using computer simulation, it was possible to know the size of the true underlying signal, as well as the size of the noise, and thereby state the information content directly.

In vivo, one can calculate information content by measuring total variance and noise variance, since although the underlying signal magnitude cannot be directly observed, it is the difference between them. We need to carry out several optimisations in the same patients. Suppose one carries out R replicate sets of optimisations in one patient. First calculate the variance of all the raw measurements (V_{raw}). Then one can calculate the mean measurement at each pacemaker setting, and then the variance (V_m) of these means. V_m will tend to be smaller than V_{raw}, because the impact of noise is reduced by the averaging process. The lower the information content in the measurement, the larger its noise in comparison to its signal, and therefore the more markedly V_m will differ from V_{raw}. In brief, the information content is approximately the ratio V_m / V_{raw}, when R is large. More elaborately, accommodating for R not always being large (Equation 4-2).

Equation 4-2

$$\textbf{Information content} = \frac{R}{R-1}\frac{V_m}{V_{raw}} - \frac{1}{R-1}$$

An example of how to calculate information content in a single patient, using only standard spreadsheet software, is shown in Figure 4-7 Calculation of information content using raw clinical data from a single patient..

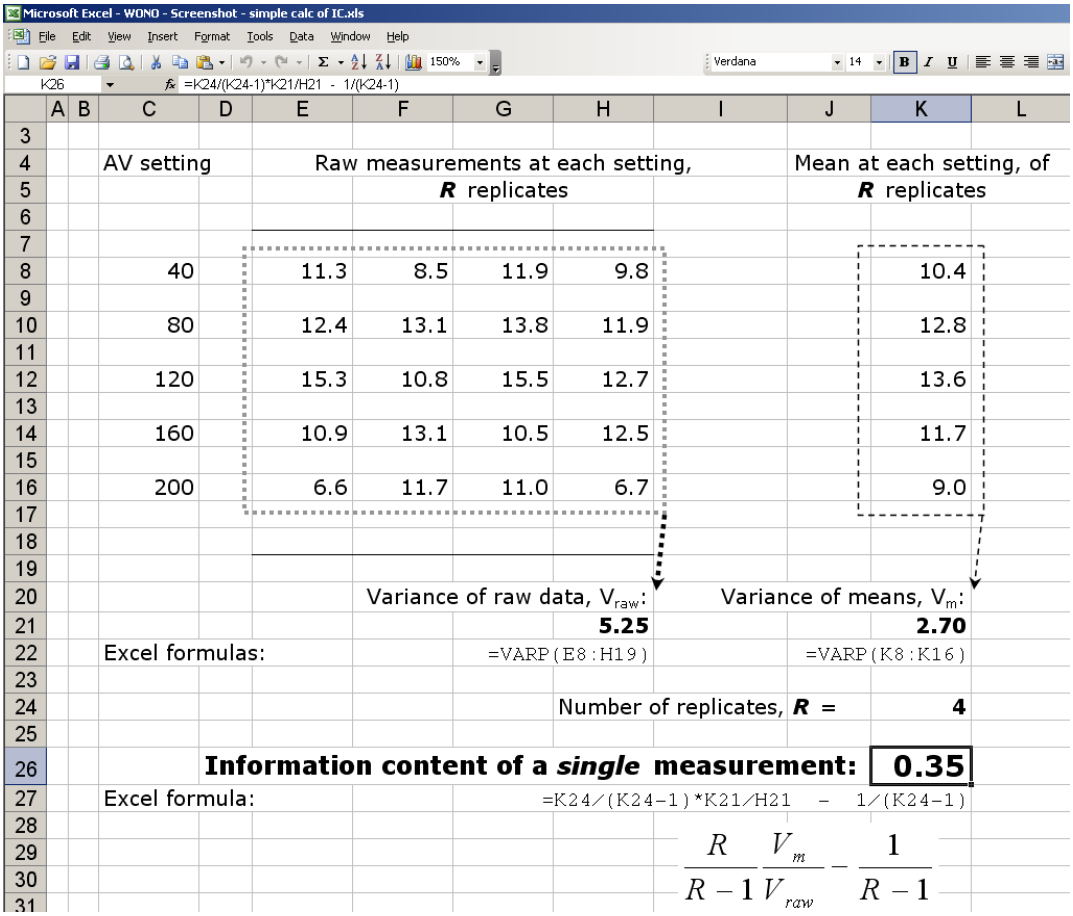


Figure 4-7 Calculation of information content using raw clinical data from a single patient.

In this example of real-life data from one patient, 4 replicates of measurements at 5 settings are entered into a table (columns E to H), and the mean at each setting calculated (column K). The variance of the raw data is calculated (Cell H21) using the formula shown immediately below it; the same is done for the variance of the means (Cell K21). Information content is calculated in cell K26 using the formula shown immediately below it. The formulae shown are in appropriate form for standard spreadsheet software such as OpenOffice or Microsoft Excel. In particular, information content for several early patients (who would need to undergo replicate measurements) can be averaged to allow the laboratory to calculate typical confidence intervals to be reported alongside optimal settings in future patients.

In practice, the examples of published data on information content in Table 4-1 show that even with time-consuming methodology including a high number of replicates and many beats measured per replicate, information content can still be low.

4.9 How many replicates are really needed in clinical practice?

Clinicians cannot afford to waste time in clinical practice on performing unnecessary numerous measurements during optimisation. Nor, though, can they waste time performing apparent optimisations that they should know will be worthless before the patient even lies down on the couch. To choose rationally the number of replicates to perform it is vital to decide how precisely the patient's optima is to be identified.

In clinical practice each individual physician can decide what level of precision is suitable in their context and can easily calculate the number of replicates required to achieve this as long as the information content of a single replicate of their local method is known. The numbers of replicates required for a range of such combinations is shown in Table 4-4.

For example, a clinician may wish to know the AV optimum with a 95% confidence interval of ± 10 ms. How many replicates are needed depends on the heart rate at which optimisation is to be carried out (Table 4-1). Studies at resting heart rate have found rather low information contents around 0.3.

If a confidence interval of optimisation of ±10 ms is wanted in this context, from Table 4-4 it can be seen that the number of replicates needing to be conducted at each setting is 59.

Table 4-4 The number of replicates required when optimisation is performed to reduce the scatter of AV optima obtained to a range of acceptable confidence intervals.

	ICC of single replicate	For 95% confidence interval of an individual-patient optimum to be:				
		±5ms	±10ms	±15ms	±20ms	±30ms
		Required "effective ICC" of the multiple replicates				
		0.974	0.962	0.93	0.877	0.56
Projected **behaviour of VV optimisation algorithms which average 18 - 20 beats per single replicate**		Number of replicates required to achieve this				
	0.02	1836	**1240**	651	349	62
	0.05	712	**481**	252	135	24
	0.1	337	**228**	120	64	11
Actual **behaviour of AV optimisation algorithms which average 18 - 20 beats per single replicate (Table 1)**	0.2	150	**101**	53	29	5
	0.3	87	**59**	31	17	3
	0.4	56	**38**	20	11	2
	0.5	37	**25**	13	7	1
	0.6	25	**17**	9	5	1
	0.7	16	**11**	6	3	1

At higher heart rates such as 90 bpm, information content is approximately 0.5 to 0.7 for several methods (Table 4-1). Achieving a confidence interval of ± 10 ms now only requires 11-25 replicates, as shown in Table 4-4, which might be achievable. At higher heart rates still, the number of replicates needed continues to fall as information content increases (Figure 4-8).

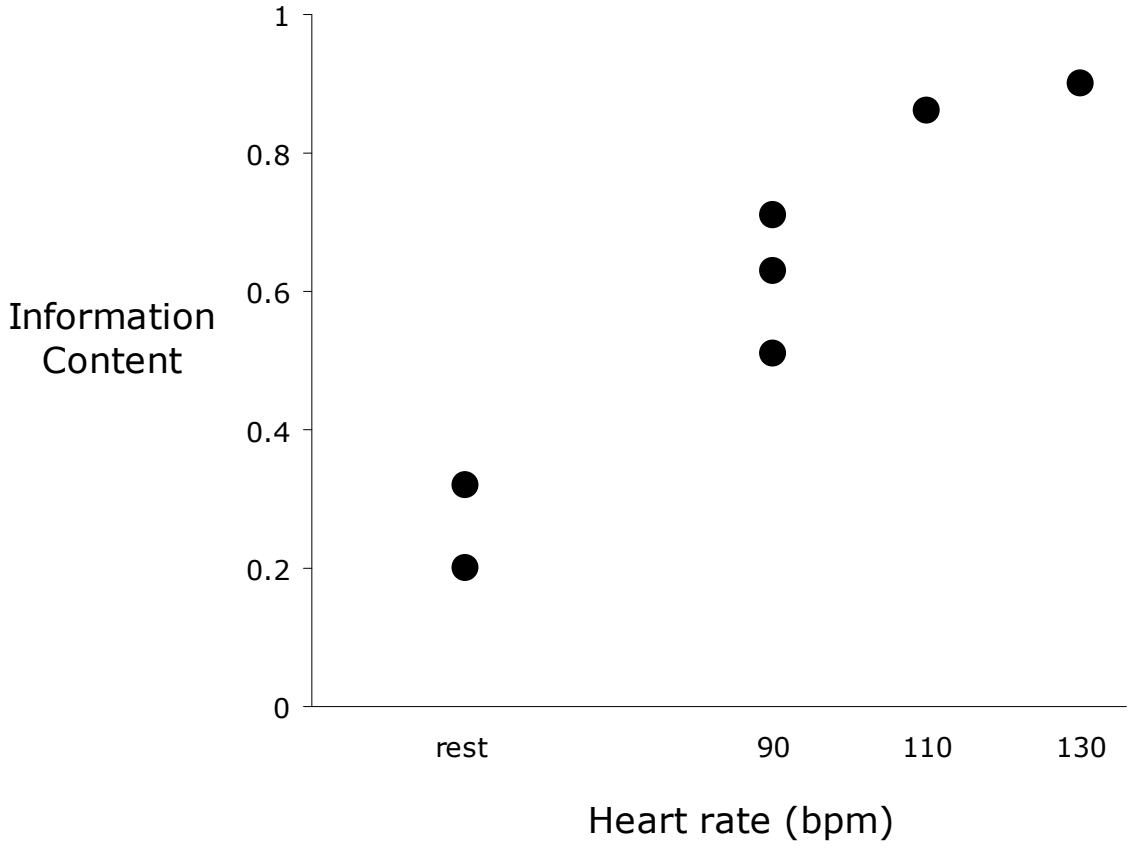


Figure 4-8 The effect of heart rate on information content from datasets of published studies regardless of method used for AV optimisation

Adjustment of VV delay, in contrast, exerts a much smaller signal effect on physiological measurements than AV adjustment; by a factor of about 5-7 fold (Whinnett et al. 2006b). Even if the variation in blood pressure is just 5-fold smaller, the information content is roughly 25-fold smaller – because it is the variances (squared deviates) that matter. Therefore, even assuming a favourably elevated (90 bpm) heart rate, a favourable range of AV optimisation information contents of 0.5-0.7, and a possible relative signal variance for VV of $(1/7)^2$ to $(1/5)^2$, the information content for VV would lie between 0.01 and 0.03. It can be seen from Table 4-4 that this necessitates well over 500 replicates at each setting to achieve the desired precision of optimisation.

Although there are detailed descriptions of meticulous protocols (Auricchio et al. 1999b) even putting together 1500 beats of data does not give high information content. Multi beat averages reduce noise, but if signal is small, information content may still be small. High heart rate raises signal magnitude (Whinnett et al. 2006a) and has allowed a higher information content to be obtained.

4.10 Clinical Implications

No clinical optimisation protocol currently specifies a number of replicates to be carried out, while giving a quantitative reasoning. This may be because the impact of noise has not been considered or measured. It may not be rational to conduct an optimisation without ensuring adequate precision of the optimum. Although there may be a clinical imperative to be seen to be doing something, we should not necessarily give in to perceived pressure to conduct a placebo procedure. Worse still, if the apparent optimisation is in fact no different to randomisation amongst a constrained range, it is inescapable that half of all such procedures worsen cardiac efficiency rather than improve it.

If we want our optimisation service to be delivering clinical valuable results, there are three generic steps we should take. First, it is important to have as large an underlying signal as possible. For blood pressure changes, it has been reported that the signal is larger in absolute terms at higher heart rates than at lower heart rates (Whinnett et al. 2006a).

Second, noise should be as small as possible. We should not criticise operators for inadequate care when they may simply be correctly measuring biological variability. Instead the measured variable and protocol should be designed to have a high information content.

Third, we can take averages of multiple replicate measurements of cardiac function at each AV delay setting. An R-fold replication will have the same beneficial effect as reducing the noise variance of individual measurements by R-fold. This can be applied to any measurement technique, but of course carries the cost of increased labour.

Realisation of these inherent properties of optimisation, should encourage us all to mandatorily report the noise and information content of our monitored variable in our hands. We should be able to therefore present the confidence interval with every optimisation we carry out. This may be uncomfortable.

I emphasize that in this article I am not recommending one method of measurement (e.g. VTI or pressure) over another, nor suggesting whether measurement should be invasive or non-invasive. The choice of measurement modality for optimisation should be prejudged by personal preference or based on whim, but rather selected on the balance of relevant properties. The most important property of optimisation (a process that recommends small adjustments to pacemaker settings) is the precision with which the recommendation is given. This chapter is neutral and simply provides a language to rationally evaluate, discuss and improve this precision.

4.10.1 Practical Recommendations

This analysis is completely general to all optimisation schemes which test a range of settings and select the one with the greatest measurement. Any laboratory conducting optimisation can use Equation 4-2 and Figure 4-7 to calculate their typical information content. In concert with device physicians, who can recommend an acceptable confidence interval, the laboratory can see how many replicates are required.

Such an estimated number of replicates required only applies to an "average" patient in the population. The size of the signal may vary between patients. For example one patient may have a particularly critical dependency on AV setting, and another a below-average amount of dependency. The former would need fewer replicates to identify the optimum within a given size of confidence interval, and the later would need more. Similarly, one patient may have more noise for one of many reasons, including deeper respiration due to acute physiological distress; chronic lung disease that enhances ventilatory fluctuation in haemodynamics; obesity impairing image quality; agitation impairing probe position maintenance. This would necessitate more replicates.

But while individual patients may have different strict needs for replication, all patients will need more replicates if the optimisation technique has poor information content. Any protocol document (which specifies an optimisation technique) to be credible must at least give quantatively sound guidance as to the number of replicates needed for an average patient to obtain optimisation with a level of precision widely considered reasonable. If a protocol does not give such guidance, clinical time pressures may lead to all patients having optimisations that are, on average, worthless (helping half slightly, and harming half slightly).

4.11 Conclusions

Information content, the proportion of the observed differences in the measurements at different settings that is genuinely due to the change in settings, has an overwhelmingly important impact on the meaningfulness of the pacemaker optimisation processes. Although easy to measure, it is rarely reported or commented on, and may be surprisingly low unless steps are taken to improve it.

Low information content leads to frequent misidentification of the optimum. However, worse than this, it inflates the apparent benefit of optimisation: counter-intuitively, the worse the optimisation method, the better it will superficially appear (unless one asks about information content).

Worst of all, because low information content makes apparent optima more variable, the poorer the optimisation method, the more frequently one will feel compelled to re-optimise the patient (unless we ask about information content).

Information content is easy to improve for any technique. All that is needed is (a) to use a technique where the underlying difference between settings is as large as possible, (b) to use a technique with beat-to-beat variability as small as possible and (c) to make multiple measurements at each setting and calculate the average.

If, despite these steps, information content is still low, clinical resources could be saved by selecting a setting arbitrarily or even at random, with no additional loss to the patient's physiology. We do not make this suggestion for fun but to point out the seriousness of the present situation. Optimisation is not optimisation when it is roulette.

5 Validating markers of mechanical dyssynchrony by experimental manipulation of interventricular timings: What is needed to make them a reasonable prospect for Cardiac Resynchronisation Therapy selection?

5.1 Abstract

5.1.1 Background

In optimisation of CRT (and even selection for implantation) we may be underestimating the impact of beat-to-beat variation in echocardiographic measurements.

A simple screening test of a proposed dyssynchrony marker is that its value should be stable between heartbeats but change dramatically when interventricular delay is experimentally manipulated across a wide range; another is that an optimal interventricular delay should minimise dyssynchrony.

The optimisation process exposes this most clearly, because genuine small changes in cardiac function (signal) must be detected amongst potentially large beat-to-beat variation (noise).

5.1.2 Methods and Results

In this study of biological variability, I performed 2592 echocardigraphic measurements in 13 patients with CRT. I performed separate, replicate measurements at a series of interventricular delays using several potential optimisation modalities. These were (i) 3D systolic dyssynchrony index, (ii) Tissue Doppler imaging, (iii) aortic pre-ejection time, (iv) interventricular mechanical delay, (v) LVOT VTI and (vi) QRS duration. Variability in acquisition or between observers over time is eliminated.

Agreement between successive optimisations was low with Kappa values of 0.24 for SDI, 0.02 for TDI, 0.36 for aortic pre-ejection time, 0.14 for IVMD, 0.40 for LVOT VTI and 0.47 for QRS duration

The intraclass correlation coefficient (ICC), which quantifies measurement reproducibility, was low for all methods when single measurements were taken (ranging from 0.32 to 0.63), but improved when pairs of measurements were averaged (0.51 to 0.74, p=0.0008). The scatter between replicate optima obtained is lower when pairs of measurements were taken, p=0.007 across all methods.

5.1.3 Conclusions

The relative sensitivity of these dyssynchrony markers to detect even large changes in VV delay (as distinct from spontaneous beat-to-beat variability) appears to be very poor. Under these blinded conditions these mechanical dyssynchrony markers cannot reliably discriminate even large changes in interventricular delay, and can be quickly rejected as candidates for predicting clinical benefit from CRT.

VV delay optimisation by any of the echocardiographic techniques is not realistic unless multiple replicates are performed and averaged. More concerningly, even dyssynchrony assessment to select patients for implantation may need averaging of far more replicate measurements than is currently described, to have any hope of success.

It would save time and expense if markers considered for clinical trialling under formal scientific conditions first underwent screening for plausibility by such a stage of inexpensive, active experimentation.

5.2 Background

What properties should a marker of dyssynchrony have? First, it should be stable from one beat to the next, which can readily be tested. Second, it should change when the amount of dyssynchrony changes: this can be tested by changing the VV delays settings of a Cardiac Resynchronisation Therapy (CRT) pacemaker. For a marker of dyssynchrony to have any hope of being diagnostically useful, the impact of changes in VV delay setting must be much greater than the impact of random beat-to-beat biological variability. This "signal-to-noise" information should be helpful in evaluating potential markers of dyssynchrony (Pabari et al. 2011), and may be pivotal for designing approaches to optimisation of VV delay.

There are several echocardiographic markers which might be used to detect dyssynchrony or be used to optimise the atrioventricular or interventricular delay settings of a CRT device once implanted. Potential techniques for VV optimisation include maximisation of LVOT VTI (Mortensen et al. 2004; Thomas et al. 2009), minimisation of 3D dyssynchrony (Kapetanakis et al. 2005; Liodakis et al. 2009) or minimisation of the differences in tissue Doppler timings (Yu et al. 2004b). The latter techniques are of special interest because they are dyssynchrony markers also used for selecting patients for implantation.

In this study I carry out an acid test of markers of dyssynchrony. I eliminate between-patient variability, between-observer variability and gradual changes in clinical status. With all these kept constant, I experimentally manipulate dyssynchrony through a wide range of values to confirm or refute the ability of individual dyssynchrony markers to detect this.

I have systematically quantified the discriminant power of these echocardiographic indices in a formal head-to-head evaluation to reproducibly detect changes in mechanical dyssynchrony elicited by changes in VV delay. Specifically, I distinguish the genuine changes caused by change in VV delay from the natural random biological variation occurring between beats. To do this I measure several independent beats, at each of several VV delays, in each patient.

5.3 Methods

5.3.1 Study Population

13 outpatients with biventricular pacemakers or biventricular defibrillators previously implanted for clinical indications (Barnett et al. 2007) were invited to enter this detailed study of approximately 10 hours of echocardiographic recording per patient. All patients were free from decompensation for greater than 3 months and no changes to medication were made 4 weeks prior to entering the study and until all data had been collected. All patients had intrinsic sinus rhythm, heart rate 70 ± 17bpm, with 100% biventricular pacing (AsVp) and a fixed AV delay programmed to the standard nominal setting of 120ms.

At the time of the study, 1 patient was NYHA class I, 4 were NYHA II and 8 were NYHA III. Time between implantation and enrolment was 32 months (range 8 months to 63 months. 5 subjects were male and 8 female, age 69 years ± 8 years (mean ± sd). The cause of heart failure was ischemic in 5 subjects and idiopathic dilated in 8 subjects. 11 patients were taking angiotensin-converting enzyme inhibitors or angiotensin-II receptor antagonists, 10 were taking beta-blockers, 6 were taking spironolactone and 6 were taking a diuretic (loop or thiazide). Patients gave informed consent for this study which was approved by the local ethical committee.

5.4 Protocol

In each patient optimisation of VV delay was performed using each of the potential markers: (i) 3D systolic dyssynchrony index (SDI), (ii) Tissue Doppler imaging (TDI), (iii) aortic pre-ejection time, (iv) interventricular mechanical delay (IVMD), (v) LVOT VTI and (vi) QRS duration . The range of VV delays investigated was RV first 40ms, 20ms, VV 0ms, LV first 20ms, 40ms and 60ms. and replicate measurements were performed as described below. Continuous ECG monitoring from limb leads confirmed that all patients were 100% AsBiVp throughout the measurements.

For most patients this required 3 or 4 visits of several hours each, which were carried out within 2 weeks of each other. All the measurements using a particular modality (e.g. tissue Doppler) across all replicates and all pacemaker settings were conducted within one visit by one experienced operator for each patient. This ensured that the only variability being quantified was beat-to-beat biological variability, a variation that cannot be attributed to between-operator disagreements or to a substantial change in patient status. 2592 measurements were made in total and I analysed all results.

5.5 Echocardiography

Echocardiography was performed in all patients using the Philips IE33 machine able to acquire the advanced images for 3 dimensional and tissue Doppler echocardiography, in addition to the more common practice Doppler measurements. Tissue Doppler studies and pulsed wave (PW) Doppler images were acquired using the Philips S5-1 sector ultrasound transducer. 3D echocardiography was performed using a X3-1 transducer.

5.5.1 Modalities used for optimising VV delay

5.5.1.1 LVOT – Velocity Time Integral

LVOT VTI images were taken from the apical 5-chamber view with the pulse wave Doppler cursor at the left ventricular outflow tract approximately 1 cm below the aortic annulus. Eight beats were acquired at each VV delay, by a single operator, at end expiration. Offline analysis was manually traced and resulted in digitised VTI tracings. The optimal VV delay setting was defined as that at which the LVOT VTI was greatest in magnitude.

5.5.1.2 Ejection flow dyssynchrony

Ejection flow dyssynchrony was measured in 2 ways. Firstly aortic pre-ejection time (APET) by measuring the time delay from onset of QRS to onset of flow (Lane et al. 2004). Second we calculated the interventricular mechanical delay (IVMD) as the delay between RV ejection and LV ejection using pulsed wave Doppler across the LVOT and RVOT respectively. Eight beats were acquired at each VV delay. All acquisitions were made at end-expiration to minimise respiratory variation. For both APET and IVMD methods, the optimum VV setting was defined as that which made the delay the smallest.

5.5.1.3 Tissue Doppler imaging

Colour coded tissue Doppler imaging was used for acquisition and analysis, using the 2 segment model from the onset of QRS complex to the peak of the systolic wave form (Bax et al. 2003a; van Bommel et al. 2010b). Apical images of the left ventricle including the level of the mitral annulus were acquired with a minimum frame rate of 90 Hz (Gorcsan et al. 2004b). Images were acquired with LV cavity positioned in the centre of the sector showing clear myocardial definition with the patient holding their

breath at end-expiration. A minimum of 4 cycles were acquired, allowing 4 separate analyses. Time–velocity curves were obtained using a sample volume placed within the segments at the region of interest at the basal lateral and septal (Bax et al. 2003a; Chung et al. 2008a) segments. Analysis was performed offline using Philips Qlab 8.1 with cardiac motion analysis (CMQ) software.

The time-to-peak systolic velocity was measured for each wall from the onset of the QRS to the highest peak on the velocity curve within the aortic ejection period. If two peaks were seen, we used the timing of the highest peak, and if two peaks of equal velocity were seen, the timing of the earliest peak was used. The optimal VV setting was that at which the delay between septal and lateral peak S' velocities was smallest.

5.5.1.4 3D Echocardiography: Systolic dyssynchrony index

3D echocardiography was performed using a X3-1 transducer. Four-chamber apical views of the heart were acquired with the left ventricle clearly identified. Full-volume datasets were acquired; each recorded over 4 consecutive beats. Acquisitions had a minimum frame rate of 20 frame/s in keeping with ASE guidelines, with the patient breath holding at end expiration. For each VV, delay four replicate measurements were taken.

Offline analysis was performed using Tomtec analysis software, version 2.0 (Tomtec GmBH, Germany). The three views used for primary analysis are the 4 chamber view, 3 chamber and 2 chamber simultaneously at 60 degree views. End-systolic and end-diastolic frames were identified, once the ventricle was aligned to eliminate foreshortening. Automatic border detection was then carried out and corrected as appropriate. A segmented shell is created (Figure 5-1) and time-volume curves and the systolic dyssynchrony index were calculated (Gimenes et al. 2008), with the

analysis performed by myself. The optimal VV setting was that at which the systolic dyssynchrony index was smallest.

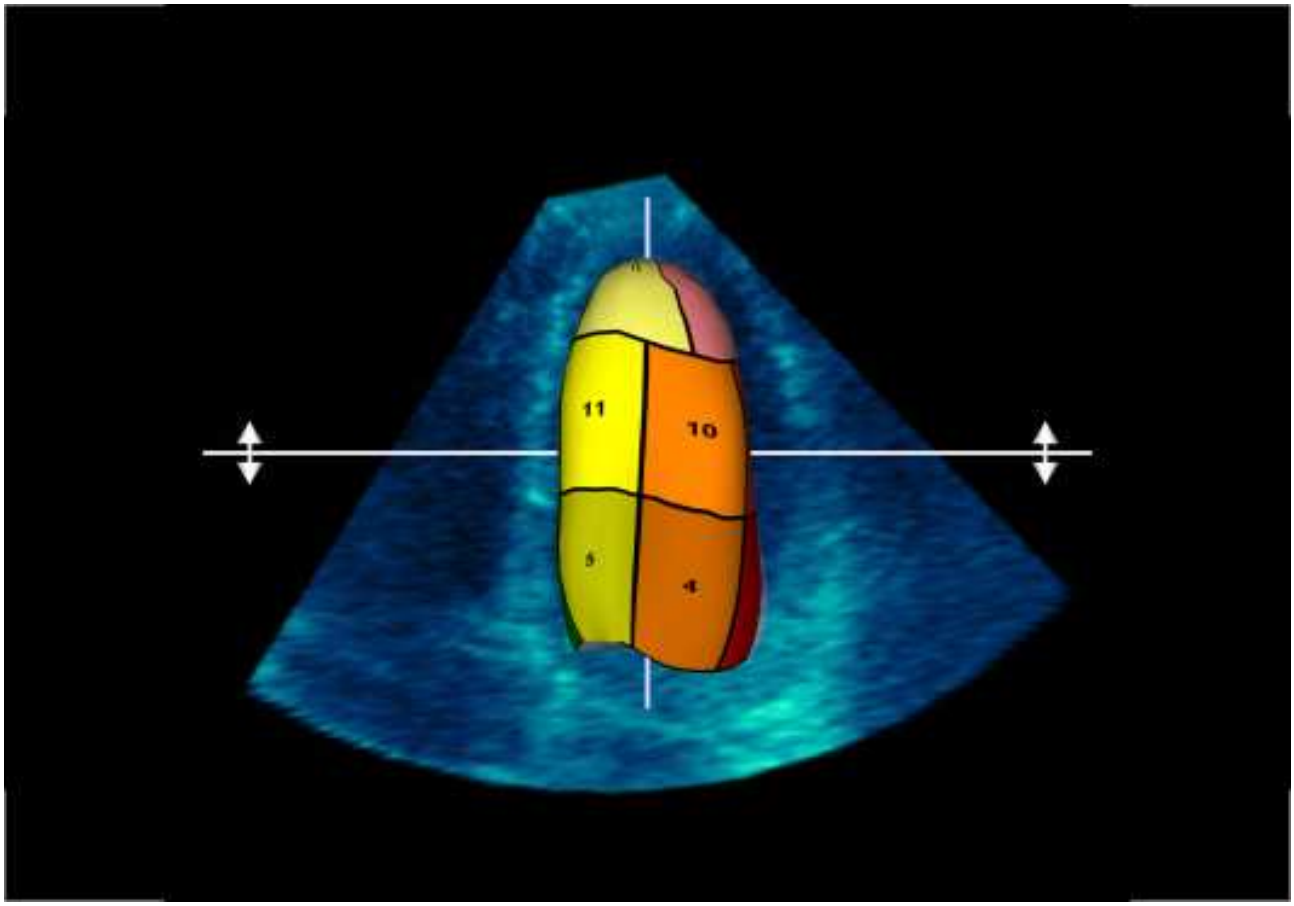


Figure 5-1 A shell is created with all segments numbered.
We can see segments 4,5,10 and 11 in the example above. Each of the 16 segments have the time to minimal volume measured.

5.5.2 Calculating the agreement of selecting the optimal VV delay setting within the same modality

For each of the dyssynchrony markers, I made multiple replicate acquisitions at each of the tested VV delays thus mimicking multiple optimisations and allowing me to measure the extent to which variation between measurements at different VV delays were genuinely due to changes in VV delay, versus simply natural biological variation. For aortic pre-ejection time, interventricular mechanical delay and LVOT VTI I made 8 replicate measurements. For 3D SDI and TDI, which are more time consuming, I made 4 replicate measurements.

The agreement is quantified as the number of pairwise comparisons between replicate optimisations, by the same method, that yield the same optima, Figure 5-2. I calculated the number of times the same optima was obtained when repeat optimisation processes were performed within the same modality, and displayed this as a percentage of the number of pairwise comparisons made.

5.5.3 Calculating Kappa within the same modality

The Kappa statistic gives the excess proportion of agreement. above the agreement that would arise by chance alone, for each individual modality studied. I initially calculated the distribution of 'optimal' settings across the replicate measurements for each modality, $P_{setting}$. Using this value, I then calculated the expected level of agreement by chance, P_{chance}, for that modality, as follows:

Equation 5-1

$$P_{chance} = \sum \left(P_{setting}{}^{2} \right)$$

I then calculated the raw agreement for that modality, defined as the number of occasions when two replicates gave the same optimum in the same patient, divided by the number of comparisons made. Expressed formally, this is:

Equation 5-2

$$P_{raw} = \frac{\text{Number of agreements within patients}}{N_{patients} \times N_{replicates} \frac{N_{replicates} - 1}{2}}$$

Finally I obtained Fleiss' Kappa in the usual way:

Equation 5-3

$$\kappa = \frac{P_{raw} - P_{chance}}{1 - P_{chance}}$$

5.6 Calculating the scatter of detected optima between repeated optimisations for each separate modality

Using the individual 'optimum' VV delay settings for each of the replicate optimisations for a specific modality, we calculated their scatter, i.e. the standard deviation between the replicate optima, and compared these values of 'scatter' between the 6 separate modalities.

5.7 Quantification of the genuine effect of interventricular delay versus random variability

The intraclass correlation coefficient (ICC) was used to quantify how well changes in synchrony were detected in amongst background noise and spontaneous variability. This was calculated for each modality across all patients. As previously described, an ICC close to 1 indicates that changes in settings make a relatively big change in the measurement parameter in comparison to the random beat-to-beat variability. An ICC close to 0 indicates that the random beat-to-beat differences are relatively big in comparison to the genuine effect of changing settings. ICC was defined as the variance of the means of the values from each of the tested VV delays, divided by the variance of all the values within that patient for that particular modality.

5.8 Results

5.8.1 Assessment of each of the 6 dyssynchrony markers as candidates for VV optimisation using single measurements

First I present the distributions of optimisation measures for single measurements. This mimics a simple, rapid clinical optimisation process in which each VV setting is applied in turn with a single measurement made for each.

As described, for each of the dyssynchrony markers, I made not one but multiple replicate acquisitions at each of the tested VV with 8 replicate measurements for aortic pre-ejection time, interventricular mechanical delay and LVOT VTI, and 4 replicate measurements for 3D SDI and TDI.

5.8.1.1 Probability of selecting the same optimum on replicate optimisations

Within method agreement of optima could not be described as good for any of the modalities. 3D systolic dyssynchrony index had 38% agreement (Figure 5-2), tissue Doppler imaging 21%, aortic pre-ejection time 49%, interventicular mechanical delay 31%, LVOT VTI 54% and QRS width 58%.

The kappa values for the within modality agreement are 0.24 for 3D systolic dyssynchrony index, 0.02 for tissue Doppler imaging, 0.36 for aortic pre-ejection time, 0.14 for interventricular mechanical delay, 0.40 for LVOT VTI and 0.47 for QRS duration. These low values indicate the poor agreement within each dyssynchrony marker and itself on successive beats.

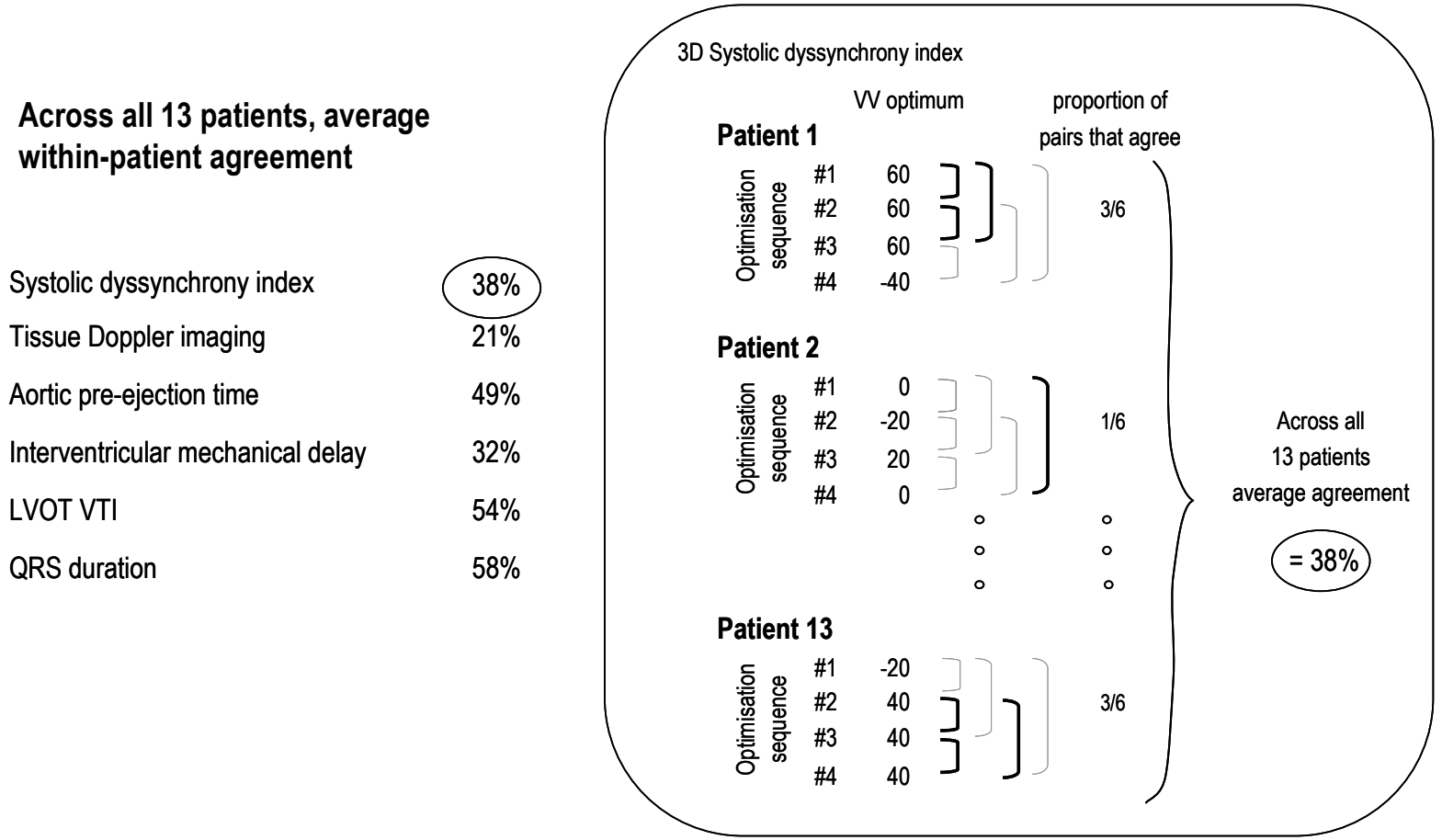


Figure 5-2 Agreement between replicate VV optimisations by the same method. Inset shows method of calculating this value for 3D SDI.

Similar calculations were done for the other methods. The agreement is quantified as the number of pairwise comparisons between replicate optimisations, by the same method, that yield the same optima.

5.8.1.2 Probability of two different optimisation methods agreeing on an optimum

I then examined the agreement, between one method and another, on what the optimum VV setting was. I used each of the single replicates in turn. The agreement between methods was poor, within many cases the level of agreement being not much different from pure chance.

Figure 5-3 demonstrates how the agreement was calculated and shows both the raw agreements and kappa valued between methods.

Agreement between modaities for selecting VV optima.

Pair of dyssynchrony markers	Raw agreement on optimum	Kappa
TDI vs APET	13%	0.15
TDI vs IVMD	11%	0.17
TDI vs LVOT VTI	11%	0.14
TDI vs QRS duration	21%	0.16
TDI vs SDI	16%	0.17
APET vs IVMD	25%	0.18
APET vs LVOT VTI	28%	0.19
APET vs QRS duration	21%	0.18
APET vs SDI	14%	0.16
IVMD vs LVOT VTI	13%	0.17
IVMD vs QRS duration	15%	0.19
IVMD vs SDI	31%	0.16
LVOT VTI vs QRS duration	16%	0.15
LVOT VTI vs SDI	7%	0.15
QRS duration vs SDI	11%	0.16

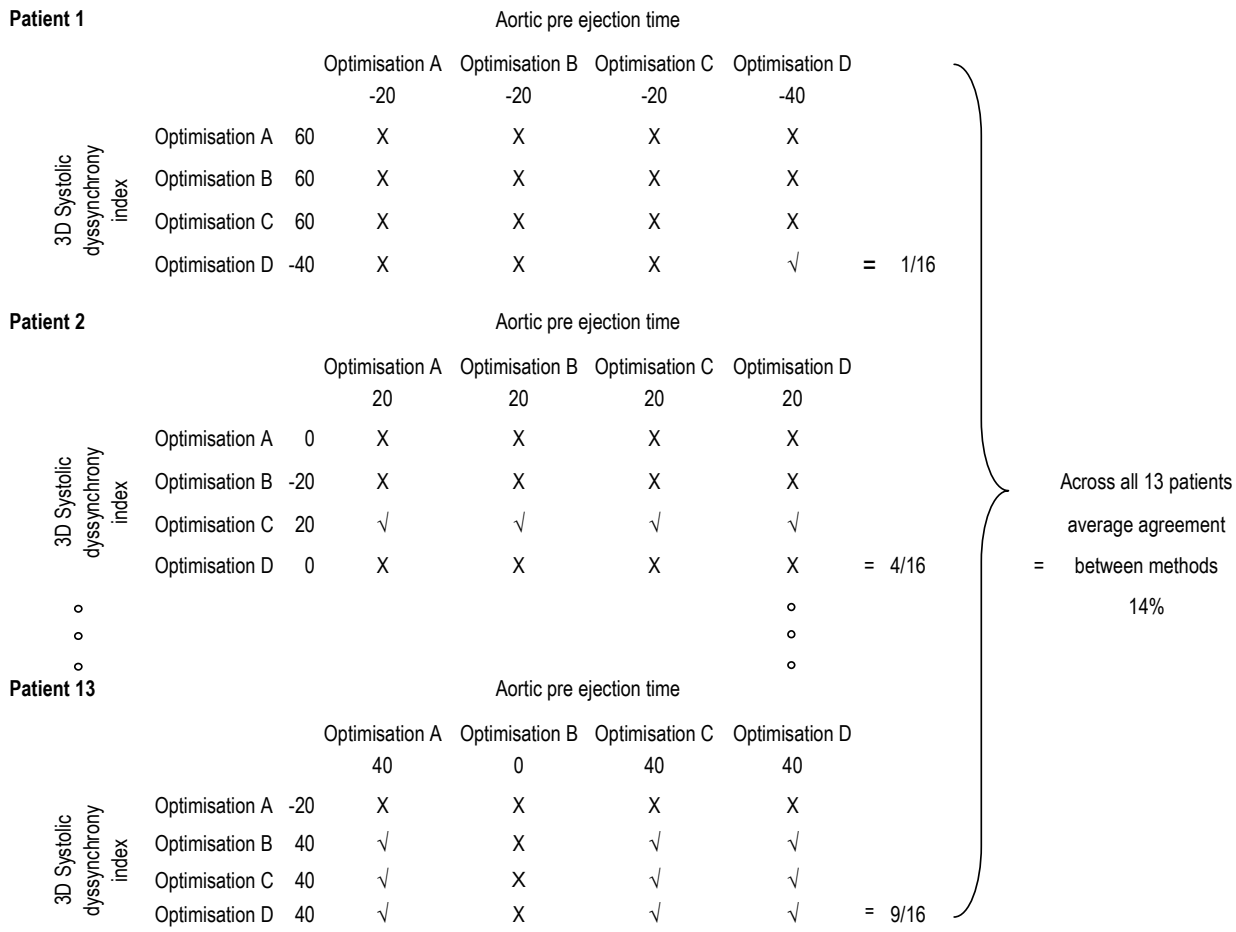


Figure 5-3 Example of how agreement between methods was calculated (3D Systolic dyssynchrony index versus Aortic pre ejection time)

The top panel shows an example of how the percentage agreement between 3D systolic dyssynchrony index measurements and APET measurements was calculated with data shown for 3 of the 13 patients. The lower panel shows the percentage agreement and kappa values between each combination of the 6 dyssynchrony markers.

5.8.1.3 The Scatter of VV optima obtained at repeated optimisations

I calculated the scatter (standard deviation) between the optimal VV delay settings on replicate optimisations for each method, as shown in a representative patient in Figure 5-4 below. The smaller the scatter for a method, the more consistent the VV optima within individual patients will be using that method.

For 3D SDI the scatter of optima was 23 ms, TDI 26 ms, aortic pre-ejection time 14 ms, IVMD 28ms, LVOT VTI 21 ms and QRS duration 14 ms.

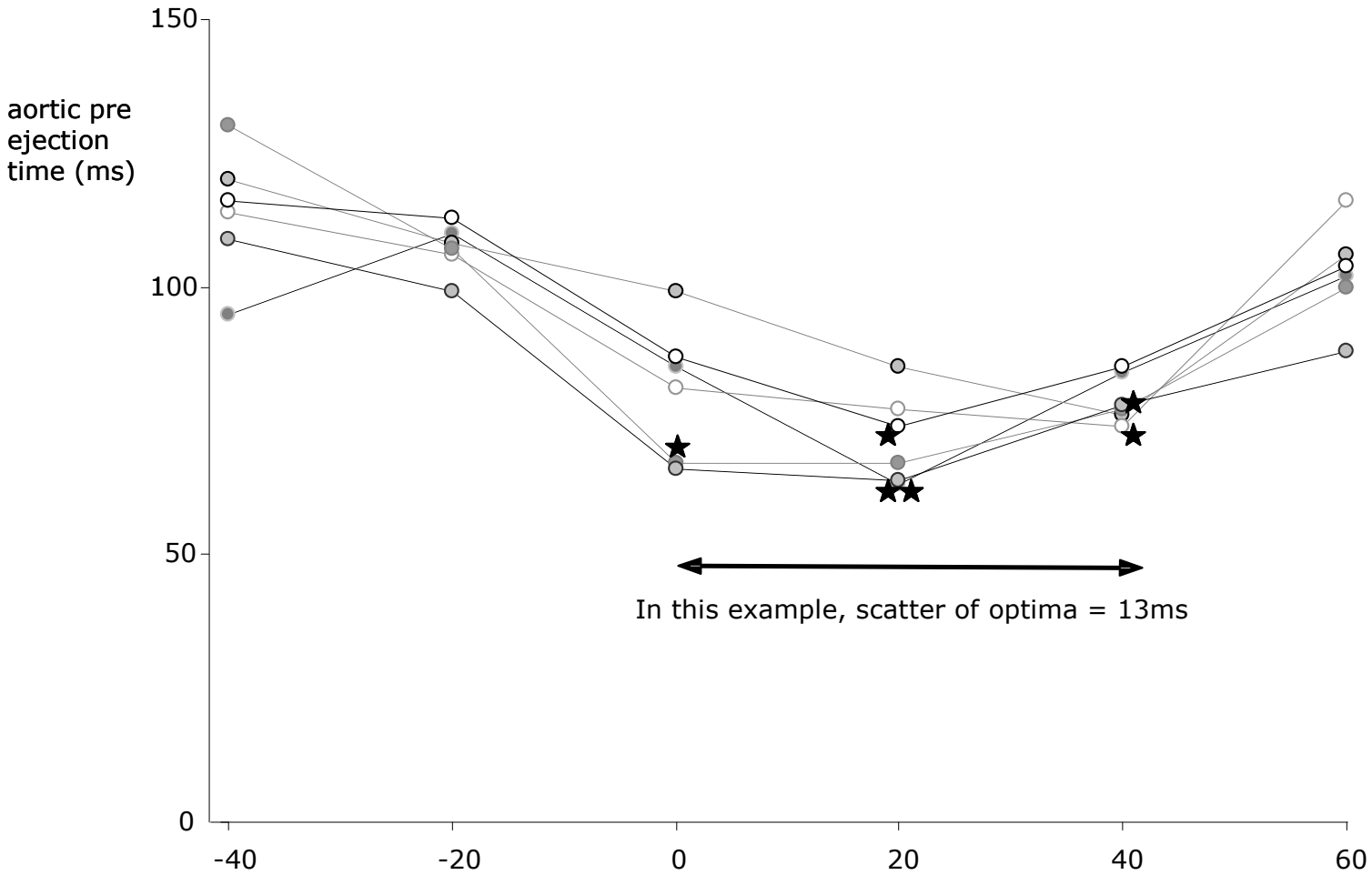


Figure 5-4 Calculation of scatter of optimum VV delays in a representative patient using repeat measurements of aortic pre-ejection time as the marker of dyssynchrony.

There were 6 separate optimisation sequences performed using APET, generating 6 optimal VV delay settings. We quantified the scatter of optima by calculating the standard deviation or scatter of the individual optima identified, each marked by ★ ,

5.8.1.4 Quantification of genuine change versus biological variation: Intraclass correlation coefficient

I calculated the intraclass correlation coefficient (ICC), which indicates the extent to which difference between single measurements at different VV delays is due to true differences between the VV delays rather than random beat-to-beat variation.

The data from a representative patient, using aortic pre-ejection time as the dyssynchrony marker, is shown in Figure 5-5. As can be seen, the scatter between replicate measurements was small in comparison to the between-setting differences, hence a high ICC was obtained in this example patient using aortic pre-ejection time as the modality.

The ICC for single replicates across all patients was 0.34 for 3D systolic dyssynchrony index, 0.32 for tissue Doppler measurements, 0.57 for aortic pre-ejection time, 0.47 for interventricular mechanical delay, 0.54 for LVOT VTI and 0.63 for QRS duration.

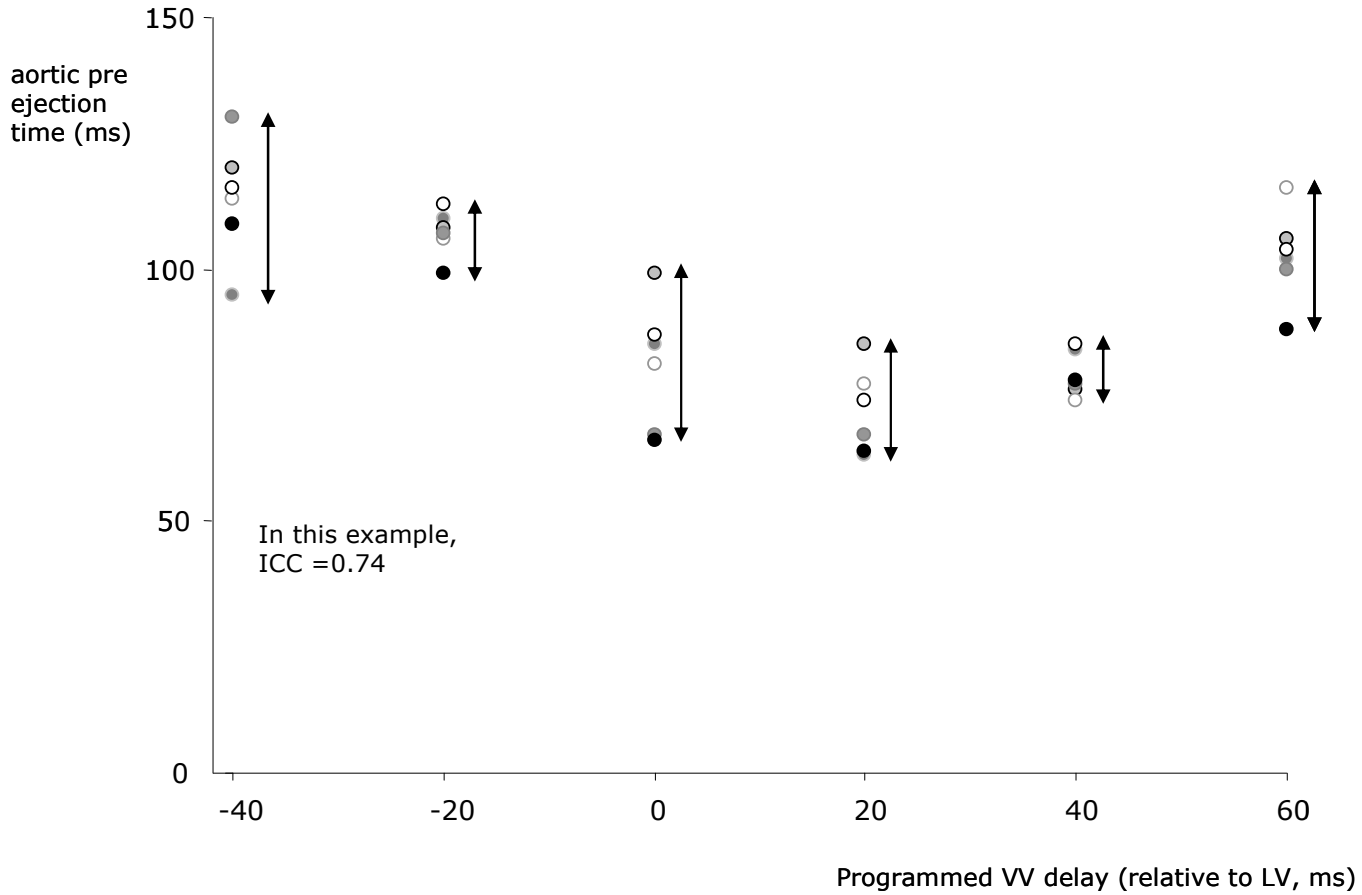


Figure 5-5 Six replicate measures at each VV delay using aortic pre-ejection time as the dyssynchrony marker.
The replicate measures at each setting are not exactly the same. The ICC is defined as the ratio between the variance of the means at each setting, and the variance of all the raw data points. In this example the ICC is 0.74.

5.9 Multiple Replicate Measurements

I then examined the effect of using averages of multiple replicate measurements of the variable being optimised, instead of single measurements.

For those modalities where 8 replicates were possible, namely aortic pre-ejection time, IVMD and LVOT VTI, I grouped the raw replicate measurements into 4 groups of 2 “averaged measurements”. For TDI and 3D SDI where 4 replicates were possible, I grouped the raw replicate measurements into 2 groups of 2 “averaged measurements”.

5.9.1 The effect of multiple measurements on the scatter of VV optima obtained

The scatter of optima obtained fell, when averages of pairs of measurements are taken, instead of single measurements. The scatter of optima fell from 23 ms to 18 ms for 3D SDI, from 28 ms to 15 ms for TDI, from 14 ms to 10 ms for aortic pre-ejection time, 28 ms to 22 ms for IVMD, 21 ms to 16 ms for LVOT VTI and 14 ms to 10 ms for QRS duration ($p = 0.00047$, Figure 5-6).

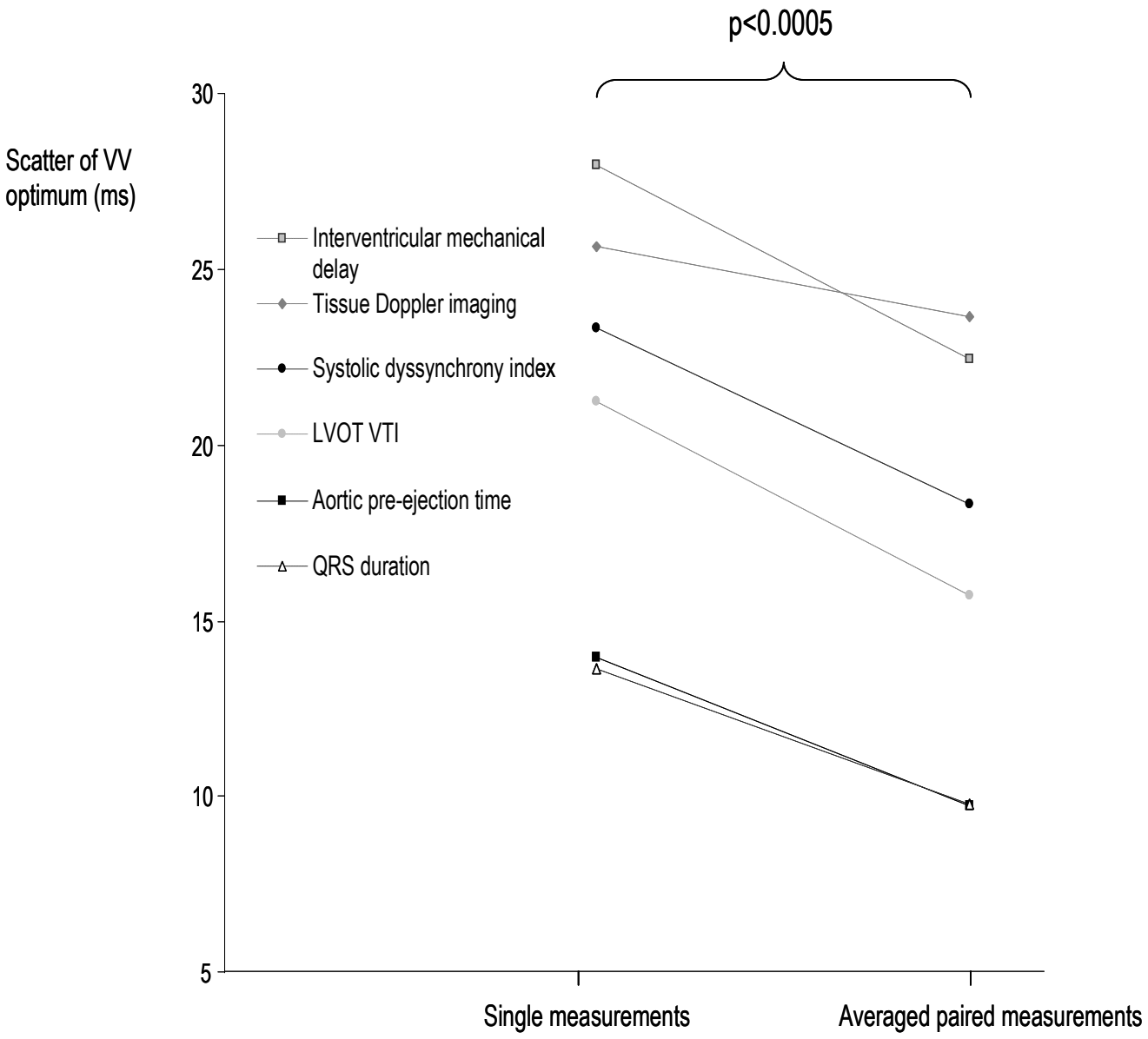


Figure 5-6 The scatter of VV optima obtained using each dyssynchrony parameter at rest with single measurements and averaging paired replicate measurements.

5.9.2 The effect of multiple measurements on the intraclass correlation coefficient

The intraclass correlation coefficient were systematically higher for the averaged measurements than single measurements across all modalities. Using single measurements compared to using averaged paired measurements increased the ICC from 0.34 to 0.59 for 3D SDI, from 0.32 to 0.51 for TDI, from 0.57 to 0.72 for aortic pre-ejection time, 0.47 to 0.59 for IVMD, 0.54 to 0.67 for LVOT VTI 0.63 to 0.74 for QRS duration ($p = 0.00084$, Figure 5-7).

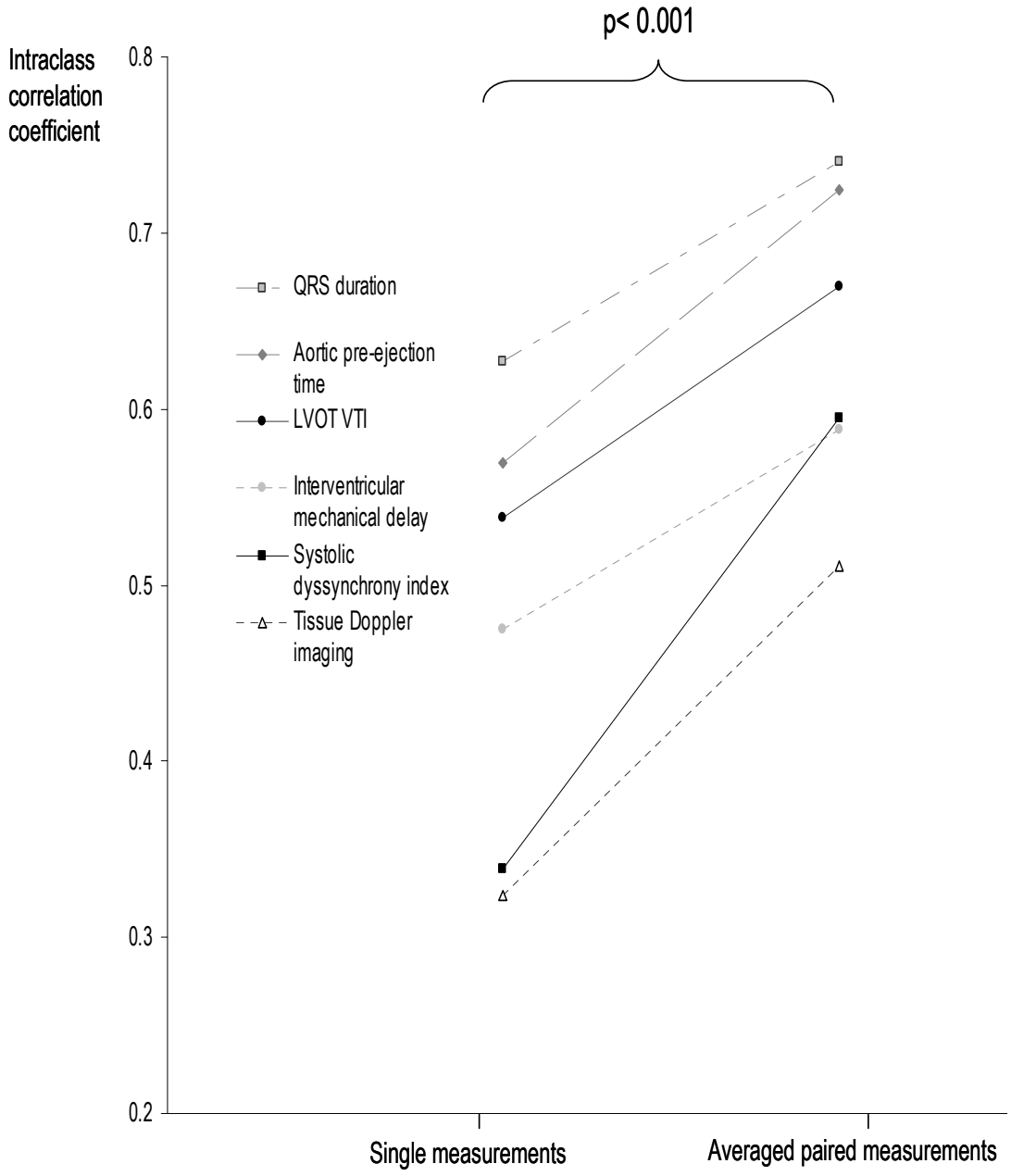


Figure 5-7 The Intraclass correlation coefficient of each modality for single and averaged paired measurements.

5.9.3 Responders and Non-responders

69% of patients (9/13) were clinical responders following CRT implantation and the other 31% patients (4/13) did not clinically improve.

There was no significant difference between the ICC for responders vs. non responders (p = 0.44) across all modalities tested. Nor was there a significant difference in the scatter of optimal VV delay obtained (p = 0.53).

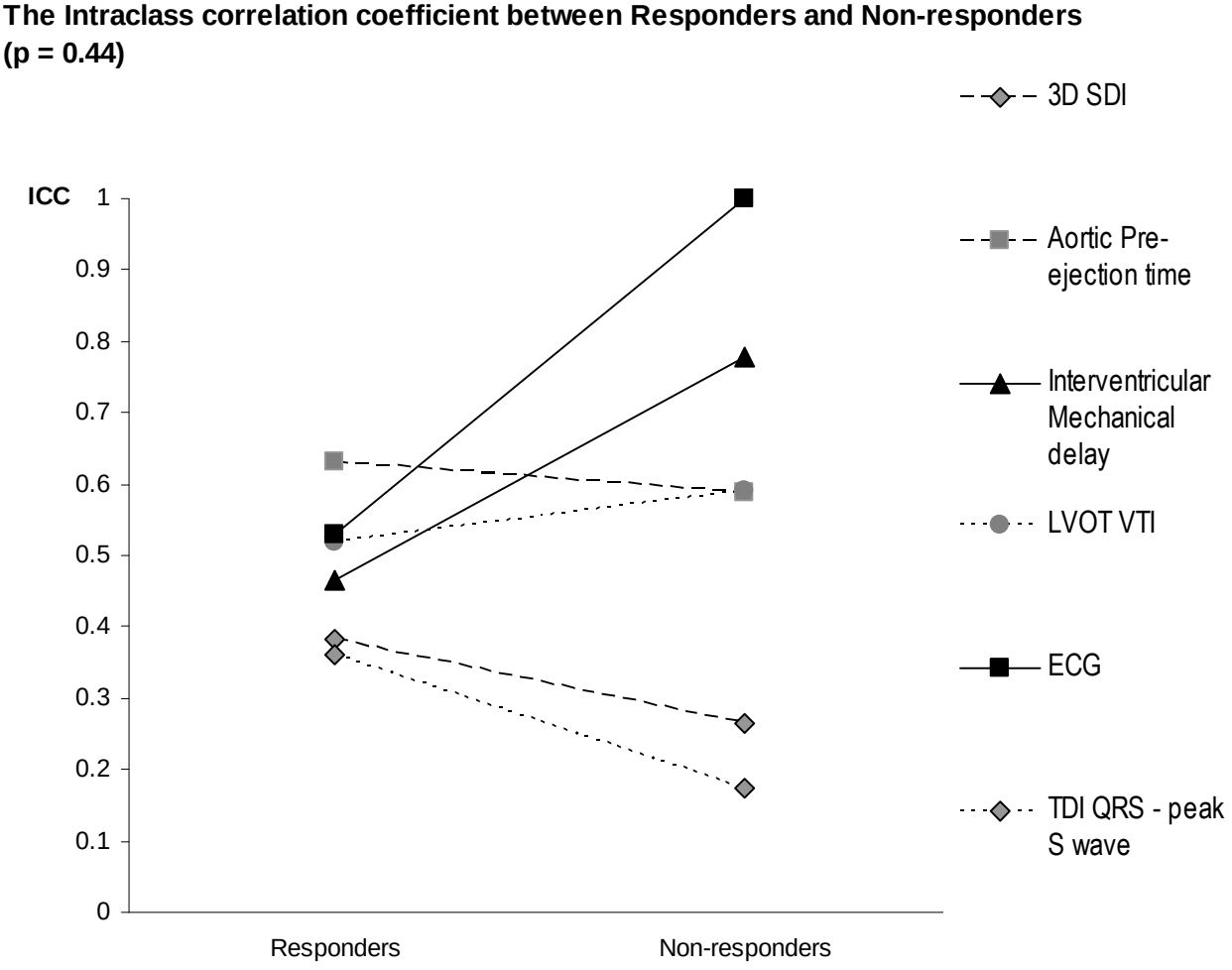


Figure 5-8: The intra class correlation coefficient between responders and non-responders showed no significant difference across all modalities (p=0.44).

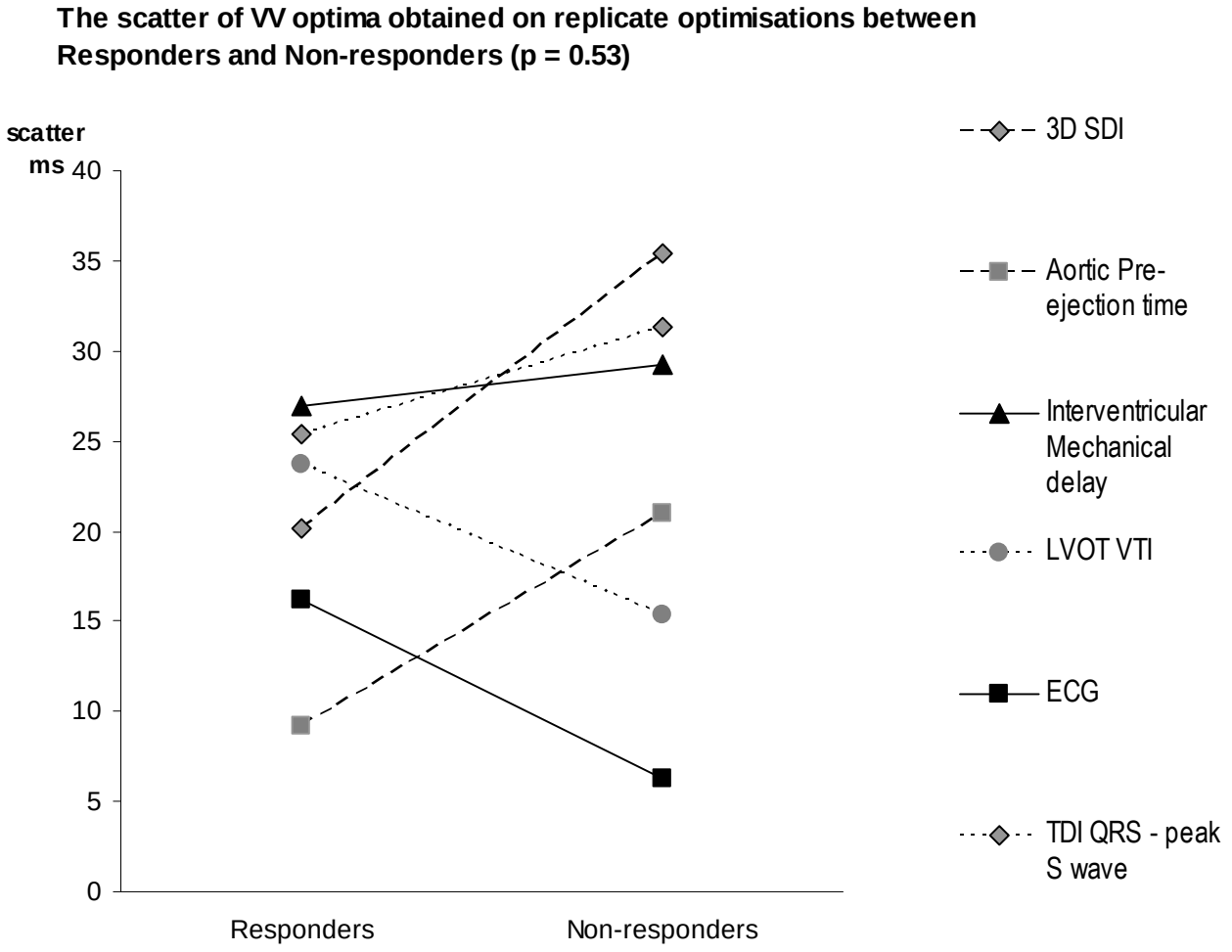


Figure 5-9: The scatter of VV optima obtained when replicate optimisations are performed showed no significant differences between responders and non-responders across all modalities tested (p=0.53).

5.10 Discussion

My study shows that each of these dyssynchrony markers has a very high beat-to-beat variability. Specifically this variability is large compared to the genuine changes in dyssynchrony elicited by changes in pacemaker timings. Together this means that single measures of any of these markers cannot realistically be used to optimise interventricular delay of a CRT device. Moreover, since such beat-to-beat variability swamps the effect of changing VV timings within an individual they are likely to even more powerfully swamp the effect of differences in dyssynchrony between individuals making it unlikely that single measurements of any of these could be a reasonable prospect for a selection criterion for implant of CRT device.

The methodology of this study permitted me to determine the absolute lower limit of measurement variability possible for each of these echocardiographic parameters, because I have eliminated any sonographer, observer or temporal influence.

5.10.1 Why this study design, rather than any other, to assess a dyssynchrony marker?

I developed this study design because it would allow the study findings to be reconfirmed (or refuted) by others without undue burden, and because it efficiently places a lower limit on the measurement variability. Even though patients, operators, and equipment may differ, when all of these are kept constant, and interventricular delay changed dramatically, markers of dyssynchrony should vary correspondingly. I had hoped that for each patient, each marker would show its own clear optimum, and that different markers might mutually agree regarding the optimum.

In fact I found that not only did markers not agree with each other, but they did not even agree with themselves: test-retest reproducibility of the optimum was poor.

The greatest problem with all of the echocardiographic modalities used for VV optimisation is the large beat-to-beat variability in relation to the genuine effect of changes in VV delay, meaning that replicate optimisation sequences (even performed a few minutes apart) produce different optima. Any technique for VV delay optimisation should reliably detect the same optimum every time it is carried out. To achieve this, the measurement must reliably discriminate genuine changes in underlying value from spontaneous beat-to-beat fluctuations (i.e. have a high ICC).

In this head-to-head comparison of these 6 modalities I found that all performed poorly: high beat-to-beat variability, high scatter between optima and low ICC. There is therefore an inevitable poor agreement of optima between different modalities. One might expect the advanced echocardiographic modalities (3D echocardiography SDI and tissue Doppler) - which directly measure the dyssynchrony between the different myocardial segments - to perform better than the other modalities that measure either upstream (e.g. QRS duration) or downstream (e.g. APET, IVMD and LVOT VTI) effects of this dyssynchrony. But they consistently performed no better, and often worse, than the other modalities.

While there have reports of tight reproducibility for these markers (Liodakis et al. 2009; Yu et al. 2004a), in other reports conducted under stringent blinded conditions test-retest reproducibility was underwhelming agreement is noted to be much less than perfect (Chung et al. 2008a; De Boeck et al. 2008; Turcott et al. 2010; Vesely et al. 2008) which is an inevitable consequence of unsatisfactory test-retest reproducibility in one or both markers. Such comparison of an optimisation technique with itself is rarely reported on, except where either the same or another operator reassesses the same data again, rather than reacquiring an entirely new dataset at a later time point. Because of this variability, there is mathematically little hope of

good agreement between modalities for detecting a consistent optimum (Thomas et al. 2009), and consequently even less chance of VV optimisation resulting in improvements in clinical endpoints. Any measurement that incorporates less separate constituent measures may perform better because it has less sources of beat-to-beat variability and hence potential for error. This may explain why LVOT VTI and APET (which both only a single measurement per VV delay setting) have, both in this study and others (Thomas et al. 2009), been shown to have a higher probability of detecting the same optimum on different optimisations, less scatter of optima and a higher ICC than techniques such as TDI and 3D SDI (which require measures of several different myocardial walls/ segments for each VV delay setting).

5.10.2 Can anything be done to improve the situation?

Averaging of replicate optimisation sequences reduces the impact of the beat-to-beat variability, and therefore reduces the scatter and improves the ICC of replicate optimisations. Although planners should incorporate this process into all optimisation protocols (and in particular into any study protocol hoping to have a chance of demonstrating any clinical benefit from VV optimisation), this is rarely performed because of the extra time needed for acquisition and data analysis per patient. The resulting trial, although possessing a high patient throughput, may have no hope of efficacy.

5.11 What do these results mean for clinical VV delay optimisation?

To be practical for clinical purposes, an optimisation process needs to consume only a reasonable amount of time, while achieving an acceptable degree of reproducibility. This study shows that for all the modalities tested, averaging replicate measurements is essential to have a reasonable reproducibility. Even with averages of 4 or 8 replicates, this study showed that reproducibility was still not at such a level that is

likely to be considered acceptable to most clinicians. More replicates and hence more time would be required.

I can estimate easily how long it would take to measure enough replicates to achieve an optimisation to a level that a clinician might find acceptable. If a raw optimisation technique had a scatter of optima (SD) of 28ms, to reduce the SD to 5ms, so that the 95% confidence interval could be ±10ms, would require $(28/5)^2$ replicates, i.e. 31 replicates. In Table 5-1 we show how many replicates would be needed for each of the methods tested. Even in experienced hands, it can take up to 5 minutes to acquire and analyse a 3D image, therefore 30 minutes for a single-replicate optimisation between 6 settings using 3D dyssynchrony. To narrow the uncertainty of the optimum to ±10ms would require 22-fold replication: this would consume 11 hours, a time so long that it would be unlikely to be clinically acceptable. Yet to carry out fewer replicates means accepting an uncertainty in the optimum that is correspondingly wider and therefore itself may be clinically unacceptable. If these dyssynchrony markers cannot consistently distinguish between dramatically different paced interventricular delays within an individual, then whether they could ever be used for CRT selection criteria is questionable?

	Scatter (ms) for single replicate optimisation across all patients	The number of replicates required to confidently provide optima ± 10ms
3D SDI	23	22
Tissue Doppler imaging - 2 seg	26	26
Aortic Pre Ejection time	14	8
Interventricular mechanical delay	29	33
LVOT VTI	21	18
QRS duration	14	7

Table 5-1 Calculation of the number of replicates required to provide reliable optima with a 95% confidence internal of ± 10ms, for each method.

I recommend that research and clinical practice with dyssynchrony markers should specify explicitly the number of replicates that should be measured and averaged, based on (blinded) measurements of beat-to-beat variability.

5.11.1 Why are our results "worse" than commonly supposed

I have not excluded any patients and hence those who had difficult images to acquire and analyse were included. This was not a study where I picked the best images, but instead those which were representative of the CRT device population who have, in some cases, more challenging image quality.

I have included successive beats without any exclusions except for ectopic beats and the subsequent beat which would be altered as a result of the ectopic.

I used no more sophistication in our analysis than we have reported. Other groups and studies may have used more subtleties, exclusions, re-analysis, unblinding collateral information but these may not been published or fully explained. This may have meant that subsequent externally-monitored, prospectively-recruited trials which were required to use openly declared methods, were doomed to failure (Chung et al. 2008a). I have used published protocols at face value.

Although no multi-centre trials have been performed for VV delay optimisation, similar factors may explain the discrepancy between the clinical results of small single-centre trials of echocardiographic measurement of dyssynchrony to guide CRT device implantation (Bax et al. 2003a; Bleeker et al. 2007c) and those of the large multicentre clinical trials such as PROSPECT, which surprised some experts by failing to identify a single echocardiographic modality that could predict CRT response (Chung et al. 2008a).

5.11.2 Would more training help?

This centre does not have below-average standards of skill, but a track record of research in echocardiography and our practice standards are not below real-world clinical practice.

I do not believe that my blinded test-retest reproducibility data are necessarily poorer than that reported by others (De Boeck et al. 2008; Turcott et al. 2010). Quite often what is reported as reproducibility turns out to be the variation between one viewing and another, of the same recorded heart beat (Zhang et al. 2011).

I have applied the published protocols for these measurements, in the way that routine clinical users are expected to implement them. The only modification I have made is blinding, i.e. withholding during analysis-time the value of the VV delay, a normal scientific precaution to prevent unintentional bias in the result.

I performed the studies, conducted the scans and made the measurements, after training within the field. While it is possible that additional training may have helped, it is not clear from the literature what further duration of training would be required, and what its specific content might be.

5.12 Limitations

In this study I have not studied every echocardiographic marker that has been proposed but rather conducted a detailed analysis of the beat-to-beat variability versus ability to detect things in VV synchrony for a wide spectrum of echocardiographic markers. I was mindful of the acquisition time I was asking these patients to undergo. Analysis time was also not inconsequential. I believe a principal finding of my study is not that any particular echo marker is inadequate as a single measurement, but rather that one cannot assume that markers will be adequate without specific assessment. For a single echocardiographic marker, a study such as this is not difficult to do, and I believe that any centre about to embark on such a marker for clinical purposes in optimisation (or even selection) might consider conducting such a brief study with their chosen marker, to gain reassurance that the marker behaves as they expect, or to modify their plans accordingly.

This study focussed on short-term variability, and may be criticised for not including a re-assessment some months or years later. However, long periods of time permit substantial biological changes to become established, to the extent that any discrepancy between initial and final echocardiographic measurements might easily be evolution of disease. Echocardiographic optimisation cannot therefore be faulted for evolving over such periods of time. Instead, my study focussed on short-term variability where there was no real possibility that changes were due to evolution of disease. In normal clinical practice, one presumably hopes that ones decision remains valid at least between one consultation and the next, which is substantially longer than the few heartbeats over which we have tested and retested optimisations and found them wanting.

I believe that this study shows that so great is the biological beat-to-beat variability of many of the echocardiographic indices proposed for use in optimisation (and also for selection) that the prime attention should be on either identifying markers not subject to such variability, or performing high numbers of replicate measurements of existing markers.

5.13 Conclusions

This study shows that all of the modalities tested for guiding VV optimisation performed poorly on an early test of validity, namely the ability to detect that VV delay has been changed across a wide range of settings. Beat-to-beat test-retest reproducibility is poor. Protocols using these measurements to assess dyssynchrony either in the context of VV delay optimisation or assessment for suitability for CRT implantation, should state beat-to-beat reproducibility and the number of replicate measurements that should be made and averaged to make the reproducibility of the average acceptable.

The complete disagreement between the candidate dyssynchrony markers within individual patients undergoing changes in synchrony means that the number of these which are a candidate for quantifying dyssynchrony is at most 1, and perhaps zero.

High beat-to-beat biological variability may explain why some echocardiographic markers, even when physiologically extremely plausible, have as single measurements been disappointing in selecting patients for CRT implantation. This study design can be applied to assess the plausibility of any proposed dyssynchrony marker, before investigators embark on clinical trials attempting to demonstrate the marker's ability to either predict CRT response or guide VV optimisation.

6 Rapid method for evaluating and improving plausibility of tissue Doppler protocols for assessment of dyssynchrony in patients being considered for cardiac resynchronisation therapy

6.1 Abstract

6.1.1 Background

Conventionally tissue Doppler imaging (TDI) protocols for quantification of dyssynchrony do not undergo early screening for relative sensitivity to genuine changes in synchrony versus spontaneous beat-to-beat variability. I show how this can be done and precision improved. I demonstrate application to the questions of whether to measure (a) onset or peak of velocity; (b) 2, 6 or 12 segments; (c) "earliest-to-latest" or standard deviation of timing; and (d) single or multiple replicates.

6.1.2 Methods and Results

Patients with CRT in situ, underwent a detailed protocol of replicate measurements and blinded analysis, at VV delays from RV-first 40ms to LV-first 60ms. I quantified ability to detect genuine changes in synchrony using intraclass correlation coefficient, ICC, (0 = completely overwhelmed by noise, 1 = perfectly consistent). Tissue Doppler velocity traces were obtained and timing was defined in each of the $2\times3\times2\times2=24$ potential combinations as listed above from (a) to (d).

Single measures of dyssynchrony had low ICC, ranging from 0.32 to 0.54. Averages of pairs of measurements improved ICC significantly from 0.50 to 0.72 (ANOVA $p< 0.001$). ICC was not significantly affected by measuring to onset or peak ($p = 0.31$), the number of segments measures 2 vs. 6 vs. 12 ($p = 0.38$) or measuring absolute differences in wall timings or standard deviation ($p = 0.86$).

Scatter between replicate optima showed broadly a similar pattern being significantly narrower with averages of two than singles (ANOVA $p<0.001$).

6.1.3 Conclusions

Despite physiological plausibility, tissue dyssynchrony assessment will never reliably categorise patients if it is unreproducible within a few seconds. Manipulating interventricular delay experimentally allows single centres to test (and improve) the ability of protocols to detect mechanical dyssynchrony. Averaging at least 2 measurements appears essential, but alone may not be enough, since ICC remains low and the width of the 95% confidence limit of optima (4 x standard error) very wide. Reliable long term prediction is unrealistic while sensitivity to dyssynchrony remains low.

6.2 Background

As shown in many studies, tissue Doppler imaging (TDI) is a potential choice for selection of patients suitable for Cardiac resynchronisation therapy. This analysis demonstrates that there is inherent beat-by-beat variability and the TDI 2-segment analysis was disappointing as demonstrated in Chapter 5.

This chapter focuses in more depth to the alternatives available with tissue Doppler imaging, namely the analysis of multiple segments, which some perceive to give a better, more comprehensive and fuller assessment of the dyssynchronous heart.

PROSPECT, the largest, prospective, externally-monitored multicentre trial of predictability of response to CRT from dyssynchrony markers, showed no significant predictability (Chung et al. 2008a). This at first seemed a surprising contrast to numerous unblinded studies which had consistently shown strong predictability of response from baseline tissue Doppler measurements (Bax et al. 2003a; Bleeker et al. 2007c; Gorcsan et al. 2004a; van Bommel et al. 2010b; Yu et al. 2003; Yu et al. 2004a; Yu et al. 2005).

It is not clear whether the error lies in one single study, or simultaneously in numerous others, but the lack of a gold standard for tissue Doppler mechanical dyssynchrony has limited scope for a convincing conclusion. It is not even clear of how to establish a gold standard for dyssynchrony, so that we can determine how precisely each proposed Tissue Doppler index is measuring it, before embarking on an expensive trial.

Blinded test-retest reproducibility is crucial for any technique to have any hope of precise predictive power. This is rarely reported for tissue Doppler measurements: instead, authors often report two analyses of the same digital data, i.e. the same

heartbeat (Zhang et al. 2011). Predictive methods with poor test-retest reproducibility cannot reliably predict the future, because a patient with many very different values in the present will only have one future. There may be methods to improve reproducibility but unless it is recognised to be poor in the first place, clinical researchers may not think to implement them.

But how reproducible is reproducible enough? The level of irreproducibility that is acceptable depends on the size of the genuine differences that the test is designed to measure. While the current confusion prevents us from knowing the true degree of dyssynchrony in any individual patient, we do know that it must change when the VV delay of a CRT pacemaker is changed especially if systematically varied over a wide range from one extreme to another (De Boeck et al. 2008; Pabari et al. 2011; Turcott et al. 2010).

In this study I present a method for quantifying the plausibility of clinical tissue Doppler protocols for assessing mechanical dyssynchrony. I used this method to address several open questions of how dyssynchrony should be quantified and show how this method could be used in the future to improve dyssynchrony assessment protocols efficiently, so that future trials might have prospect for success.

6.2.1 Which tissue Doppler model is correct for assessing dyssynchrony?

This question is still under active debate (Chung et al. 2008a; De Boeck et al. 2008; Palmieri et al. 2010; Zhang et al. 2011). Individual centres tend to champion design choices adopted locally without much evidence that those choices are advantageous over the alternatives and little appetite to carry out small, simple and relatively inexpensive blinded experiments on these questions.

In this study I experimentally manipulate synchrony in patients who already have CRT and quantify the sensitivity of detection of these changes in synchrony (as opposed to random beat-to-beat variability) of various proposed methods for tissue Doppler assessment of dyssynchrony. Using a prolonged protocol of many hours of detailed acquisition, followed by blinded measurement, I simultaneously address, in factorial form, 4 choices in design of tissue Doppler assessment protocols:

- Single measurements or averaging multiple measurements per VV delay?
- Measuring to onset, or peak, of S wave?
- 2, 6 or 12 segmental analysis?
- Earliest-to-latest interval or standard deviation of timings?

6.3 Methods

6.3.1 Tissue Doppler dyssynchrony assessment

There are many ways of using tissue Doppler imaging to assess dyssynchrony. While there are numerous aspects to tissue Doppler, even with the most routine of approaches, namely heterogeneity colour-coded tissue Doppler timings, there remain many choices that a protocol designer can make. These choices include (a) how many wall segments to consider, (b) whether to measure timing from QRS to onset of systolic velocity or to peak, (c) whether to calculate the difference in timing from earliest to latest or as a standard deviation of wall timings, and finally (d) whether it is worth making the effort to measure more than one heart beat at each wall and take the average.

6.3.1.1 2, 6 or 12-segment model

The simplest subdivision of the ventricle is into two walls, septal and lateral, with measurements made at basal levels which has been reported to be diagnostic for mechanical dyssynchrony when it exceeds 60ms (Bax et al. 2003a; Bax et al. 2003b).

A 6-segment approach can also be used (Chung et al. 2008a), by adding the basal segments of the anterior, inferior, posterior and anteroseptal walls to the 2-segment model.

A 12-segment model, which extends the 6-segment model by additionally including measurements at mid-wall level is also widely recommended (Yu et al. 2002; Zhang et al. 2011).

6.3.1.2 Earliest to latest or standard deviation of wall timings

Earliest to latest timings is the difference between the wall which contracts earliest and the wall which contracts last (Bax et al. 2003b)

A standard deviation of timings is also used as a marker of dyssynchrony, which calculates the standard deviation of multiple walls i.e. 6 or 12 (Yu et al. 2002; Zhang et al. 2011).

6.4 Study population

11 outpatients with biventricular pacemakers or biventricular defibrillators previously implanted for clinical indications (Barnett et al. 2007) were invited to enter this detailed study of tissue Doppler imaging. All patients were free from decompensation for greater than 3 months and no changes to medication were made 4 weeks prior to entering the study and until all data had been collected. All patients had intrinsic sinus rhythm, with 100% biventricular pacing (AsVp) and a fixed AV delay programmed to the standard nominal setting of 120ms. 2 patients had only septal and lateral walls imaged due to poor visualisation of the other walls. This is in keeping with feasibility reported in other studies (Bax & Gorcsan 2009). They were included in the appropriate analysis i.e. 2 segment analysis.

At the time of the study, 1 patient was NYHA class I, 4 were NYHA II and 6 were NYHA III. Time between implantation and enrolment was 31 months (range 8 months to 63 months. 4 subjects were male and 7 female, age 67 years ± 8 years. The cause of heart failure was ischemic in 4 subjects and idiopathic dilated in 7 subjects. 10 patients were taking angiotensin-converting enzyme inhibitors or angiotensin-II receptor antagonists, 8 were taking beta-blockers, 7 were taking spironolactone and 4 were taking a diuretic (loop or thiazide). Patients gave informed consent for this study which was approved by the local ethical committee.

Each underwent a detailed protocol of replicate measurements and analysis under blinded conditions, at a series of VV delays from RV-first 40ms to LV-first 60ms.

6.5 Measurements

6.5.1 Image Acquisition

Tissue Doppler imaging was performed with a 2.5 MHz phase array transducer. The gain settings, filters and pulse repetition frequency were optimised. Apical images of the left were acquired with the sector size and depth optimised for the best frame rate, with a minimum frame rate of 90 Hz (Gorcsan et al. 2008) . Images were acquired with the areas of interest i.e. the mitral valve annulus, in the centre of the sector showing clear myocardial definition. Tissue Doppler velocity traces were obtained at the standard 12 sites by acquiring 4 chamber, 3 chamber and 2 chamber views of the left ventricle with the patient holding their breath at end-expiration.

Pacemaker settings were changed in the usual way to program a range of VV delays: RV first at 40ms, 20ms, VV 0ms, LV first 20ms, 40ms and 60ms. I acquired 4 sets of replicate measurements (from separate beats) of colour tissue Doppler images. For each of the 4 replicate optimisations a separate optimum VV delay was identified as the VV delay giving the least dispersion in timing of segmental contraction. This yielded 4 replicate assessments of VV optimisation.

6.5.2 Image Analysis

Images were digitised and analysed offline using Philips Qlab 7.0 3DQ and 3DQ advanced software. Several questions were to be answered in my analysis:

1. I used either (1) the first and (2) the average of 2 replicate measurements per VV delay setting to identify the optimum VV delay setting

2. Timing was defined as either (a) the time from onset of QRS to onset of S wave and (b) the time from onset of QRS to the peak of S wave (Figure 6-1).

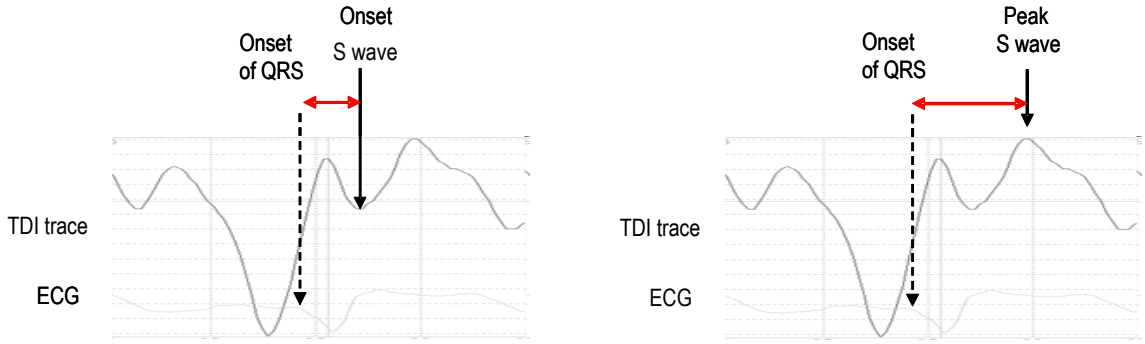


Figure 6-1 Demonstration of measuring from onset of QRS to onset of S wave (left panel) and from onset of QRS to peak of S wave (right panel)

3. Multiple segmental analysis was performed. I used 2, 6 and 12 segment tissue Doppler analysis. To obtain time–velocity curves, a sample volume was placed within the segments at the region of interest i.e. for the septal wall as shown below in Figure 6-2.

4. Dyssynchrony was defined as either (i) absolute difference between the earliest and latest segments or (ii) standard deviation of the segments.

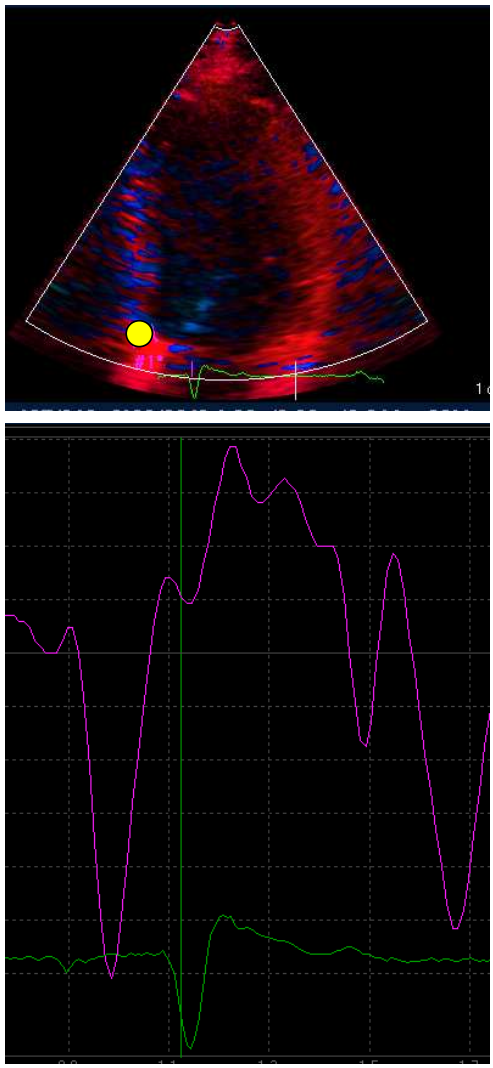


Figure 6-2 showing a tissue Doppler image of the 4 chamber view, with the basal septal wall highlighted. Below is the time-velocity curve which is produced.

2 segments involved the basal septum and basal lateral walls; 6 segments involved the basal segments of the septum, lateral, anterior, inferior, posterior and anteroseptal walls; and 12 segments involved the basal and mid segments of the septum, lateral, anterior, inferior, posterior and anteroseptal walls. All measurements were taken by me.

6.6 Results

6.7 Using single measurements versus averages of paired measurements

Using averages of paired measurements compared to single measurements when performing optimisation of VV delay for all models of tissue Doppler imaging analysis showed a higher ICC as shown in Table 6-1 and Figure 6-3, ANOVA repeated measures $p<0.001$, $F=361$.

Table 6-1 Across all patients there was a statistically significant improvement in the intraclass correlation coefficient when averages of paired measurements were used compared to single measurements.

Intraclass correlation coefficient

	Using single measurements	Using averaged paired measurements
Measuring to Onset of S wave		
2 segment	0.32	0.57
6 segment (earliest to latest)	0.49	0.67
12 segment (earliest to latest)	0.47	0.69
6 segment (standard deviation)	0.48	0.63
12 segment (standard deviation)	0.54	0.72
Measuring to Peak of S wave		
2 segment	0.32	0.50
6 segment (earliest to latest)	0.42	0.64
12 segment (earliest to latest)	0.43	0.66
6 segment (standard deviation)	0.39	0.61
12 segment (standard deviation)	0.44	0.63

ANOVA $p<0.001$ ($F=361$)

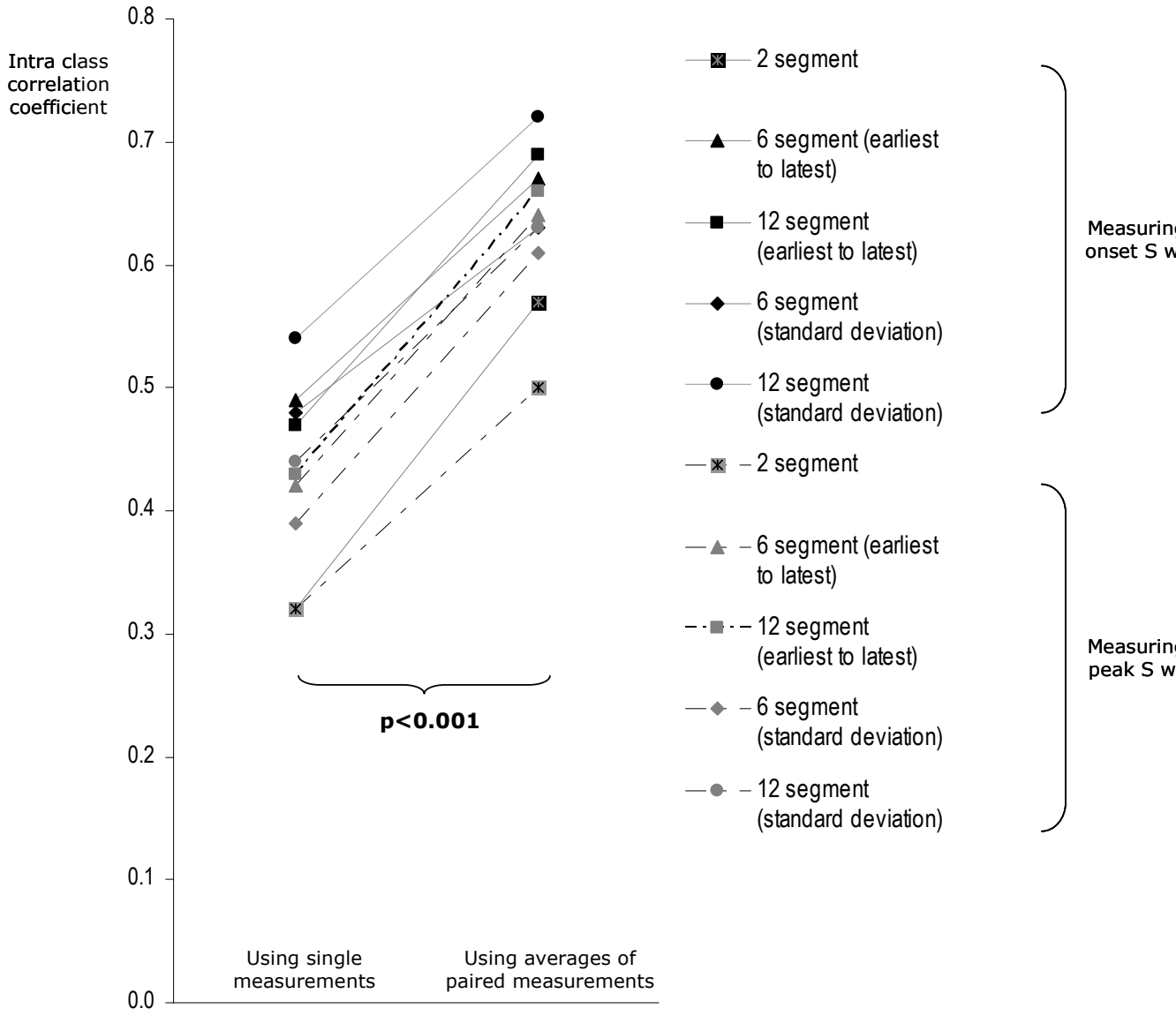


Figure 6-3 The ICC increases significantly across all patients and all measurement methods, when averages of paired measurements are used in comparison to single measurements.

Scatter of optimum

	Using single measurements (ms)	Using averaged paired measurements (ms)
Measuring to Onset of S wave		
2 segment	28	15
6 segment (absolute difference)	18	9
12 segment (absolute difference)	23	12
6 segment (standard deviation)	16	8
12 segment (standard deviation)	21	6
Measuring to Peak of S wave		
2 segment	26	24
6 segment (absolute difference)	19	11
12 segment (absolute difference)	13	4
6 segment (standard deviation)	22	14
12 segment (standard deviation)	12	8

ANOVA p=0.001 (F=25.8)

Table 6-2 The scatter of the optima obtained across all patients decreases significantly when averages of paired measurements are used compared to single measures.

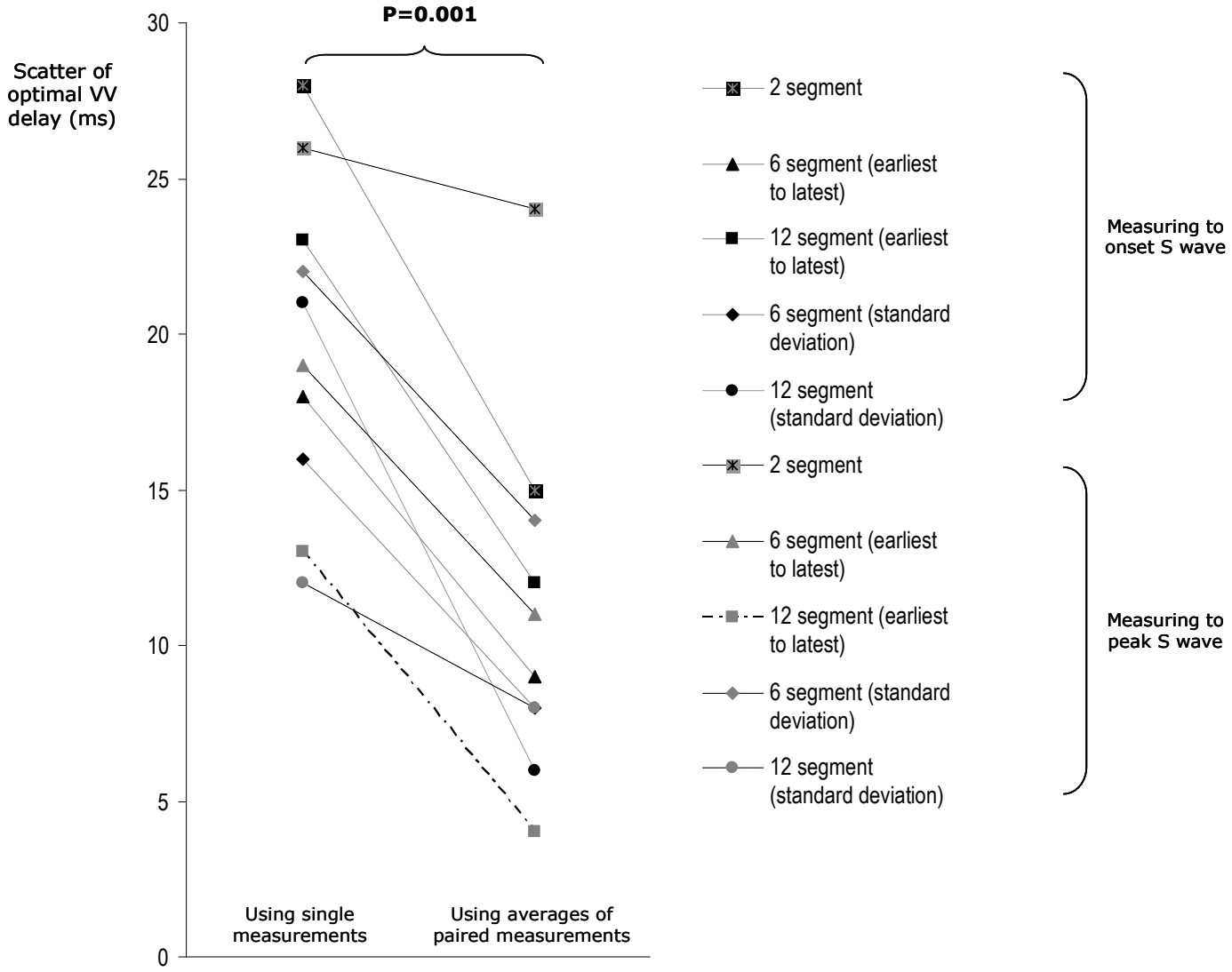


Figure 6-4 The scatter of optima is significantly narrower when averages of pairs of measurements are used compared to single measurements.

6.7.1 Measuring to onset of S wave or to peak of S wave.

The intraclass correlation coefficient (ICC), which indicates the extent to which difference between single measurements of TDI at different VV delays is due to true differences between the VV delays was calculated. The ICC averaged across all patients for 2, 6 and 12 segment tissue Doppler imaging is shown in Table 6-3. There was no significant difference when measuring either from the onset of QRS to the onset of S wave or from onset of QRS to the peak of QRS.

Table 6-3 Intra class correlation coefficients of measuring to the onset of systolic wave and to the peak of systolic wave. There is no significant difference in the intra class correlation whether you measure to the onset of systolic wave or to the peak.

Intraclass correlation coefficient

	Measuring to Onset of S wave	Measuring to Peak of S wave
Using single measurements		
2 segment	0.32	0.32
6 segment (earliest to latest)	0.49	0.42
12 segment (earliest to latest)	0.47	0.43
6 segment (standard deviation)	0.48	0.39
12 segment (standard deviation)	0.54	0.44
Using averages of paired measurements		
2 segment	0.57	0.50
6 segment (earliest to latest)	0.67	0.64
12 segment (earliest to latest)	0.69	0.66
6 segment (standard deviation)	0.63	0.61
12 segment (standard deviation)	0.72	0.63

ANOVA p=ns (0.31)

The scatter of the optimum obtained when measurement of tissue Doppler were take at a series of VV delays are shown in Table 6-4. Across all methods of analysis of 2,6 and 12 segment analysis showed no significant difference between measuring to the onset of the S wave or measuring to the peak of the S wave, ANOVA p= 0.90.

Table 6-4 Scatter of optimum measuring to the onset of systolic wave and to the peak of systolic wave.

There is no significant difference in the scatter of optima obtained whether you measure to the onset of systolic wave or to the peak.

Scatter of optimum

	Measuring to Onset of S wave	Measuring to Peak of S wave
Using single measurements		
2 segment	28	26
6 segment (earliest to latest)	18	19
12 segment (earliest to latest)	23	13
6 segment (standard deviation)	16	22
12 segment (standard deviation)	21	12
Using averages of paired measurements		
2 segment	15	24
6 segment (earliest to latest)	9	11
12 segment (earliest to latest)	12	4
6 segment (standard deviation)	8	14
12 segment (standard deviation)	6	8

ANOVA p=ns (0.90)

6.7.2 Comparison between differences of absolute values and standard deviation calculations of the timings of 6 and 12 segment models of tissue Doppler analysis

Measuring the absolute difference between the walls of the segments measured compared to the standard deviation of the timings of the walls made no difference to the ICC *or* to the scatter of optima obtained across 6 or 12 segmental analysis. This is shown in Table 6-5.

Table 6-5 Measuring earliest to latest versus standard deviation between wall timings. The Intraclass correlation coefficient and the scatter of the optima are not significantly different when the earliest to latest timings are used, or whenever the standard deviation of timings are used for 6 or 12 segment TDI models.

	Intraclass correlation coefficient		**Scatter of optimum (ms)**	
	Earliest to latest between walls	Standard deviation between wall timings	Earliest to latest between walls	Standard deviation between wall timings
Using single measurements				
6 segment (onset)	0.49	0.48	18	16
12 segment (onset)	0.47	0.54	23	21
6 segment (peak)	0.42	0.39	19	22
12 segment (peak)	0.43	0.44	13	12
Using averages of paired measurements				
6 segment (onset)	0.67	0.63	9	8
12 segment (onset)	0.69	0.72	12	6
6 segment (peak)	0.64	0.61	11	14
12 segment (peak)	0.66	0.63	4	8
	ANOVA p=ns (0.86)		ANOVA p=ns (0.68)	

6.7.3 2 versus 6 versus 12 segments

There is no effect on the intra class correlation coefficient whether 2, 6 or 12 segments were used. There is weak evidence on the scatter of the optimum identified (ANOVA p=0.035) with a trend to a wider scatter with a 2-segment model (Table 6-6).

Table 6-6 The effect of measuring 2, 6 or 12 segments for TDI analysis.

The top panel shows the effect on the intra class correlation coefficient and the lower panel shows the effect of the scatter of the optimum identified when repeated optimisations have been performed.

Intraclass correlation coefficient

	Using single measurements (ms)		Using averages of paired measurements (ms)		
	Measuring to Onset of S wave	*Measuring to Peak of S wave*	*Measuring to Onset of S wave*	*Measuring to Peak of S wave*	
2 segment	0.32	0.32	0.57	0.50	comparing 2, 6 and 12-segments
6 segment	0.49	0.42	0.67	0.64	ANOVA p=ns (0.38)
12 segment	0.47	0.43	0.69	0.66	

Scatter of optimum

	Using single measurements (ms)		Using averages of paired measurements (ms)		
	Measuring to Onset of S wave	*Measuring to Peak of S wave*	*Measuring to Onset of S wave*	*Measuring to Peak of S wave*	
2 segment	28	26	15	24	comparing 2, 6 and 12-segments
6 segment	18	19	9	11	ANOVA p=0.035
12 segment	23	13	12	4	

6.8 Discussion

This study shows that for all methods of analysing colour coded tissue Doppler imaging measurements there is high spontaneous beat-to-beat variability and a correspondingly low intra class correlation coefficient i.e. poor signal-to-noise ratio. Single measures of any of the approaches here cannot realistically claim to be reliable predictors of the future because they do not even predict themselves a few seconds later.

6.8.1 Beat-to-beat variability taboo

Blinded beat-to-beat variability is rarely discussed in unblinded studies that report marked predictor values of dyssynchrony for clinical outcomes.

Sometimes reanalysis of the same heart beat (by the same or different operator) is reported (Zhang et al. 2011) but this is not test-retest reproducibility because it omits biological variability between beats. Reanalysis of the same beat yields correlation coefficients in the range of 0.95 to 0.99 (Palmieri et al. 2010) which can therefore be seen to be only a very small minority of the clinically relevant total variability.

It is possible that some centres have developed special positioning of the patient and / or probe, or have special protocols for selecting beats that have unusual consistencies but, if so, these appear not to have been disclosed and put readers at a disadvantage when reproducing the work locally.

6.8.2 Inability to detect resynchronisation when it occurs

Difficulties in detecting resynchronisation arise from the lack of certainty in the results obtained. This is particularly concerning, when a measurement is performed on 2 consecutive occasions and grossly different results are obtained. The method used, whether it be tissue Doppler imaging, as in this study, or an alternative is then brought into question. To ensure that we are reliably detecting resynchronisation, we need to have narrow confidence intervals, within which is an acceptable range.

In this study I have shown the scatter of optimum that have arisen when tissue Doppler imaging has been used in various models. I had tested a wide range of tissue Doppler definitions of dyssynchrony including testing the definitions of timings based on the onset of S-wave, and the peak of S-wave to allow a comprehensive analysis of the various possibilities. I tried hard not to favour on criterion over another.

The 95% confidence interval of these values equips us with the knowledge of the range of values we may obtain whether we need to use:

- single measurements or averaging multiple measurements per VV delay?
- measure to onset, or peak, of S wave?
- 2, 6 or 12 segmental analysis?
- earliest-to-latest interval or standard deviation of timings?

Figure 6-5 shows the confidence intervals of each tissue Doppler parameter and the reduction in width of confidence intervals when paired measurements are averaged.

Width of 95% Confidence Intervals

	Using single measurements (ms)	Using averaged paired measurements (ms)
Measuring to Onset of S wave		
2 segment	112	60
6 segment (earliest to latest)	72	36
12 segment (earliest to latest)	92	48
6 segment (standard deviation)	64	32
12 segment (standard deviation)	84	24
Measuring to Peak of S wave		
2 segment	104	96
6 segment (earliest to latest)	76	44
12 segment (earliest to latest)	52	16
6 segment (standard deviation)	88	56
12 segment (standard deviation)	48	32

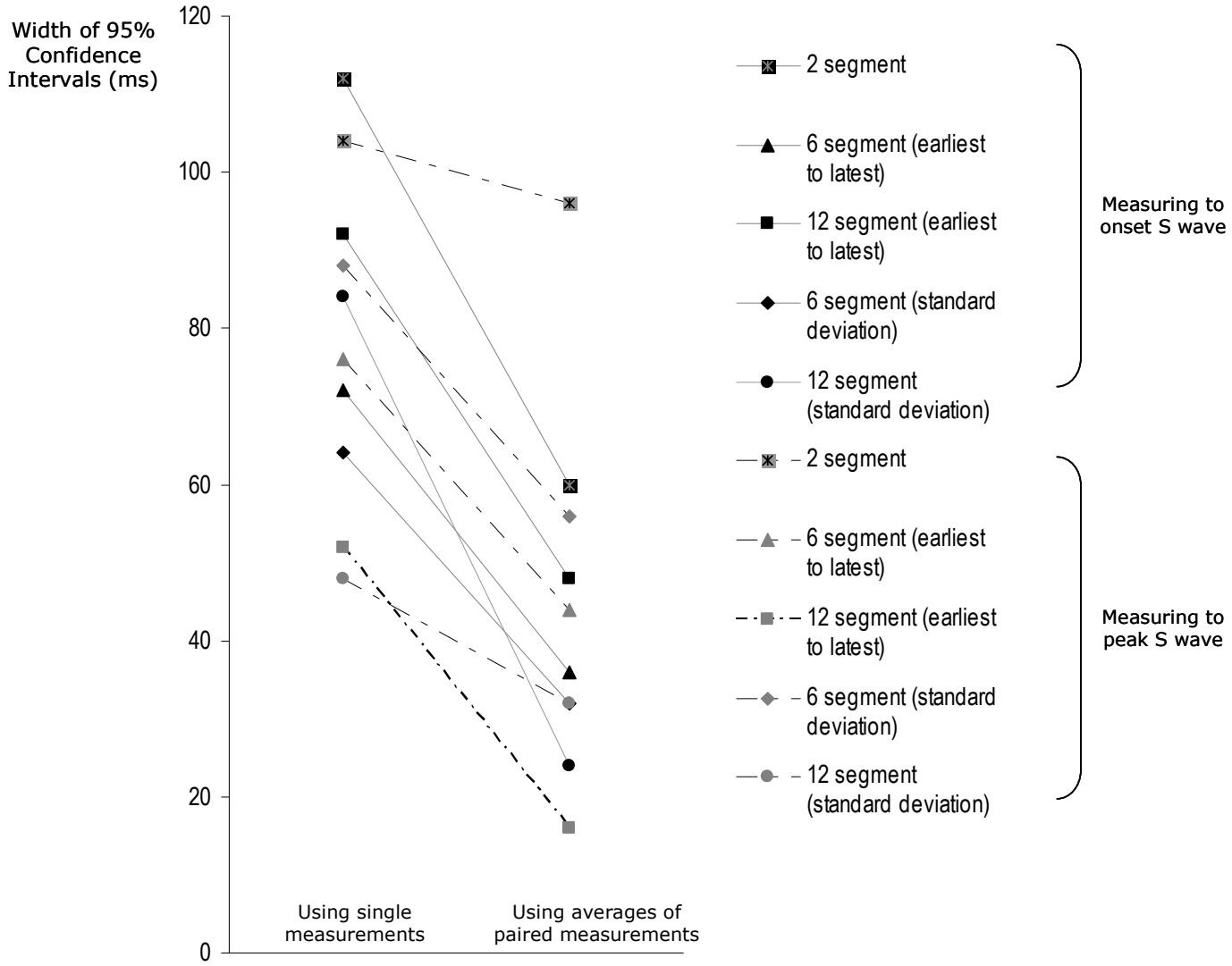


Figure 6-5 The width of 95% confidence intervals which will surround the optimum detected are shown above for each possibility of tissue Doppler measurements used. The graphical representation shows the dramatic decline in this width when averages of pairs are used. Numerically we can see the smallest width with 12 segment earliest to latest using paired measurements.

6.8.3 Implications for protocol designs and reproducibility

To ensure that a design protocol will reliably test what it is meant to i.e. identify the optimal setting with the method used, then it is important to establish the test retest reproducibility to ensure we are happy with the results, and that this will then give a result which has an acceptable confidence interval. To date, the highly transparent studies looking at test-retest reproducibility have failed to confidently point towards tissue Doppler imaging as a method which can be used for this purpose (De Boeck et al. 2008; Vesely et al. 2008).

However, every method can be improved upon and hence use averages of multiple settings to improve the sensitivity and reduce the confidence interval of the optimum obtained (Pabari et al. 2011).

Some groups have repeatedly demonstrated that colour coded TDI is suitable and able to identify patients for CRT implantation. Other groups have found that this modality is limited in its test-retest reliability (Vesely et al. 2008) and in a multicentre trial the results for all criteria of TDI analysis were modest (Chung et al. 2008a). These differences could be clarified if protocols were designed using the suggestions arising from this study

6.9 Limitations

This study may be criticised for having too few patients, but this was a detailed meticulous protocol which was time consuming for the patients and not ethical to apply to even more patients until methods are improved – which could then be tested by this new method, on even a single pilot patient, to establish at least one adequate ICC before formally designing a future study.

We did not follow up the patients for a period of time afterwards and rescan them, with good reason. Because each patient (at each setting) had a wide variety of possible optimal values, any setting could be claimed to be optimal by one criterion or another. Without establishing a specified protocol for selecting an optimal value, looking back from the future to find a fit would risk confirmation bias which would artificially enhance correlation coefficients.

TDI has many limitations however it is still one of the most extensively studied echocardiographic methods for assessment of LV dyssynchrony. Limitations include (1) angle dependency which can alter velocities significantly, (2) measurements of wall motion does not indicate whether it is active contraction or passive motion, with even akinetic walls being dragged across, and (3) the fact that several views are required for all walls to be evaluated because all walls cannot be captured in a single apical plane. Multiple acquisitions mean that different beats are being used for different walls (Agarwal et al. 2009).

6.10 Conclusions

Tissue Doppler imaging as a marker for optimising CRT is disappointing. It has a low sensitivity even for detecting large changes applied exponentially with everything else held constant. With all the potential combinations of measurements available to us namely measuring (a) onset or peak of velocity; (b) 2, 6 or 12 segments and (c) "earliest-to-latest" or standard deviation of timing, none convincingly improved the sensitivity of this marker to a level that might be reasonable to apply in day-to-day clinical practice. Averaging multiple replicates is currently the only plausible approach to reducing the wide confidence intervals and improving ICC in order to make tissue Doppler imaging a reliable predictor and marker for optimising CRT. However the number of replicates needed may not be palatable.

7 Comparison of the reproducibility of left ventricular outflow tract velocity time integral (LVOT VTI) versus peak velocity (V_{max}) as targets for optimisation of VV delay in cardiac resynchronisation therapy.

7.1 Abstract

7.1.1 Background

Clinically practised echocardiographic optimisation of VV delay in Cardiac Resynchronisation therapy (CRT) typically seeks to maximise flow through the left ventricular outflow tract (LVOT), conventionally assessed using the velocity–time integral (VTI). In this study I compare VTI maximisation against maximisation of peak velocity (V_{max}), to determine which provides a more reproducible optimum.

7.1.2 Methods

I measured LVOT VTI and V_{max} at 6 different settings of VV delay in 40 subjects and determined the optimal VV delay settings using each Doppler measure. We compared the scatter of the identified optima and the intra-class correlation coefficient (ICC, a measure of signal-to-noise ratio) between the two techniques.

7.1.3 Results

Optimisation by V_{max} showed a smaller scatter between repeat optimisations (SD 4ms narrower, $p=0.001$) than optimisation by VTI. Using averages of triplicate measures for optimisation showed a smaller scatter than single measures (SD 5ms narrower, $p<0.0001$).

ICC was significantly stronger for V_{max} than VTI (by 0.13, $p<0.0001$). It was higher for averages of triplicate measures (by 0.05, $p=0.001$).

Despite the greater precision, time taken to analyse a 6-setting optimisation was shorter using V_{max} than VTI (17.5 versus 59.1 seconds, $p<0.0001$).

7.1.4 Conclusions

For optimisation purposes, it is critical to be able to detect small differences in signal reliably, and time is at a premium. Measuring peak velocities takes one-third of the time to analyse and delivers significantly more precise localisation of the optimum than measuring VTI. Clinicians choosing to optimise VV delay by LVOT Doppler might therefore be advised to use peak velocity and average multiple replicates if economy and precision are desired.

7.2 Background

There remains controversy regarding whether programming different ventriculo-ventricular delay (VV delay) settings makes any difference to patients with cardiac resynchronisation therapy (CRT) devices. If CRT is truly about resynchronisation, intuitively the programmed VV delay must be important. However, the question of the reproducibility of the standard methods for optimisation is rarely addressed on a formal blinded basis. If VV delay optimisation is to be rejected as worthless because of apparent lack of differences in haemodynamic effects between different VV delay settings, then the term resynchronisation might also be called into question. Although the range of VV delay settings tested during optimisation may not have included one that achieves perfect resynchronisation, it must have surely included some highly undesirable ones. If it does not matter which setting is selected then the marked benefit proven for biventricular pacing might be arising from mechanisms other than resynchronisation of the LV walls (Kyriacou et al. 2011).

Cardiac resynchronisation therapy provides clinical and prognostic benefit to patients with heart failure (Abraham et al. 2002; Boriani et al. 2006; Cazeau et al. 2001) therefore it is rational to attempt to maximise the hemodynamic improvement derived from the devices, by optimising the programmed settings. Clinically however, echocardiographic VV delay optimisation is not routinely performed for two main reasons; because of the time involved to perform optimisation algorithms and because the endpoint studies so far have not consistently demonstrated clinical improvement (Boriani et al. 2006; Vanderheyden et al. 2005).

It is difficult however to find in the literature the early simple studies demonstrating that, in an objective environment, VV optimisation can actually be carried out (i.e.

yields the same results a few minutes later). Without this confirmation, it is not possible to be certain that the VV optimisation procedures now being carried out are reliably identifying a consistent value for each individual, or just yielding a random value. If current VV optimisation protocols are unreliable, then failure of VV optimisation to yield consistent clinical improvements might reflect on the protocols rather than on the concept of resynchronisation itself.

To be of practical value, a protocol must yield the same VV delay optimum on successive optimisations, especially over the short term. If it does not then either the true optimum is varying over such short time intervals (in which case optimisation can never realistically keep up with it to allow consistent resynchronisation), or the protocol is unreliable at detecting the true optimum (in which case the protocol should be improved). A likely strong driver of the relative precision or variability of a VV delay optimisation protocol is the signal-to-noise ratio of the variable being measured during optimisation (Pabari et al. 2011; Whinnett et al. 2008).

Currently there is no gold standard echocardiographic measure used to guide the non-invasive optimisation of VV delay in CRT devices, although left ventricular outflow tract – velocity time integral (Geibel et al. 1991; Sawhney et al. 2004; Sogaard et al. 2002) (LVOT VTI, also called 'stroke distance'), as a practical surrogate for stroke volume is widely used, especially in the research setting (Bertini et al. 2010; Bleeker et al. 2007c; Boriani et al. 2006; Duvall et al. 2010; Geibel et al. 1991; Sawhney et al. 2004; Sogaard et al. 2002; van Gelder et al. 2004; Vanderheyden et al. 2005; Zuber et al. 2008). However if LVOT VTI is insufficiently reproducible (Jansen et al. 2006a), then using it to guide optimisation will yield variable "optima", and conclusions drawn from studies using this method may underestimate the benefit of optimisation. Even if it takes more time to deliver a more reproducible technique for optimisation,

such investment is a sine qua non before expecting clinical benefits. If we had a better way of detecting the optimum, that had been rigorously evaluated, the clinical improvement and endpoints may be stronger and more convincing therefore encourage clinicians to do more routinely.

An inherent property of Doppler echocardiographic assessment of LVOT flow is the physiological variability between beats. Optimisation of VV delay requires detection of genuine persistent changes in cardiac function between VV settings ("signal"), which may be small in comparison to the beat-to-beat variability (noise) of the Doppler trace. A likely strong driver of precision of VV optimisation is the signal-to-noise ratio of the variable being measured (Pabari et al. 2011).

In this study I conduct head-to-head comparison of LVOT VTI and peak velocity (V_{max}) for optimisation of VV delay in patients with CRT devices. First, I quantify the relative immunity of each method to noise, using the intra class correlation (ICC) (Muller & Buttner 1994). This is a number between 0 (indicating a method for which a biological variability completely overwhelms any meaningful signal) and 1 (indicating a measure in which differences between settings are perfectly identical on each evaluation with no disruption by biological variability).

Second I describe the reproducibility of the two optimisation techniques by determining the scatter of the VV optima obtained on repeat optimisations. I calculated the standard deviation (SD) of VV optima from serial repeat optimisations within the same individuals on the same day. Ideally successive optimisations should derive the same optima, hence the standard deviation (SD) or scatter would be zero. In reality there may be slight differences in the optima when successive optimisations are performed therefore if 6 optimisations are performed, 6 optima will be obtained.

The standard deviation (SD), or scatter, is the spread of these results. To measure the change in signal when VV delays are changed, we need to choose a measure with the greatest reproducibility and signal-to-noise ratio (Pabari et al. 2011; Whinnett et al. 2008).

Third, I evaluate the impact of conducting and averaging multiple replicates of each measure (instead of a single measure) on the noise immunity (ICC) and scatter (SD) between replicate optimisations. This will minimise the scatter between optima on repeat optimisation (Pabari et al. 2011). I would predict that in order to further improve the reproducibility of the algorithm, averaging repeat optimisations would be necessary, further increasing the time taken for acquisition and analysis of data during an optimisation procedure.

Finally I determine whether using LVOT V_{max} instead of LVOT VTI reduces the acquisition and analysis time, hence enhancing the feasibility of clinical VV optimisation of CRT devices.

7.3 Methods

7.3.1 Study Population

40 consecutive outpatients with biventricular pacemakers or biventricular defibrillators previously implanted for standard clinical indications (Barnett et al. 2007) were enrolled into this study. 28 subjects were male and 12 female, age 73 ± 7.7 years (mean ± sd).

The cause of heart failure was ischemic in 15 subjects and idiopathic dilated in 25 subjects. At the time of the study, 1 patient was NYHA class I, 17 were NYHA II and 22 were NYHA III. 34 patients were taking angiotensin-converting enzyme inhibitors or angiotensin-II receptor antagonists, 32 were taking beta-blockers, 18 were taking spironolactone and 27 were taking a diuretic (loop or thiazide).

All patients were 100% biventricularly paced and free from decompensation for greater than 3 months.

The time between CRT device implantation and enrolment into the study was 32.3months (range 3 months to 75 months). The LV lead was placed in the optimal position that was possible during implantation. All patients were 100% biventricularly paced during the study and AV delay was programmed to the nominal setting of AV 120ms.

Patients gave prior written informed consent for this study which was approved by the local ethical committee.

7.3.2 Protocol

Patients lay awake, recumbent in the left lateral position on a couch having rested supine for at least 10 minutes. 2 operators were present – one to reprogram the CRT devices via a telemetry head, and the second to acquire the Doppler echocardiographic images. The echocardiographer was blinded as to the pacemaker settings programmed by the other operator.

The programmed VV delay was changed in a random order, and the VV delay settings that were tested were left ventricular activation first by 60ms, 40ms, 20ms and 0ms and right ventricular activation first by 20ms and 40 ms.

7.3.3 Measurements

Echocardiography was performed on a single visit to the hospital with the same experienced operator conducting all echocardiographic measurements using a Philips IE33 machine with a S5-1 transducer. Measurements were taken from the apical 5-chamber view with pulse wave Doppler sampled at the left ventricular outflow tract approximately 1 cm below the aortic annulus as per published guidelines (Gorcsan et al. 2008). Six beats were acquired at each VV delay, at end expiration.

7.4 Analysis

Measurements of peak velocity and VTI were conducted offline, under timed conditions, in a random order, with the operator blinded to the programmed VV delay. VTI measurements were made by tracing around the flow envelope manually as per usual practice. The peak measurement was made from the same captured images but independently of, and blinded to, the VTI tracing.

7.4.1 Determination of Optimal VV delay

The VV delay at which the Doppler signal was maximal was identified for both the V_{max} and VTI measurements, producing an optimal VV delay for each of the 2 parameters for each patient. This was conducted 6 times so that six replicate optimisation sequences were performed in each patient.

7.4.2 Calculation of the effect of averaging replicate optimisation sequences on the scatter of identified optima

I determined the effect of averaging replicate optimisation data, ensuring that no raw measurement that was used to determine one optimum, contributed to the determination of another optimum. This ensured that the optimum VV delay identified from each replicate optimisation sequence was independent.

In each patient, I averaged the raw VTI and V_{max} measurements at each setting for each of the 6 replicate optimisation sequences, in pairs (averages of two measurements) or triplets (average of 3 measurements). The 6 raw data measurements provided three sets of "averages of two" values for each VV delay. From this data, I then identified the three optima arising, for both VTI and peak velocity. When I averaged the triplicate sets of raw data, this provided two sets of "averages of three", and from this we were then able to identify two optima for each of VTI and V_{max}, Table 7-1.

Table 7-1: Data from one patient showing all raw measurements of VTI from 6 separate optimisation sequences. This demonstrates how we averaged raw data as pairs and triplets, and calculated the between optima scatter (SD)

		VV delay setting (ms)							
		-40	-20	0	20	40	60	Optimum VV delay (ms)	Scatter of optimum (SD)
Analyzed as Singles		LVOT VTI (cm)							
Optimization sequence	#1	13.3	[15.9]	15.1	15.6	15.8	15.6	-20	23.1 ms
	#2	13.0	15.0	13.9	15.4	[16.2]	15.3	40	
	#3	13.7	14.1	[15.4]	15.3	15.2	14.1	0	
	#4	13.7	15.0	15.3	[15.9]	15.1	14.9	20	
	#5	13.0	14.8	15.3	15.2	[15.9]	15.0	40	
	#6	13.2	14.4	15.6	15.2	[16.8]	15.2	40	
Analyzed as averages of Pairs of data									
Optimization sequence	#1, #2	13.1	15.5	14.5	15.5	[16.0]	15.4	40	9.4 ms
	#3, #4	13.7	14.6	15.4	[15.6]	15.1	14.5	20	
	#5, #6	13.1	14.6	15.5	15.2	[16.3]	15.1	40	
Analyzed as averages of Triplets of data									
Optimization sequence	#1, #2, #3	13.3	15.0	14.8	15.4	[15.7]	15.0	40	0 ms
	#4, #5, #6	13.3	14.8	15.4	15.4	[15.9]	15.1	40	

Using the optimal VV delay settings identified from each of the 6 replicate sequences, we then calculated the scatter of optima using standard deviation, for both LVOT VTI and V_{max}. An example patient is shown in Figure 7-1 demonstrating the scatter of optimal VV delay obtained by VTI and V_{max}.

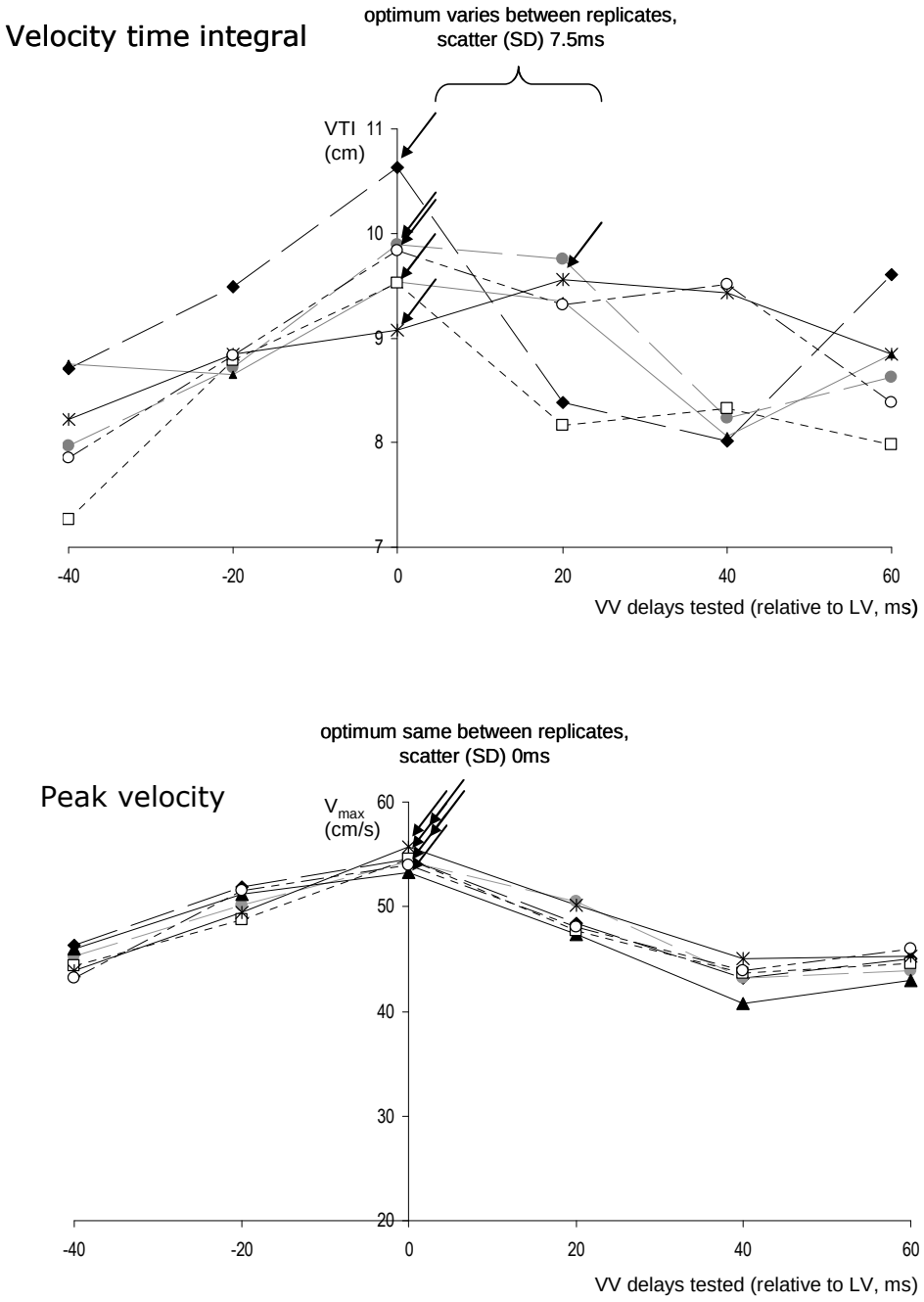


Figure 7-1: An example patient comparing the scatter of the values for optimal VV delay between velocity time integral and peak velocity.

The top panel shows the optima obtained when 6 replicate optimisations were performed using VTI and the bottom panel shows the optima obtained when 6 replicate optimisations were performed using $V_{max.}$ The scatter (SD) of the values of optima was calculated from this data to be 7.5ms for VTI and 0ms for V_{max}.

I then assessed the effect of averaging replicate optimisations on the scatter of optima, using the 6 individual values, the 3 pairs of replicate averages and the 2 groups of triplicate averages. While a narrow scatter (good reproducibility) does not guarantee that the method is good, a wide scatter (poor reproducibility) does guarantee that it is poor.

7.4.3 Calculation of Intraclass Correlation Coefficient

The intraclass correlation coefficient (ICC) quantifies the extent to which measurements differ between settings because of genuine difference between settings versus random biological variation. As described in the methodology section, the ICC quantifies how well each modality could detect a genuine improvement arising from a change in VV delay, as distinct from background spontaneous beat-to-beat variability. ICC close to 1 indicates that the changes are detected when settings are altered, whereas ICC close to 0 indicates that spontaneous variability is large in comparison to changes which occur hence true detection of change is unlikely.

7.4.4 Calculation of within-patient agreement of replicate optimisations

I calculated the within-patient agreement of optima obtained using repeated data acquisition, using Fleiss's Kappa, for VTI and for V_{max} separately. If each time an optimisation was performed the same optimum was obtained, then Kappa = 1. If the optima obtained on replicate optimisations agree only as often as expected by chance, then Kappa would be zero.

7.4.5 Calculation of between method agreement of VV optimisation

I calculated the agreement between VTI and V_{max} derived VV optimum using the absolute percentage agreement between the two methods, and using Fleiss kappa as described above.

7.4.6 Calculation of time taken to perform analysis of VTI and V_{max}.

I timed how long it took to analyse a full optimisation process. We measured the time taken to trace the LVOT velocity time integral at each VV delay setting in each patient. This was then repeated for measurements of V_{max}.

7.4.7 Statistical Analysis

Stata version 11.0 for Windows (StataCorp LP, College Station, Texas) was used for statistical analysis. A 2 x 3 mixed linear model was used to quantify the effects of using VTI versus V_{max} to guide optimisation, and of using single measurements versus averages of pairs or triplicate measurements for analysis. Fleiss kappa values were calculated for within and between patient results. Comparison of time taken using V_{max} versus VTI for analysis was performed using a paired t-test. A p value <0.05 was considered significant.

7.5 Results

In all enrolled patients, acquisition of images and measurements of LVOT V_{max} and LVOT VTI were possible for all VV delay settings.

7.5.1 Relative characteristics of VTI based optimum versus V_{max} based optimum

The scatter between repeated optima obtained was better (smaller) for V_{max}-based optimisation than for VTI-based optimisation (expressed as standard deviation, p = 0.001, Table 7-2, top panel).

V_{max} measurements were more immune to spontaneous variability than VTI measurements. The ICC, was significantly higher for V_{max} than for VTI ($p<0.0001$),Table 7-2 bottom panel.

7.5.2 Consequences of averaging replicate measurements

The scatter of replicate optimisations was improved (narrower) by moving from single measurements to averages of pairs or triplets of the raw VTI and V_{max} measurements ($p<0.0001$), as shown in Table 7-2.

For both V_{max} and VTI, spontaneous variability was smaller using averages of pairs or triplets of measurements rather than single measurements, as shown by a higher ICC ($p<0.0001$), Table 7-2.

Table 7-2 The scatter and intraclass correlation coefficient for both V_{max} and VTI

The scatter (or standard deviation) for single measures, average-of-two and average-of-three measures (across all patients) is shown for both VTI and V_{max} in the top panel. Intraclass correlation coefficient is similarly compared in the bottom panel.

Scatter (Standard Deviation, ms)

	VTI	Vmax	
Singles	18.1	15.3	
Pairs	14.9	8.2	p<0.0001
Triplicates	7.8	5.4	
	p= 0.001		

Intraclass Correlation Coefficient (ICC)

	VTI	Vmax	
Singles	0.48	0.59	
Pairs	0.55	0.68	p<0.0001
Triplicates	0.57	0.72	
	p< 0.0001		

7.5.3 Within-patient agreement of VV optima obtained on replicate optimisations

The within-patient agreement of VV optima selected with repeat optimisations was better (higher) using V_{max} than using VTI when assessed using both percentage absolute agreement and Fleiss kappa methods (absolute percentage agreement V_{max} 59.8% versus VTI 47.3%, p=0.015, Fleiss kappa V_{max} 0.51 versus VTI 0.36).

Averaging triplicate measures further improved within-patient agreement (VTI from 47.3 to 62.5%, p=0.023, V_{max} from 59.8 to 72.5%, p=0.035).

7.5.4 Between method agreement of VV optimisation

The VV optimum obtained by VTI and V_{max} agreed in 23 out the 40 subjects (absolute percentage agreement 58%, Fleiss kappa 0.52). In 76% of cases the optimum by one method was within ±20ms of the optimum by the second method, when using single optimisations, Figure 7-2.

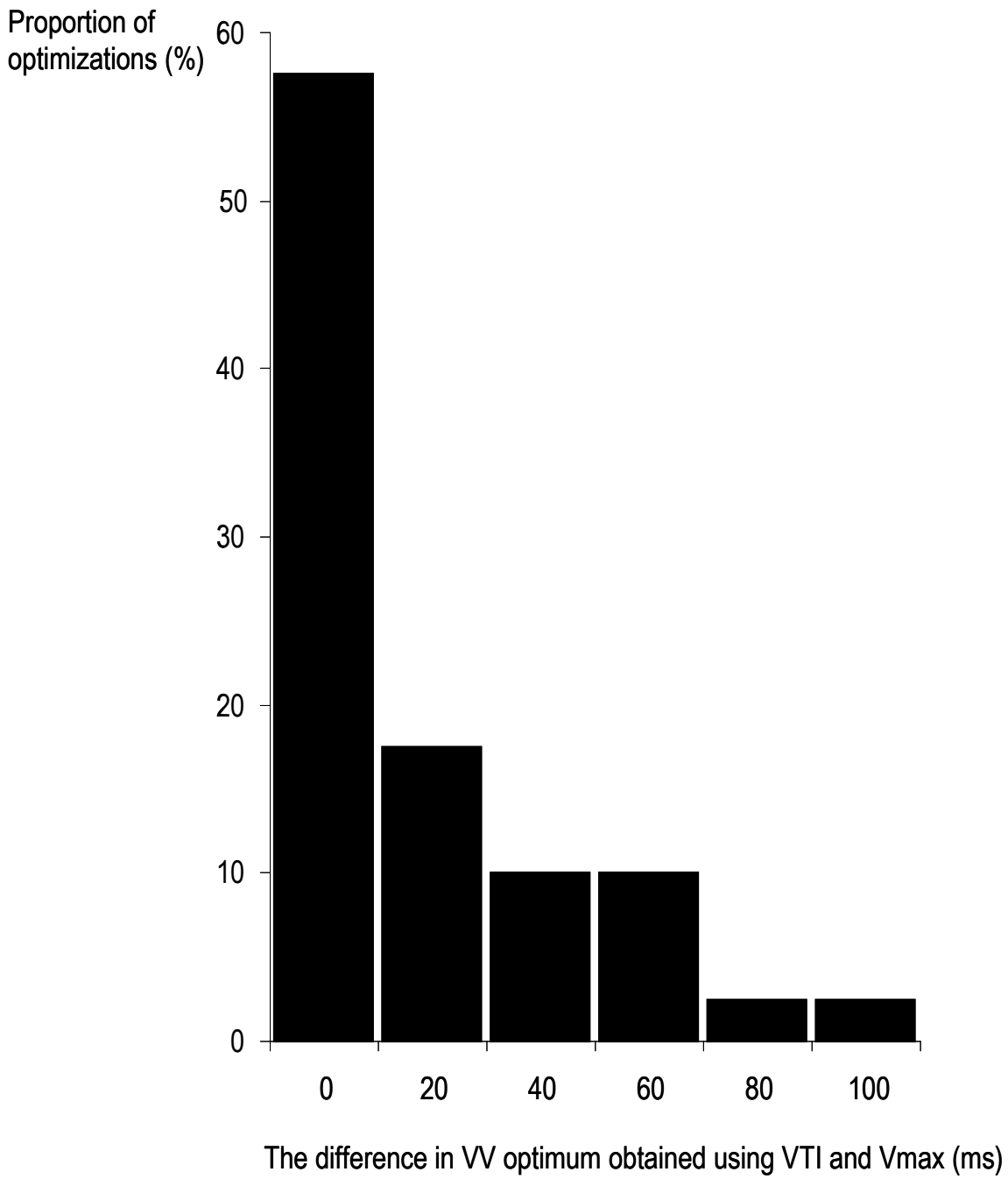


Figure 7-2 The difference in the VV optimum obtained between VTI and V_{max} across all patients. In 58% of cases the optimum was the same for both VTI and V_{max}. For 76% of cases the optimum by VTI and Vmax were within 20ms of each other. It should be noted that "± 20 ms" permits, in most cases, the second optimum to be any of 3 of the 6 tested settings.

7.5.5 Responders and non-responders

Of the 40 subjects, 28/ 40 (70%) were clinical responders. In the responder group the ICC was greater when measured using V_{max} compared to VTI, 0.65 vs 0.55, p=0.021 across 28 patients. In the non-responder group of 12 patients, there was no significant difference although a trend to a greater ICC in the V_{max} measurements vs. VTI measurements, 0.65 vs 0.51 respectively, p=0.20.

The scatter of the VV optimum obtained showed a similar pattern. The scatter was smaller using V_{max} compared to using VTI in the responder group, 14.0 ms vs. 18.8 ms (p=0.028). In the non-responder group, there was no significant difference between Vmax and VTI in the scatter of VV optima obtained 17.6ms vs 16.4ms (p=0.72).

The smaller number of patients who were clinical non-responders (12 rather than 28) is the likely reason for the lack of evidence of a statistically significant difference between responders and non responder groups.

Time taken to perform optimisation analysis

It was more than 3 times faster to perform the analysis of a full optimisation protocol using V_{max} than VTI, (17.5 vs. 59.1 seconds, $p<0.0001$), Figure 7-3.

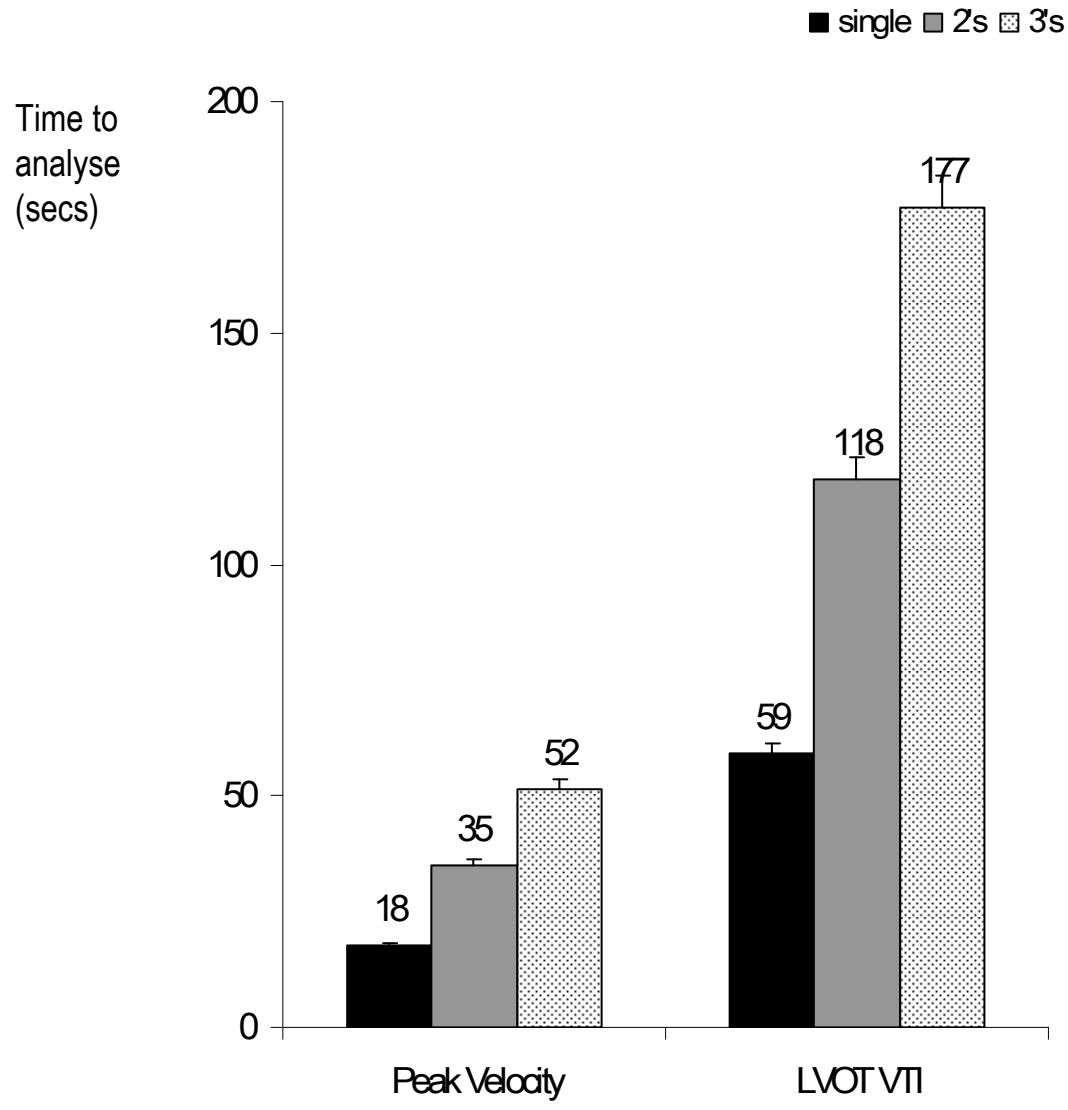


Figure 7-3 Time taken to analyse single measures, 2's and 3's as used for the single, averages-of-two and averages-of-three calculations.

7.6 Discussion

These data highlight the characteristics, both favourable and challenging, of using Doppler echocardiography to guide the optimisation of VV delay in CRT devices. The inherent variability in Doppler measurements between beats must be taken into consideration when assessing its suitability for use in optimisation protocols. If assumed to be negligible, this variability ("noise") can easily swamp the "signal", which often changes only subtly between different settings, making consistent identification of the optimal VV delay problematic (Pabari et al. 2011)

This study shows that any attempt at optimising VV delay using LVOT Doppler must be accompanied by verification that the scatter between replicate measurements is either small (which is unlikely) or quenched by sufficient replication and averaging, to make the optimisation process credible.

The head-to-head comparative data indicate that using V_{max} to guide VV optimisation takes less time and gives a more reproducible optimum, and therefore would seem to be unambiguously preferable to using VTI.

These findings have implications for the interpretation of studies where an optimisation protocol was trialled for effects on clinical endpoints (Ellenbogen et al. 2010) but for which precision of the optimisation procedure has not been evaluated (or has not been published). The necessary evaluation is test and retest (i.e. separate data, not re-evaluation of the same JPEGs) reproducibility, with observers blinded on the second occasion to the results obtained from the first.

7.6.1 How can an incomplete picture be better?

Recommending measurement of peak velocity rather than VTI for the purpose of optimisation seems to run counter to many axioms in echocardiography, and I do not do this lightly.

First, VTI gives a more complete picture, because it captures the profile of flow throughout ejection, whereas peak velocity only assesses a single instant. Second, VTI allows calculation of stroke distance (Colocousis et al. 1977), and therefore stroke volume and cardiac output, which are physiologically and clinically valuable to know.

I do not consider these considerations incorrect, nor unimportant, but rather that they are less important than delivering the fundamental requirement without which optimisation is doomed to fail before it begins: that is, to be reliable in detecting the very small changes in cardiac function occurring with alterations in VV delay. For this very specific requirement, capturing a more complete picture does not help if that bigger picture contains more additional noise than it does additional signal. Nor is it clinically essential to establish estimates of the stroke distance, stroke volume or cardiac output for each tested setting unless those measures have the best signal-to-noise ratio. If knowledge of these variables is desired, they can be assessed at the optimised setting, i.e. after (rather than as part of) the optimisation process. I would argue that a less comprehensive picture can be better for the purpose of optimisation, provided it has a higher signal to noise ratio.

7.7 The importance of assessing signal-to-noise ratio when planning optimisation algorithms

Echocardiography is attractive to use as a tool to guide optimisation because it is non-invasive and widely available with no additional operator training needed. Nevertheless the choice of what technology to use, or which variable within that technology is most suited should not be based on such superficial concerns but rather on a sound basis.

In principle, the first step in assessing the suitability of a proposed variable for guiding optimisation would be to establish its intraclass correlation coefficient while VV delay is adjusted. If the ICC is close to 1, it is a plausible candidate variable to take forward for further testing (for example, of the reproducibility of the optimum obtained at separate visits). If the ICC is ~0.5 or lower, it is not a plausible candidate (Pabari et al. 2011). This initial screening test would just need a few (for example 6) replicate measurements at each of the range of settings intended to be used clinically (for example 6 settings). These 36 measurements need not take more than half an hour to conduct per patient, unless the method is unusually complex (in which case it is unlikely to be ready for routine use) and need only be done for a dozen or so patients.

This elementary step is a prerequisite before issuing a recommendation of how to conduct clinical optimisation.

7.7.1 Averaging multiple measurements

Conducting multiple replicates is an important step for reducing the impact of unwanted noise (Francis et al. 1999) and hence maximising the precision of any measurement technique, but it is not always practical to do this under the time constraints of clinical practice. Most research protocols involve multiple replicates to be made, although many study reports do not state how many replicates have been averaged or if simply the best-looking recording of the replicates at each setting has been selected as representative.

If using LVOT Doppler signals to guide optimisation, the present study shows that using V_{max} allows more than 3 times more optimisations to be analysed within the same timeframe as analysing a single optimisation using VTI. In addition, V_{max} is more reliably sensitive to the genuine changes between settings.

7.7.2 How to have confidence in the method of optimisation

The natural variability (or noise) in the Doppler signal is likely to make identification of the true underlying VV optimum difficult. The critical information is the test-retest reproducibility of the VV optimum (from data acquired in separate visits) of any method recommended in a guideline, or carried out in an endpoint trial? The present analysis, and previous studies (Turcott et al. 2010; Whinnett et al. 2006a) suggests that the reproducibility is poor using guideline-recommended methods, and therefore may have been poor in trials (unless those trials used some undisclosed steps to reduce noise). Other centres have had greater success with better reproducibility, but the conclusion that must be drawn is that reproducibility of this technique varies either between operators and / or centres (Thomas et al. 2009; Turcott et al. 2010) . If this is the case, rather than rejecting the concept of VV optimisation altogether if the trials do not show endpoint benefits, we might better put efforts into developing

methods for optimisation that are confirmed, in independent hands, to be reproducible between visits, with appropriately blinded operators.

Such studies are very much cheaper and quicker to do, and are an essential prelude to longer-term outcome studies. That they have not been published is a surprise.

7.8 Limitations

This study did not set out to examine, or predict, response to CRT implantation. It did not set out to measure long-term outcomes at each optimised setting, because there is no point trying to do so until we have a technique that recommends a single optimum consistently for one patient (Ellenbogen et al. 2010; Pabari et al. 2011). For measuring VTI I took the standard approach which involves tracing around the outer perimeter of the VTI envelope (Bertini et al. 2010; Tribouilloy et al. 1994). There are other methods using automated algorithms which might potentially improve the precision obtainable. These however are not generally available or used in the standard clinical practice setting. Since I was addressing the practical question of how the precision of optimisation could be improved above guideline-recommended methods, in a real-world clinical setting without using any new advanced modalities, I adopted the method conventionally used during clinical optimisation.

This study does not consider more modern echocardiographic measurements than spectral Doppler, for example tissue Doppler or strain (Bertini et al. 2010) which may require additional training for both acquisition and analysis. All echocardiographers are familiar with measuring spectral Doppler peak velocity from routine assessment especially of valves (Minners et al. 2008). However, the approach taken in this study might also be taken for any number of other potential markers, and would allow

markers to be screened for plausibility before undertaking prolonged studies (Pabari et al. 2011).

I should highlight again that I am not disputing the important role for VTI measurement in clinical practice, as this permits evaluation of stroke volume and thereby cardiac output. Rather, I am just addressing the specific question of which measurement gives the best reliability to detect the small changes occurring when VV delay is changed, which is the key requirement to reliably identify an optimum.

7.9 Conclusion

Whether VV optimisation of CRT devices is worthwhile or not has not been settled. If it is not, i.e. the programmed VV delay is not important, then the term "resynchronisation" is no longer tenable. If, on the other hand, VV delay does matter, then it is essential for optimisation methods to give reproducible results.

Following the guideline protocol, which is to use LVOT VTI to guide VV delay optimisation, gives an unreliable optimum. LVOT V_{max}, obtained in a similar manner, is 3 times quicker to analyse and yields an optimum more reproducible. This advantage of using peak velocity for optimisation is seen even more clearly when multiple measurements are averaged, making the optimisations more precise.

8 Invasive pressure and flow measurements during AV delay improvement reveal a compensatory peripheral vasodilator response which attenuates the initial blood pressure increment: implications for the design of AV optimisation protocols

8.1 Abstract

8.1.1 Introduction

With synchrony of ventricular contraction already restored by cardiac resynchronisation therapy (CRT), optimisation of atrioventricular (AV) delay relies on improving filling. Although when AV delay is improved blood pressure immediately rises, there is a subsequent partial decline. Is this secondary decline because (1) non-invasive measurements are unreliable, (2) cardiac function increment is short-lived or (3) peripheral vasodilatation occurs? I conducted invasive experiments to answer this.

8.1.2 Methods and Results

Nine patients with heart failure, CRT and underlying left bundle branch block, underwent changes in AV delay from 40 to 120 ms. I simultaneously measured beat-by-beat invasive aortic pressure and flow, and non-invasive pressure (Finometer). Triplicate experiment runs were performed and averaged to minimise biological noise.

There was an immediate increment in invasively measured blood pressure of +14.7±2.0mmHg (p=0.0001), but after the initial 10 beats there was a secondary progressive decline to a lower plateau of +8.0±1.8mmHg (p=0.004). The initial increment was caused by a prompt rise in flow by +9.1±2.4% (p=0.007) which did not drop later. The secondary fall in pressure was caused by a delayed gradual decline in total peripheral resistance. Finometer-derived non-invasive blood pressure tracked invasive pressure closely (r=0.97).

8.1.3 Conclusion

When AV delay is made more favourable, only the instant pressure increment is caused by increase in stroke volume. The secondary pressure decline is caused by systemic vasodilatation. Non-invasive blood pressure changes in parallel with invasive pressure when AV delay is altered, therefore may be a convenient, accessible target for CRT optimisation. Design of AV optimisation protocols, which face severe challenge of signal versus noise, might benefit from recognition that not all beats are equally informative: the first few after a transition are most signal-rich.

8.2 Background

Following the work carried out using echocardiography, I was not entirely satisfied with the signal-to-noise ratio of the echo parameters. Even Doppler echocardiography has a relatively unsatisfactory signal-to-noise ratio, therefore I looked at ways to reduce signal further, improve the signal when intra cardiac timings are changed and hence eliminate the hindering factors which affect echocardiography as a modality in understanding changes that occur after timings are altered.

This lead to the next study which I performed; measuring invasive flow and pressure simultaneously on a beat-by-beat basis when intra cardiac timings are changed from one setting to another. To improve the signal-to-noise ratio I used AV delay settings of 120ms to 40 ms and this provides greater change in hemodynamic measure hence a greater signal is visible (Pabari et al. 2011; Whinnett et al. 2006b), and performed replicate measurements as described below to minimise the noise (Pabari et al. 2011).

Over 100,000 times per year in USA and Europe alone (Cunningham D et al. 2011), a CRT device is implanted but after each procedure the clinician is left with an undesirable conflict between theory and practice. On one hand, carefully conducted randomised controlled clinical trials have demonstrated that qualitative optimisation (by visual inspection of transmitral Doppler) does not significantly improve outcomes compared with a standard setting of 120 ms AV delay (Ellenbogen et al. 2010). Yet on the other hand the sole theoretical reason for implanting these devices is to improve intra cardiac timings in order to maximise cardiac function and improve symptoms. It would be inconsistent with the stated purpose of the devices if, in patients with appropriate biological targets (LBBB and wide QRS), AV delays ranging as widely as 31ms to 300 ms (Cleland et al. 2001; Ellenbogen et al. 2010;

Whinnett et al. 2006b) and VV delays from RV-first 40ms to LV-first 40 ms, made no difference to the physiological effect of the device (Inoue et al. 2005b).

One potential explanation for this apparent discrepancy is that many of these studies used the iterative method for AV delay optimisation – a qualitative technique that employs visual inspection of Doppler traces to detect the optimum setting. This has imperfect reproducibility (Raphael CE. et al. 2011; Turcott et al. 2010) and may not have identified the optimal setting for individual patients with high fidelity.

Quantitative optimisation is the alternative which should now be explored. It requires making measurements at each setting and establishing their uncertainty, so that it can be confirmed that the setting which appears to be the optimal is not simply the highest by chance (Pabari et al. 2011). Quantitative pressure measurements, made invasively at the time of implant, consistently show a clear maximum within individual subjects (Auricchio et al. 1999a; Perego et al. 2003) and therefore pass the first hurdle for an optimisation process to be considered valid. Historically, AV delay has been optimised using assessments in diastole such as transmitral filling pattern, whereas the primary target of LV pacing is considered to be systolic mechanical dyssynchrony.

There are a variety of approaches to develop a non-invasive quantitative process suitable for use in an outpatient setting. These include quantitative non-invasive beat-by-beat pressure monitoring (van Geldorp et al. 2011; Whinnett et al. 2006a), and quantitative echocardiographic measurements of velocity time integral (Jansen et al. 2006b; Thomas et al. 2009) (i.e. stroke volume). However, there are 2 important questions.

First, the immediate primary change in pressure after changing the AV delay is followed by a secondary decline after a few seconds (Whinnett et al. 2011). One

explanation for this is that the increased duration of filling for a single beat at the time of change in AV delay produces a transient increase in LV volume that allows a transient increase in stroke volume. If that is the case, the increment seen during such protocols may be largely reflecting transient phenomena which are not likely to affect long term physiology, and meanwhile any genuine persistent physiological benefits of the change in AV delay may be difficult to detect in the presence of this irrelevant transient increment. Alternatively, the decline may represent potentially-beneficial homeostatic responses i.e. partial relaxation of the vasoconstriction characteristic of heart failure.

Second, although non-invasive blood pressure may well be a practical mode of optimisation, it is not clear if non-invasive measurements adequately track invasive measurements in the context of pacemaker optimisation with sufficient fidelity.

8.3 Methods

8.3.1 Protocol design elements for reliability of results

8.3.1.1 Invasive acquisition.

I examined carefully the temporal pattern of concurrent changes of aortic pressure and flow when a pacemaker setting is changed. I used well-established invasive methods that deliver beat-to-beat measurements. Simultaneously I used non-invasive beat-to-beat blood pressure measurements to assess their fidelity in tracking the invasive blood pressure measurements, and therefore their validity for measuring the blood pressure response during AV delay optimisation.

8.3.1.2 Signal and noise.

Measurement imprecision and physiological variability frequently obscure the underlying true, albeit often small, changes during a clinical optimisation procedure and therefore we took special steps to maximise the signal (Whinnett et al. 2011) (by guaranteeing a large haemodynamic effect) and to minimise noise (Pabari et al. 2011) (by performing multiple replicates in each patient which could then be averaged).

8.3.1.3 Minimising noise: replication and selection of variables.

The purpose of this experiment was to establish invasively and with temporal precision, the time course of the sequence of changes in aortic pressure and flow after a change in AV delay of a CRT device. I did not want to study the effect of resynchronisation itself, but only that of a change in AV delay, and therefore I compared states that were biventricularly paced. The reason no previous study has reported on this time course is that the changes are subtle in comparison to spontaneous beat-by-beat biological variability. Precision is increased (i.e. standard error of the mean is reduced) by using a measure with a small spontaneous biological

variability and by averaging multiple replicate experimental runs. I chose to track changes in stroke volume using Doppler flow velocity because using standard techniques its spontaneous variability is ~5% in comparison to ~11% for thermodilution (Lehmann & Platt 1999); and because the Doppler provides beat-by-beat information which is necessary to examine this time course.

8.3.2 Maximising signal: only two AV delays.

I chose two AV delay settings that, despite being predefined, would consistently provide contrasting cardiac outputs. One was 40 ms which would cause truncation of the A wave and therefore poor cardiac output. The other was 120 ms which was designed to be different by enough to provide substantially better physiology and yet not so much longer as to risk fusion in any of the patients. I did not make measurements in any modes without biventricular pacing because I wanted to lavish all the available experimental time on answering as precisely as possible the question of AV delay alteration within the range during which full biventricular pacing was definitely in place.

I did not test any further AV delays (and specifically did not explore the “good” region closely) because increasing the number of AV delays by N-fold would increase the acquisition time for the patient not just N-fold but approximately N 5-fold. The reason is that the shape of haemodynamic response is curved approximately parabolically (Whinnett et al. 2006b) near the optimum, so an AV delay N times closer to the optimum would have haemodynamic effects ~N 2 times closer to the optimal, and would therefore by the Central Limit Theorem require ~N 4-fold more replicate measurements per AV delay to resolve the difference to the same level of certainty. Therefore, for example, doubling of the number of AV delays tested

increases the number of measurements required per patient to achieve the same resolution between settings by ~25=32-fold.

8.4 Study population

Patients eligible for CRT implantation according to published guidelines (Barnett et al. 2007) (NYHA III or IV heart failure, QRS> 150ms, EF < 40% and on maximal medical therapy) and required coronary angiography as part of their standard clinical investigations were recruited from Imperial College NHS trust cardiology outpatients. Exclusion criteria were significant coronary stenosis requiring revascularisation, or any degree of aortic stenosis or aortic regurgitation assessed as being greater than mild, atrial fibrillation or incomplete ventricular capture when pacing. Patients gave informed consent for this study which was approved by the local ethical committee.

8.5 Measurements

8.5.1 (a) Invasive measurements

Aortic pressure and flow were respectively assessed using a fluid filled catheter and Doppler flow wire (Volcano FloWire 1400) positioned in the aorta approximately 5 cm from the aortic valve. The experiment was designed to use standard Doppler flow velocity measurements rather than standard thermodilution measurements for two reasons. First, the coefficient of variation between measurements was 2-fold smaller (5.6% versus 11.0%(Lehmann & Platt 1999) respectively) and therefore the changes in flow could be detected with the same precision with 4 times fewer repetitions of the protocol. Second, Doppler flow velocity measurements provide data with very high temporal precision in comparison with thermodilution, and thereby permit the time course of effect to be more precisely evaluated.

8.5.2 (b) Non-invasive blood pressure measurements

Beat-to-beat systolic and diastolic blood pressures (SBP and DBP) were also measured non-invasively simultaneously using a photoplethysmograph device (Finometer, Finometer Medical Systems, Netherlands). This system uses a cuff placed around the finger with a built-in photo-electric plethysmograph and volume-clamp circuit which dynamically follows arterial pressure, to yield the beat-to-beat arterial pressure waveform (Imholz et al. 1998; Penzel et al. 1992; Smith et al. 1985; van et al. 1985).

8.5.3 Pacing protocol

During coronary angiography, temporary biventricular pacing was performed in eligible subjects. A quadripolar electrode catheter (Josephson Curve, Bard Viking) was placed in the right atrium (usually the right atrial appendage) and a pentapole electrode catheter was placed at the right ventricular apex. An AL1 and/or a channel sheath was used to gain access to the coronary sinus and an ATW wire was positioned in a lateral or posterior-lateral coronary sinus branch for LV temporary pacing(Lane et al. 2008). LV capture was verified using a 12 lead ECG.

A CRT pacemaker (Medtronic InSynch III 8042) was connected extracorporally and transitions made via a standard pacemaker programmer. DDD pacing mode was used with heart rate was fixed at 100bpm in all patients by atrially pacing. Stable pacing and sensing for all 3 pacing wires was monitored throughout the protocol in all patients. By lengthening the paced AV delay progressively, we checked in each patient that the AV delay at which the right ventricular lead began to sense or the QRS began to change shape was well beyond 120 ms (range 240 ms to 320 ms). This ensured full biventricular pacing at paced AV 120 ms.

The protocol consisted of a series of transitions from AV delay 40ms to 120 ms. Measurements were recorded for at least 100 beats before and 100 beats after the change in AV delay. To minimise the effect of random variation, triplicate runs of the experimental protocol were conducted so that each patient's dataset was composed of an average of 3 runs, aligned by beat and registered to the time point of transition from 40 ms to 120 ms.

8.5.4 Total Peripheral Resistance

To detect changes in total peripheral resistance (TPR) we used the formula:

$$\text{Total Peripheral Resistance} = \left(\frac{\text{Mean Arterial Pressure}}{\text{Stroke distance} \times \text{LVOT cross sectional area} \times \text{heart rate}} \right)$$

Because LVOT cross sectional area and heart rate remained constant, fractional changes in total peripheral resistance could be calculated as:

$$\text{Fractional change in TPR} = \text{Fractional change in} \left(\frac{\text{Mean Arterial Pressure}}{\text{Stroke volume}} \right)$$

8.5.5 Data Acquisition and Analysis

Simultaneous recordings of invasive flow, aortic invasive blood pressure and non-invasive blood pressure were measured on a beat-by-beat basis throughout the protocol.

Data were acquired using a NIDAQ AI-16E-4 analogue-to-digital card (National Instruments, Austin, TX) and Labview (National Instruments, Austin, TX). We analysed data using custom software, based on the Matlab platform (MathWorks, Natick, MA)(Davies et al. 1999) .

8.5.6 Statistics

Values are presented as mean ± sem, unless otherwise stated. Paired comparisons of continuous variables were made using Student's paired t test. A p-value of <0.05 was taken as statistically significant. Statview 5.0 (SAS Institute Inc., Cary, NC) was used for statistical analysis.

8.6 Results

8.6.1 Patient characteristics

A total of 9 subjects (8 male, mean age 63 years, range 42 to 80 years underwent this study. At the time of study, 7 were NYHA class III and 2 were NYHA IV. The cause of heart failure was ischemic in 3 and idiopathic dilated in 6. All patients had left bundle branch block on their native ECG with a QRS duration mean of 175 ms and standard deviation of 16 ms. All patients were taking angiotensin-converting enzyme inhibitors or angiotensin-II receptor antagonists, 6 were taking beta-blockers, and 8 were taking a diuretic (loop or thiazide). 8 out of 9 patients had an echocardiogram at our institution. Non-invasive assessment of cardiac output by the LVOT integral method was 3.8 l/min ± 0.24 l/min at rest pre implantation.

8.7 Phases of Response to change in AV delay from 40 to 120 ms.

We observed a multiphasic pattern of response in the invasive measurements acquired during the change in AV delay from 40-120 ms, in all subjects. There were 4 phases to the response including 2 plateau phases. First, the initial beat had an unnaturally long diastolic period preceding it and therefore had a large stroke volume Figure 8-2.

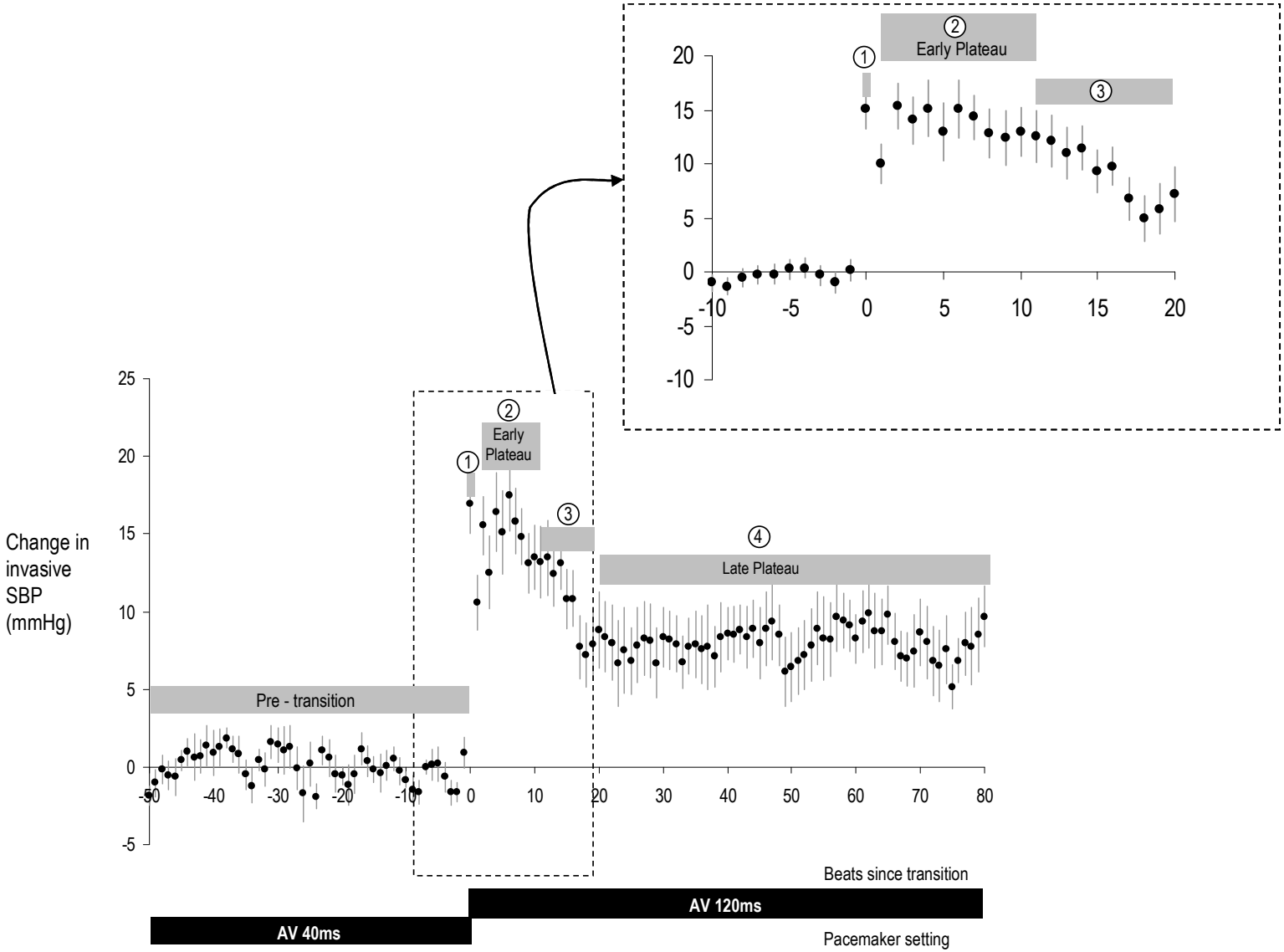


Figure 8-1 Beat-to-beat changes in invasive blood pressure when AV changed from 40 ms to 120 ms across all patients. After the transition the four phases are shown: (1) the first beat, (2) the early plateau (beats 2-11), (3) a progressive decline and (4) the late plateau (beats 20-80).

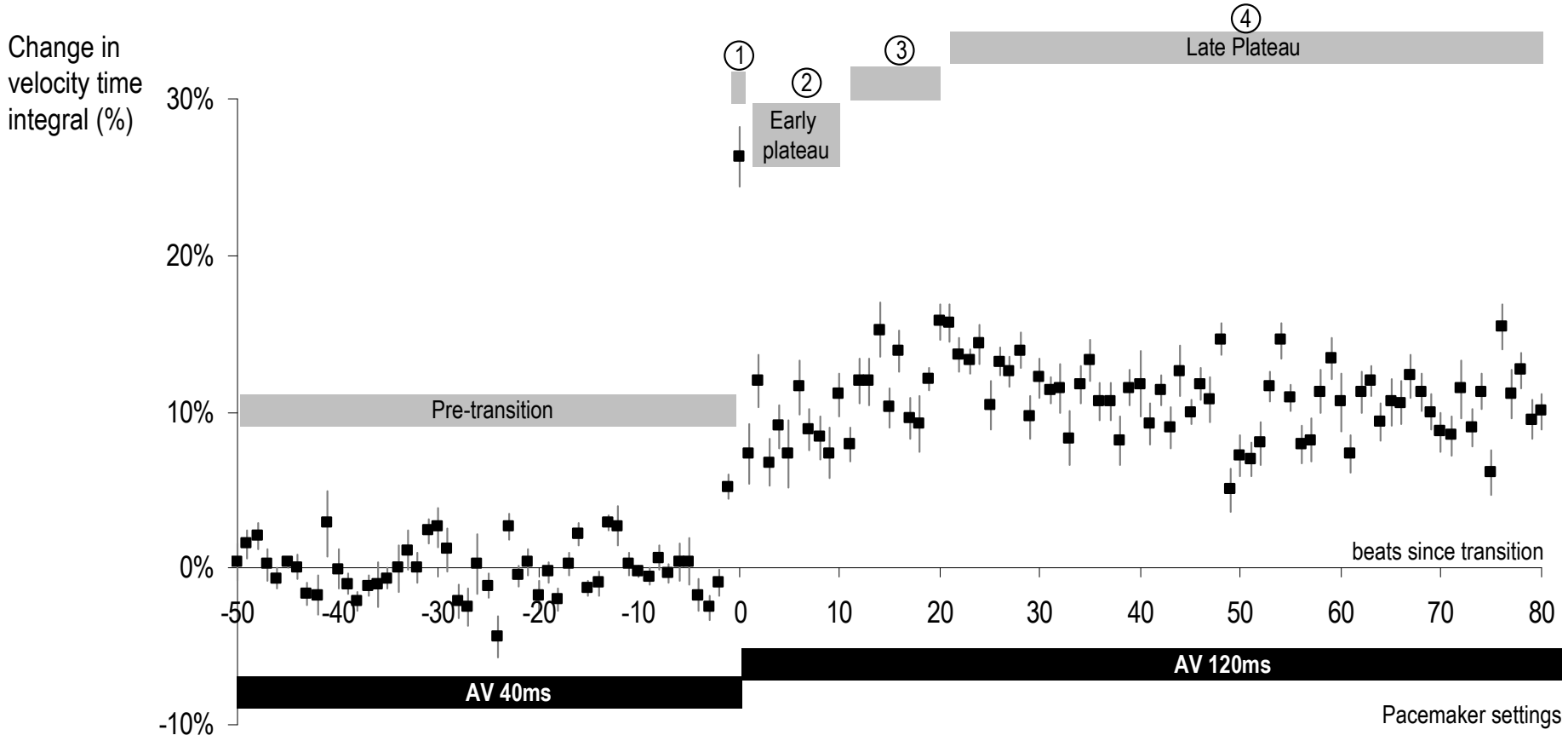


Figure 8-2 Change in invasive flow, shown beat-by-beat, as AV delay is changed from 40 to 120ms across all patients.
The early plateau (beats 2-11) and late plateau (beats 20-80) are clearly demarcated with a significant increase between AV 40 ms and the early plateau and a further increase between the early plateau and late plateau.

Second, there was a period of about 10 beats (beats 2 – 11) with a substantially elevated blood pressure (Figure 8-1) and somewhat elevated flow (Figure 8-2). Over the next approximately 10 beats there were further changes, and then a new plateau was evident as a fourth phase from beat 20 onwards.

The invasive pressure early plateau, defined as beats 2 to 11, was elevated at 14.7 mmHg ± 2.0 mmHg, (p=0.0001) above baseline. The late plateau (phase 4) (from beat 20 onwards) was 8.0 mmHg ± 1.8 mmHg above baseline (p=0.003). This was significantly lower than the early plateau (p=0.004), Figure 8-1.

Invasive flow velocity integral had an early plateau of 9.1% ± 2.4% (p=0.007) above baseline. After this it did not show a fall but in fact was maintained at this elevated level with a maximum in the late phase of 10.3% ± 1.6% (p=0.004) above baseline.

8.7.1 Comparison between invasive and non-invasive blood pressure measurements following the change in AV delay from 40-120ms

Non-invasive blood pressure showed an identical pattern of response to that of invasive pressure, as shown in Figure 8-3. The blood pressure early and late plateaus from non-invasive measurements were 14.3 mmHg ± 2.3 mmHg and 8.2 mmHg ± 3.2 mmHg respectively above baseline. These are similar to the invasively measured increments, Figure 8-3. The beat-by-beat data for non-invasive and aortic pressure showed parallel behaviour (Figure 8-4, r=0.97).

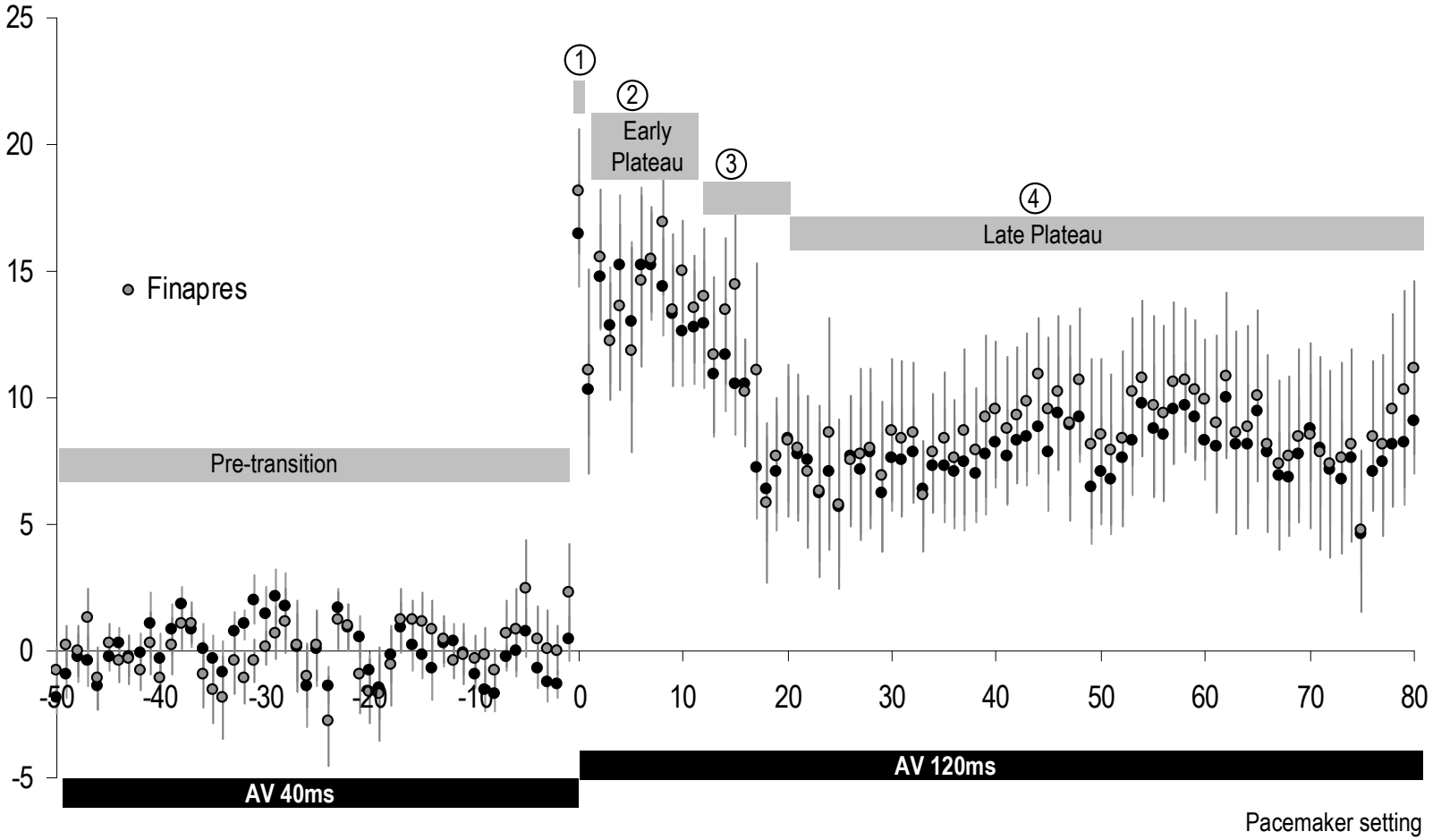


Figure 8-3 Beat-to-beat data for invasive and Finometer (non-invasive) systolic blood pressure across all patients during a programmed change in AV delay from 40 to120 ms, averaged across all subjects.

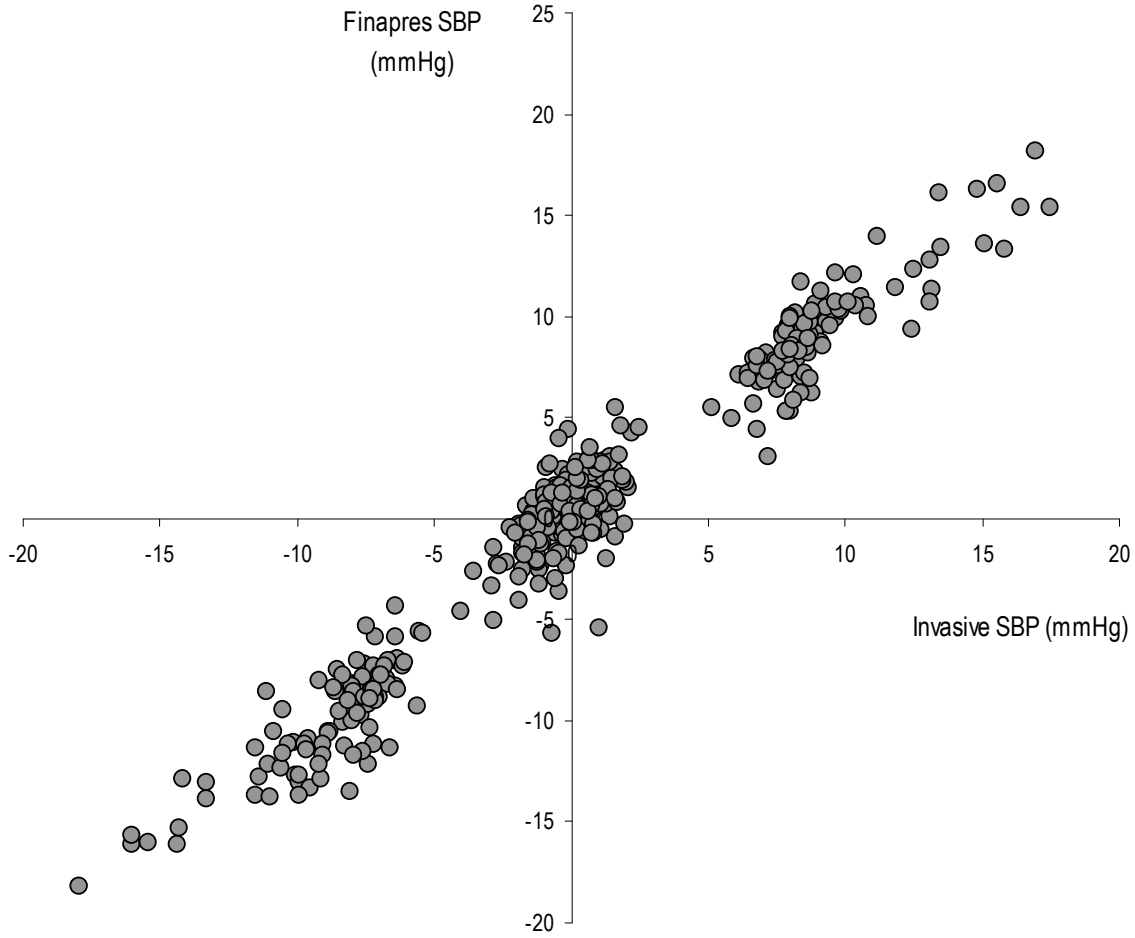


Figure 8-4 Relationship between change in beat-to-beat invasive SBP and Finometer SBP.
Data is averaged across all subjects between transitions AV 40ms to 120 ms, and AV 120 ms to 40 ms, from Figure 3 (r=0.97).

8.7.2 Reverse transitions: 120 ms to 40 ms

In a post-hoc analysis we examined the data that had been acquired during reverse transition, i.e. from 120 ms to 40 ms. There is a drop in invasive blood pressure to −14.4 mmHg ±2.3 mmHg below baseline (p=0.0004) in the period from beats 2 to 11 (Figure 8-5), and a parallel drop in flow VTI to −11.2% ± 1.6 % (p=0.0001), Figure 8-6. Pressure the rises a little (p=0.0006) so that in the late phase of 20 beats onwards it is −7.8mmHg relative to baseline (p=0.002). This partial recovery is not caused by a rise in flow VTI, which is not significantly different in the late phase and remains clearly lower than before the transition (p=0.00003, Figure 8-7).

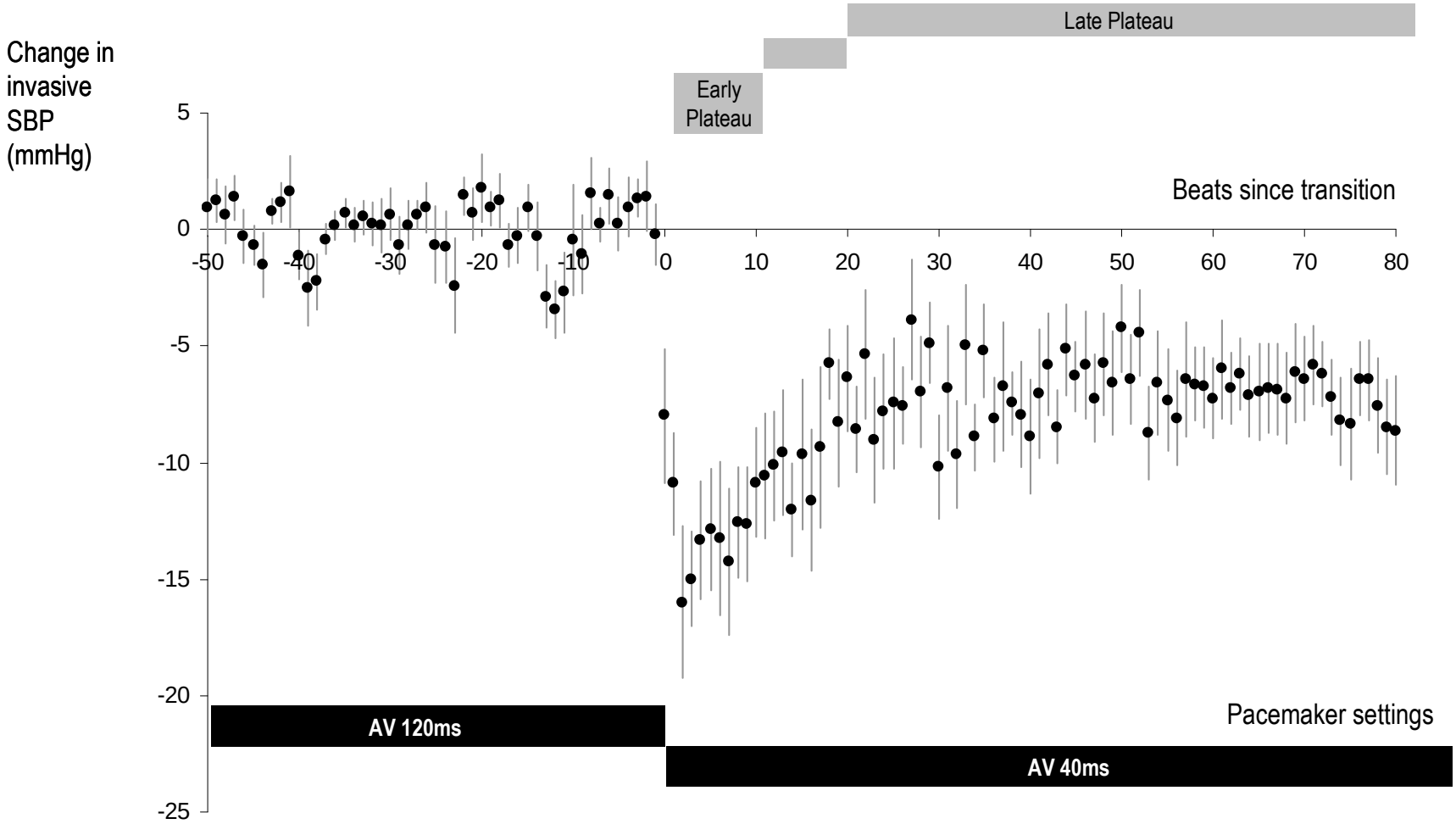


Figure 8-5 Beat-to-beat changes in invasive blood pressure when AV changed from 120 to 40 ms across all patients. There is a drop in invasive blood pressure to −14.4 mmHg ±2.3 mmHg below baseline (p=0.0004) in the period from beats 2 to 11

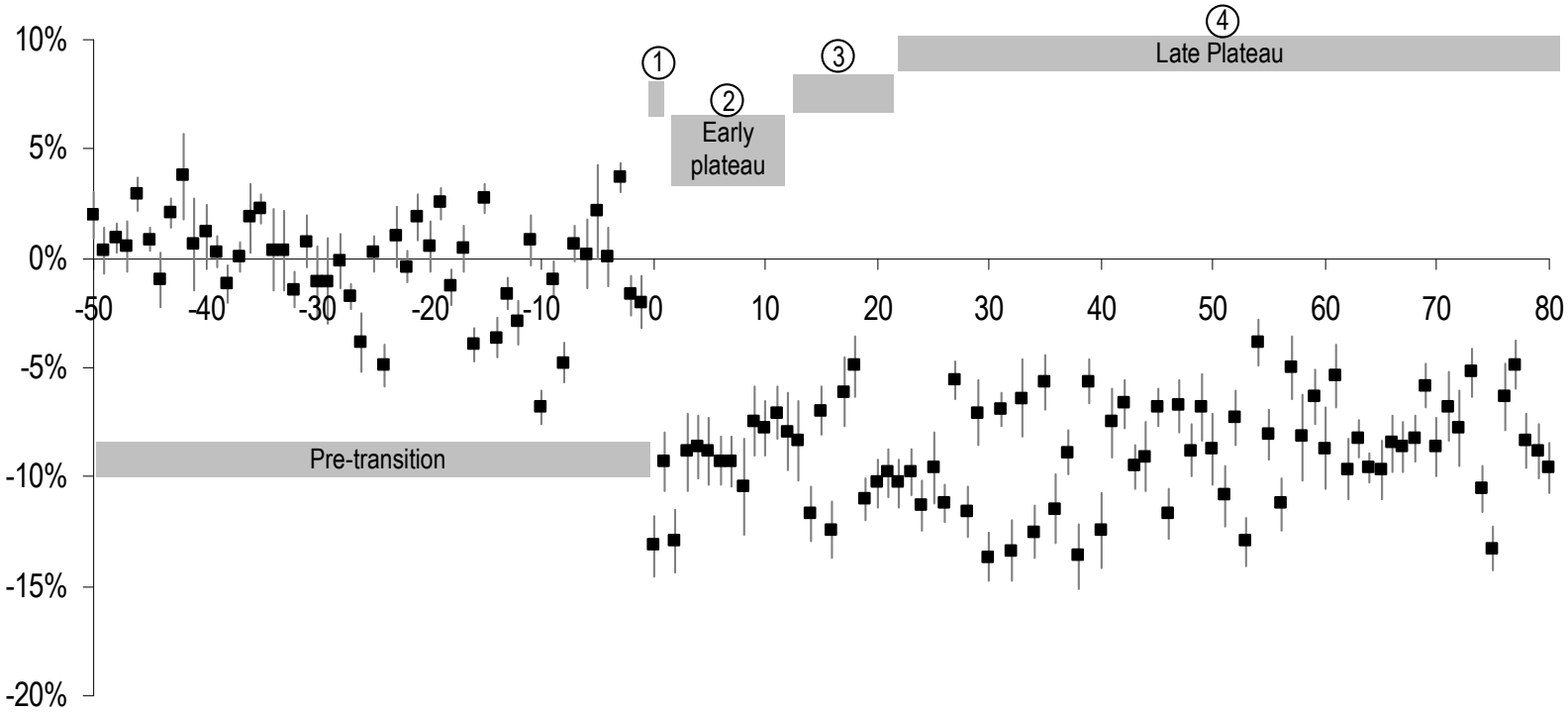


Figure 8-6 Change in invasive flow, shown beat-by-beat, as AV delay is changed from 120 to 40ms across all patients.
The early plateau (beats 2-11) and late plateau (beats 20-80) are clearly demarcated with a decrease in flow between AV 40 ms and the early plateau. Of note there is a lack of the prominent first beat effect seen when the transition occurs from 40 to 120ms.

8.7.3 Mechanisms: Delayed change in Total Peripheral Resistance

From the pre-transition state, total peripheral resistance (TPR) is no different in the early plateau phase but then falls during the late plateau. This is shown in Figure 8-7 demonstrating the effect on TPR with the transition from 40ms to 120 ms, and the reverse effect when transition made from AV 120 ms to 40 ms.

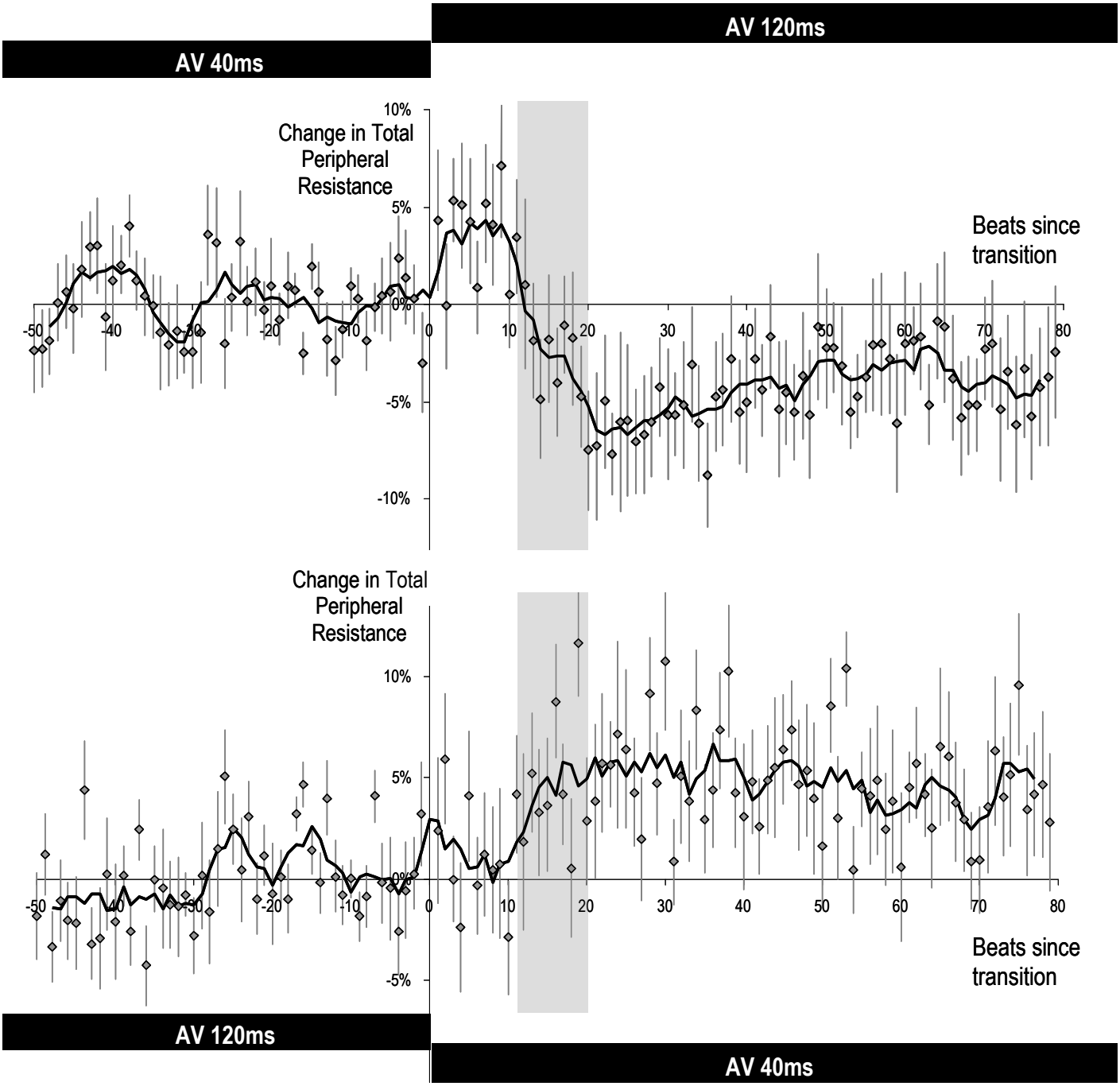


Figure 8-7 Beat-to-beat changes in Total Peripheral Resistance across all patients for change in AV delay from 40ms to 120ms (top panel) and 120ms to 40ms (bottom panel).

There is a delay of a few beats after which there is a progressive change in total peripheral resistance between AV delay 40 ms and AV delay 120 ms, and rise in resistance between 120ms 40ms which is highlighted by the grey shaded area. It then change then stabilises over subsequent beats at that level reached after 20 beats. A 5-beat moving average line is also shown.

8.8 Discussion

This study employs an experimental methodology with a deliberately high signal-to-noise ratio to permit temporal resolution of the changes in blood pressure and aortic flow following changes in AV delay. It does not address the mechanism of action of cardiac resynchronisation itself, since all patients had full CRT pacing throughout, but rather focuses only on the effect of changes in AV timing on downstream physiology in the systemic arteries. It shows an immediate increase in aortic pressure and flow forming an early plateau following transition from AV 40 ms to AV 120 ms. A few seconds later there is a further increment in flow accompanied by a fall in pressure, which is likely to represent secondary systemic vasodilatation. Beat-by-beat non-invasive blood pressure closely tracks changes in invasive aortic blood pressure during this manipulation confirming the validity of this approach for blood pressure monitoring during haemodynamic optimisation of CRT. These observations may be helpful in designing CRT optimisation protocols.

8.8.1 Relevance to quantitative optimisation of CRT

CRT exerts its benefits by changing intra cardiac timings. Even small improvements in cardiac function if sustained can be worthwhile (Cleland et al. 2005), but measurement imprecision and spontaneous physiological variability within individuals makes optimisation difficult (Francis 2011; Pabari et al. 2011). Averaging multiple replicate measurements to improve precision (Le Cam L 1986) is laborious if not automated. Protocols are typically not designed and verified to deliver precision of optimisation, and so even carefully-conducted endpoint research (Cleland et al. 2005; Ellenbogen et al. 2010; Pabari et al. 2011) may not be definitively testing the clinically important hypothesis that precise AV optimisation is helpful.

Non-invasive measuring of blood pressure is not uncomfortable and might be fully automated. Nevertheless, it faces two challenges. First, there are two classes of consequences of improving AV delay: sustained improvement of cardiac function, and a mere transient "bump" caused by a single cardiac cycle with a longer diastole. Only sustained consequences matter. The initial blood pressure increment with improvement in AV delay is followed by a partial decline after a few beats (Whinnett et al. 2011). If the secondary decline is of both pressure and flow in concert then the earlier initial apparent haemodynamic benefit may be just a transient "bump" rather than sustained.

Second, non-invasive measurements might not faithfully reflect invasively defined changes. Although beat-by-beat non-invasive blood pressure (Imholz et al. 1998; Lansdorp et al. 2011; Schutte et al. 2004) has been validated extensively it has generally not been in patients undergoing manipulation of intra cardiac timings and so validity in this context is less clear. So far, most pressure-based haemodynamic optimisation have therefore used incontrovertible, invasive measurements (Auricchio et al. 1999a; Kass et al. 1999; Nelson et al. 2000)

8.9 Immediate and delayed effects of changes in pacemaker settings

This study is unique in performing beat-by-beat simultaneous invasive measurements of aortic pressure and flow, and being explicitly designed to deliver a large "bump" in pressure and flow by changing AV delay (Whinnett et al. 2006b). The cardiac cycle at the instant of transition has an abnormally long R-R interval for that single beat only, because the pacemaker maintains the P-P interval and lengthens the P-R interval instantly. One beat therefore has very prolonged filling and a much greater stroke volume (Figure 8-2), but in the reverse state i.e. 120 ms to 40 ms, where the PP

interval is prolonged and RR interval is unaltered this prominent first beat change does not occur. This is shown in the schematic diagram below (Figure 8-8), where the changes which occur between the “PP” and “RR” intervals when the AV delay is changed from 40ms to 120ms and from 120ms to 40ms are clearly illustrated.

From the second beat onwards there is a plateau of limited duration during which stroke volume is elevated and aortic systolic pressure are both increased. A similar rise of non-invasive systolic pressure was also seen. This early plateau is presumed to be an instant, direct cardiac effect of changes in pacemaker setting that increases cardiac output.

The early plateau is followed by a phase of significant fall in pressure caused by a fall in total peripheral resistance (Figure 8-7) and not by a fall in flow. The most likely explanation is progressive vasodilatation in the circulation, which does not begin immediately after the AV delay transition but after a short delay. Its onset (defined by the early plateau ending and being replaced by a downward trend in pressure) is approximately beat 13, and it is established (onset of the late plateau) by approximately beat 18. The R-R interval is 600 ms and therefore we can say that this secondary vasodilatation process begins to manifest at ~8 seconds and seems to be established by ~11 seconds.

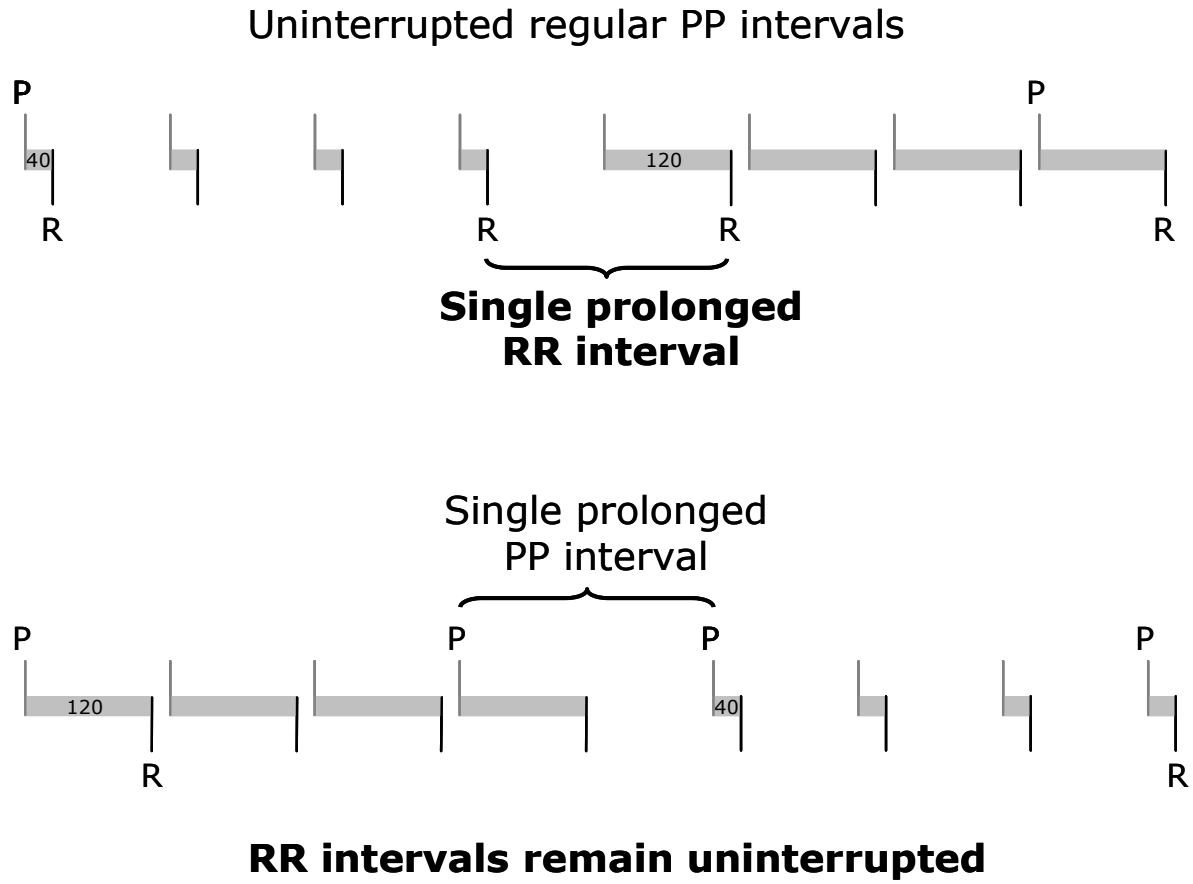


Figure 8-8 Asymmetrical impact of AV-lengthening and AV-shortening. In the top panel change occurs from AV 40 to 120 ms. The PP intervals are uninterrupted because at the point of transition there is a single prolonged RR interval. This results in a greater diastolic filling period and hence a greater stroke volume. The lower panel shows the transition from 120 to 40 ms. In this case, the RR interval is uninterrupted and the stroke volume is therefore maintained relatively consistently within the first beat at transition. Instead it is the PP interval that is increases.

8.9.1 Is AV delay optimisation dead in the wake of SMART AV?

The SMART AV trial which showed that neither qualitative selection of preferred transmitral Doppler flow pattern nor an electrical process designed to mimic this process, conferred statistically significant benefit on any of the endpoints measured. However, widely-used optimisation methods can disagree in many patients (Turcott et al. 2010), has two implications. First, disagreeing methods cannot all be correct: almost all must be incorrect. Second, however well-conducted, a trial of one method is no guide to the expectation from another.

A 3-phase process for developing optimisation science may be fruitful. First, list methods that have satisfactory test-retest reproducibility under blinded conditions. (Transmitral Doppler pattern optimisation may not belong on this list, since its between-observer and within-observer agreement from identical data is poor (Raphael CE. et al. 2011), so that its test-retest reproducibility must certainly be worse). Second, identify those that cluster together, agreeing closely on the optimum under blinded conditions, and separating those that (despite being reproducible) disagree with the main cluster of methods. Third, of the cluster of reproducible, mutually-agreeing methods, trial the most clinically-convenient, using the careful rigorous approach of SMART-AV.

8.10 Limitations

These experiments were conducted in humans and were designed to elicit information with sufficient precision using an experimental protocol of acceptable duration. As a result, there are several limitations.

First, I did not conduct thermodilution cardiac output measurements. As a result, I cannot confirm what the beat-by-beat effect of thermodilution cardiac output would be. However, thermodilution is not well suited for such sensitive temporal resolution and its susceptibility to spontaneous variation (noise) even in expert hands is twice that of beat-by-beat Doppler aortic flow VTI, and therefore would require four times as many patients to be studied and yield information with the same confidence.

Second, I only studied a single transition: AV 40ms to 120ms. This study was not an optimisation protocol but sought quantitative temporal information about the downstream physiology to inform design of optimisation protocols. I therefore cannot state with certainty that the time courses of effects of changes between other pairs of settings would be the same.

Third, this study of the acute invasive haemodynamics cannot confirm whether these changes are sustained over days or years. To test whether there are even later phases would be possible by this approach in humans.

Finally, the number of patients studied was not large but the design was able to resolve the time course with high fidelity because each patient underwent repeated beat-to-beat direct invasive measurements (attenuating noise), and the signal was arranged to be large. My patients were not selected for any characteristic other than those listed in the enrolment criteria and are representative of patients referred for CRT.

8.11 Conclusions

Non-invasive blood pressure has been proposed for optimising AV delay in CRT, but the immediate pressure increment partially fades within a few seconds. These simultaneous invasive blood pressure and flow measurements show that this secondary fall in non-invasive blood pressure is not a failure of non-invasive monitoring, nor the end of a short-lived improvement in cardiac output, but rather a manifestation of reduction in peripheral resistance that may indeed be desirable. Aortic flow, invasive aortic pressure and non-invasive finger blood pressure all rise together immediately following a change in AV delay and therefore early measurements of these are plausible candidates for optimisation. However, after a few seconds pressure delays somewhat. Planners designing optimisation protocols may choose to measure pressure for its favourable noise characteristics and potential for non-invasive measurements, but they should do so without delay.

9 Synthesis

In this thesis I have identified that beat-to-beat biological variability of all the widely recommended markers of dyssynchrony is far too large to be realistically used in optimisation. Although making multiple repetitions can alleviate this problem, the number of replicates needed is very large and might not be clinically practical unless either markers become available that are measurable with higher precision, or automated methods become available to make multiple measurements less time-consuming.

Since these echocardiographic markers have such large beat-to-beat variability there is no chance they will predict response to the level of reliability that is claimed in many publications. The response markers also have intrinsic variability which is wide compared to the spectrum of changes seen in randomised controlled trials. Moreover, from my local data, such variability appears to arise immediately and not be particularly dependant on passage of a long time. This indicates that the variability should not be assumed to be genuine variation in clinical status, but rather should be suspected to be spontaneous beat-to-beat variability that does not have clinical meaning.

As a result of the above, many highly skilled specialist centres have reported strong predictors of response which must either be a coincidence, inadvertent self deception or intentional mis-representation, because they exceed the limit of mathematical possibility. Studies conducted to normal scientific standards consistently stay within these limits. Therefore a large number of these papers need to be abandoned (and perhaps retracted). In future, studies of poor design should not be initiated, or if initiated and completed, not given credence in the literature.

With my colleagues I have identified two key pre requirements, without which it makes no sense to embark upon an attempt of reliable prediction:

- Reliable measures of response i.e. those which remain stable within individuals
- Candidate dyssynchrony markers with good blinded test-retest reproducibility

In parallel with my non-invasive work, I have also conducted an invasive study to establish, without fear of contradiction, the relative potential of flow and pressure measurements in optimisation of CRT. This study demonstrated that flow measurements, which might potentially be measured by echocardiographic measures, show a sustained increase following a change to a favourable AV delay, whereas pressure initially rises but then partially declines. This shows that measuring flow could be used for optimisation purposes provided a high reproducibility was obtained. Pressure measurements are a plausible alternative, and can be measured just as reliably non-invasively, but are best measured in the early phase after the change in the intra-cardiac settings.

9.1 The puzzle of dyssynchrony and response

I have identified that there is substantial difference between centres of dyssynchrony markers to determine response. By mathematically calculating the maximal possible R^2 values of these markers and the maximum R^2 values of ejection fraction variability, I have calculated the best possible R^2 value for prediction: the "contraction factor of R^2". Therefore, I am able to confidently state that it is not realistic to expect a good correlation between dyssynchrony markers and prediction of ΔEF.

The ΔEF, when repeated echocardiographic measurements were performed at our local hospital, showed that the standard deviation between these repeated values ranged between 10 and 20%. This is not dissimilar to published values of EF variability of at least 10% (Otterstad et al. 1997). This demonstrates the implausibility of definitions of response of 10-15% being reliably predicted by any baseline marker or combination of markers. The real puzzle is not why dyssynchrony does not successfully predict response in some studies, but why anyone takes seriously any studies that say it does.

The responsibility may lie more properly with us as an audience for several weaknesses in application of normal scientific critique.

Sound-bite Susceptibility. As a community we accepted uncritically the name "cardiac resynchronization therapy". With repetition it became obvious that quantifying mechanical dyssynchrony (which can only refer to ventricular timings, since atrium and ventricle should not be synchronous) would quantify degree of benefit. Obvious, but not necessarily true. Experimental investigation of the mechanism of action remains at an early stage. A neutral term such as "biventricular pacing" might reduce cognitive distortion.

Authority Awe. Physical science audiences judge a scientific finding by the precise nature of the experiment, the attention to detail and the track record of previous results being verified by others. Cardiological audiences may not apply the same level of scrutiny (e.g. bias-resistance is rarely debated) and may suffer from the availability heuristic (judging the credibility of sources from their visibility rather than their track record of reliability). Audiences could usefully restore habits from their earlier scientific training.

Amnesia. Hearing of each successive novel predictive marker with progressively more excellent predictive capacities, cardiological audiences often forget to ask what happened to yesterday's markers. If two different markers predict excellently, they must agree almost perfectly; if the latter is not the case, the former is not credible. Audience memory would help resist successive overstatements.

Paralysis. Physical scientists, hearing of an efficacious new approach, rush back to try it out in small experiments. Cardiological audiences may feel unable to do this. Yet simple experiments taking only minutes can quickly reject some claims. One example is evaluation of blinded test-retest reproducibility, and application of the formulae in this thesis. Another is adjustment of interventricular delay in a single biventricular pacemaker patient, with blinded measurement of mechanical dyssynchrony. Any reliable marker will show a clear, reproducible minimum. If not, it cannot credibly classify dyssynchrony status across different patients (Pabari PA et al. 2012).

Wishful thinking. We all want our specialty of echocardiography to be relevant. Reports of successful application are therefore intrinsically popular. But this is failure to separate our individual skill as echocardiographers, from the ability of one echocardiographic technique to deliver what is claimed. Distinguishing falsity of a

hypothesis from personal inadequacy requires courage, but is worthwhile, since it is the hallmark of science.

Cryptic commenting. Even when experts carefully review available methods, and tabulate that dyssynchrony markers are intensely vulnerable to noise and sometimes choice of measurement location so that there is risk of "dialling in" any clinically desired level of dyssynchrony (Abraham et al. 2007), nevertheless the published conclusion may not be quantitatively clear on the implication for claims of response prediction (MacRoberts MH & MacRoberts BR 1984).

Bias blindness. We frequently confuse bias (which arises from study design) with chance (which is addressed by p values). Large study sizes reduce the impact of chance, but increase the false statistical significance of the effects of bias. Even the most methodical of clinician audiences – guideline bodies – consistently consider observational studies (if large) as the same level of evidence ("B") as a randomised controlled trial. We should become alert to the fact that large study size increases susceptibility to bias.

Test-retest taboo. We confuse re-measurement of identical digital images with genuine test-retest reproducibility, entirely ignoring the majority of variability which occurs between beats. Test-retest variability can be readily checked by clinicians (Finegold et al. 2012).

9.2 The puzzle of optimisation

The concept of optimisation was initially thought to be fairly simple. By changing intra-cardiac timings to a more favourable setting in an individualised manner per patient, it would surely be possible to give the patient some advantage. This may sound plausible, but in reality I have identified some flaws with this concept. In an ideal world, if we were measuring a noise-free parameter as settings were changed, then we could confidently state that that we were improving patient settings. However any real life parameter which is measured has a significant noise component, and this noise (when measured as a single beat), is in almost all cases much larger than the genuine difference between settings. For this reason, the protocols that are claimed to permit optimisation by making a single measurement at each of several settings and then choosing the setting with the highest measurement, are almost all incorrect and the claimed positive results are almost all the results of wishful thinking on behalf of the researchers.

I have mathematically shown the impact of signal and noise on the effectiveness of optimisation. I have demonstrated that the information content (the relative amount of useful information versus random variability) has an enormous effect on reliability of optimisation. I have identified a series of illusions regarding optimisation of CRT that frequently afflict the literature, demonstrating the importance of asking for the confidence interval of any observed optimum. This allows us all to evaluate how much we should trust an apparent change in optimum after a second optimisation process.

I have shown what steps are important in improving the quality of the process of optimisation. These are as follows:

(a) reducing the noise component by performing replicate measurements at each setting

(b) maximising the signal, for which we know the most reliable approach is increasing the heart rate in some parameters

(c) being aware that large confidence intervals can cause misinterpretation of the reliability of optimisation of CRT.

9.3 Biological variability of optimisation using a spectrum of echocardiographic parameters

In this thesis I studied a selection of echocardiographic markers, head-to-head, as potential tools for VV optimisation of CRT devices. I have scrutinised each parameter looking at the biological variability which impacts on the feasibility of each as a potential optimisation tool. I have identified that the advanced modalities of 3D echocardiography and tissue Doppler imaging demonstrate a poor signal-to-noise ratio yet take longer to acquire and analyse. The more simplistic pulsed Doppler methods, namely LVOT VTI and aortic pre-ejection times show a better signal-to-noise ratio, resulting in a lower spread of results when repeated optimisation processes are performed. All parameters improved their signal-to-noise ratio when multiple replicates of each setting are used. Time taken, for acquiring and measuring each traces at each setting, then becomes the limiting step restricting the level of precision that can be achieved.

Any of the above parameters could, in principle, be used for selection of those suitable for CRT and for optimisation of VV delay. But multiple measures would be required to achieve adequate signal-to-noise ratio. Recommendations for protocols are usually made without regard to how many replicates would be needed to achieve the bare minimum level of precision that might be considered clinically adequate. Calculating that number of replicates is surprising: it is very much higher indeed than the numbers described in protocols or recommended in guidelines.

Finally, exploring simple opportunities to improve upon conventional parameters, I compared the velocity time integral to the peak velocity of LVOT Doppler traces as alternatives for optimisation of CRT. Better signal-to-noise ratio was obtained with

peak velocity. Meanwhile, time to analyse was also shorter with peak velocity than VTI, so that more replicates could be averaged in the same time frame, making peak velocity an even better candidate for optimisation of CRT devices.

9.4 The effect of AV delay optimisation on invasive pressure and flow measurements: implications for the design of AV optimisation protocols

The detailed invasive study I performed showed, on a beat-by-beat basis, the changes which occur when AV delay is altered. Pressure increased initially but then dipped away over the subsequent few beats to settle at a lower plateau. This behaviour is mirrored closely by the non-invasive blood pressure, which means the non-invasive signal is not in error. This opens up the possibility of using non-invasive Finapres - type signal as a method for CRT optimisation.

My data showed that flow rises initially with a further rise temporally aligned with a fall in pressure. I have explained this secondary dip in pressure and rise in flow, by a change in peripheral resistance which I interpret as a favourable phenomenon in heart failure patients who are characteristically in a state of pathological vasoconstriction. From a practical point of view, for the design of optimisation protocols, my work shows that the greatest signal size in pressure is at the first few beats post transition of settings. In contrast if flow is used as the optimisation parameter, there is less urgency to make the measurements because in later beats the signal is not systematically smaller.

9.5 Recommendations for performing optimisation of Cardiac Resynchronisation Therapy devices

When performing an optimisation process, the aim is to detect small changes in the measured parameter as intra cardiac timings are altered. For this to be reliable, i.e. to achieve a small scatter on the optima obtained on repeated optimisations requires a high intraclass correlation in the measurements being made. My thesis demonstrates that this is mathematically possible but would be very laborious and time consuming because a large number of replicates will be required. Therefore it is rational to adopt a method whose intrinsic signal-to-noise ratio is as high as possible, so that the number of replicates needed are as small as possible. For example peak velocity is on these grounds a better candidate than VTI.

A second consideration is that one should favour methods where the measurement and analysis time per beat is small. In clinical practice it is the analysis time which is by far the most time consuming component. In future a process of automating some or all of this would be likely to increase speed and /or precision of the optimisation process. Algorithms which could automatically measure the peak velocities would be attractive.

The difference between beats is a greater source of variability than operator uncertainty regarding how to draw round the beats. Therefore I would not expect automation of measurement of VTI to necessarily greatly help through elimination of individual-beat measurement error, but by making it possible to average multiple beats at little cost in time it could make it clinically realistic to perform high-replicate, reliable optimisations in routine practice.

A pilot study is already underway to see if VTI can be automatically calculated. A similar algorithm might also be implemented for peak velocity. This would assist those using outflow measurements for optimisation of CRT devices in which beat-to-beat variability can mainly be reduced by multiple replicate measurements.

There is no reason to be limited to either Peak velocity or VTI as a parameter. Any variable would be a possibility but those with greater noise would require a much greater level of repetition to achieve the same level of precision of optimum. For each doubling of noise, we know that the number of replicates required increases 4 fold. Looked at another way, if the degree of uncertainty of the optimum is considered unsatisfactorily wide, the number of replicates can be increased to narrow it. But achieving a halving of this uncertainty requires approximately 32 times more replicates.

9.6 Impact on selection criteria for Cardiac Resynchronisation Therapy devices

As has been demonstrated, signal-to-noise of the advanced echocardiographic modalities is poor when single measurements are performed. Therefore despite claims made in the literature, it is not plausible that a single assessment of mechanical dyssynchrony can tell us anything useful about the patient (since it varies so extensively between beats). Many of the claims made for the success of predicting response are definitely incorrect and should be withdrawn.

Current practice and guidelines focus on ECG criteria, and at present there is little basis to press for echocardiography to be added as a large contributor. Chapter 5 demonstrates that QRS duration has a good signal-to-noise ratio which further improves upon multiple replicates, and is quick and QRS duration can be automatically calculated. However the criteria of biological plausibility of optimum obtained is not always fulfilled, even if the criteria of reproducibility and agreement with repeated optimisations are met (Kyriacou et al. 2012). For this reason I believe that QRS duration alone cannot confidently be used as a selection criterion.

Additional use of an echocardiographic criterion requires careful thought as there needs to be grounds for confidence in the parameter used. Each department is likely to have a preferred method, which, they feel if used carefully could assist with selection. However, the record - even amongst carefully argued clinical research papers that have undergone peer review - shows that most of these feelings are misplaced optimism. Feelings of a local department, without even perfunctory statistical scrutiny or peer review, are even less likely to be reliable. Gut instinct should not be believed to be reliable, because the record shows clearly that it is unreliable.

Proposals for a future selection criterion would need very careful stepwise evaluation. The first step would be establishment of the genuine, honest test-retest reproducibility of both the marker and the planned marker of response. After this a protocol could be designed, to have sufficient replicates. The number of replicates should be determined not by just guessing, asking an expert (who in turn guesses), or plucking a number from the air but rather by calculation from the marker's properties and the level of precision required (Pabari et al. 2011).

9.7 Recommendations for future studies

Although most optimisations are performed at rest, patients experience their most severe symptoms when they are exercising. Therefore, identifying the benefits of optimisation during exercise would be the next logical step forward. I have been collecting pilot data in 11 patients to demonstrate the effect of AV optimisation on peak cardiac power output measured whilst cycling on a supine bicycle.

I have demonstrated that in 10/11 patients (91%) an independently established optimal AV delay shows individually detectable increases in both flow and pressure. In all patients cardiac power output increases when compared with a non-optimal delay.

In this thesis I have demonstrated that ultimately, almost any parameter can in principle be used for optimisation of CRT but obtaining a sufficiently large signal-to-noise ratio is essential. This is shown in this thesis when measurements are made at rest, and likely to be even more important on measurements taken during exercise where noise is only going to be a greater component.

Beat-to-beat variability, which is inherent and unavoidable, means that a bias resistant protocol of systematically averaging multiple replicate measurements is essential. My thesis lists strategies for achieving this.

10 Appendices

10.1 Appendix 1: Invitation letter

St Mary's NHS Trust

St Mary's Hospital
International Centre for Circulatory Health,
59-61 North Wharf Road,
London
W2 1LA

Direct Dial: 020 7594 1093
Direct Fax: 0707 505 5876

Dear Sir / Madam

You are being invited to participate in a research study. You have been recruited from either a Heart failure or Pacemaker clinic where it has been identified that you may be suitable to participate in this research project.

Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

Please find attached a Patient Information sheet and contact details should you have any further questions

Thank you for considering participating in this project

Kind Regards

Dr Punam Pabari

10.2 Appendix 2: Patient information leaflet

Non Invasive Haemodynamics to Probe Physiology and Echocardiographic Dyssynchrony in Chronic Heart Failure

What is the purpose of the study?

Patients with heart failure are limited greatly by their symptoms and often require emergency admissions to hospital. Day to day activity is restricted and there is an increased risk of death associated with the condition. We know that the heart pumping mechanism is often not in a coordinated fashion, therefore "dyssynchronous" leading to inefficient working of the heart.

Specialist pacemakers, Cardiac Resynchronisation Therapy, (CRT), have been shown to improve the symptoms and survival of certain groups of patients, in addition to maximum medical treatment. There are selection methods in place for choosing patients who will benefit from these CRT devices; however we are unclear which will predict the best long-term outcome. A beneficial response, by improving co ordination of ventricle contraction, is seen by an improvement in blood pressure on altering pacemaker settings. At present a proportion of patients do not respond as predicted after CRT implantations or show the expected benefits.

This study aims to research the different imaging methods of echocardiography to maximise the best outcome in terms of blood pressure response at different pacemaker settings. The study will then progress to look at changes in blood pressure once the optimal, the best, pacemaker setting had been identified. This will give us a greater understanding into the long term benefits of improving pacemaker settings in patients with CRT.

We hope this study will give a greater insight into which methods of selecting patients for CRT devices provides greatest benefit, and whether optimising pacemaker settings will give long term benefits and therefore a worth while pursuit.

Why have I been chosen?

You will probably have been contacted and asked whether you would be interested in helping with our research following an outpatients' appointment at the heart failure and / or pacemaker clinic, where we have noted you currently have a CRT pacemaker.

You can participate in our research providing that you have a Cardiac Resynchronisation Therapy pacemaker fitted and heart failure.

Do I have to take part?

No. Your decision to participate or not to participate within this study is completely free and voluntary. You have the right to refuse as well as to withdraw your participation at any time (even if you agree today) without giving a reason. If you decide not to participate or to withdraw, it will not affect the quality of your care or treatment, nor the relationship you have with your doctor and nursing team. Should you withdraw, your doctor will recommend appropriate clinical treatment and follow-up to ensure your safety.

What will happen to me if I take part?

When you arrive for the testing, we will ask you to lie comfortably on a bed, remove your top half clothing, and we will attach several pieces of monitoring equipment such as an ECG machine and a blood pressure monitor to your finger.

The study takes between 2 – 3 hours and you will be asked to lie on your side and back while we take various echocardiographic images, specialised ultrasound scans of your heart. During this time the pacemaker settings will be altered and the images taken at each setting. We will be taking continuous measurements of blood pressure during the testing, and these along with the images taken will give us information to analyse. The pacemaker settings will be changed in the same way as the technicians use when you attend the pacemaker clinic, in order to check the pacemakers. The settings will be returned to what they were when you arrived at the hospital, before you go home.

This is a totally painless procedure however, if at any time during the study you feel uncomfortable, or wish to stop, then we will terminate immediately.

What are the side effects of this treatment when taking part?

This research does not involve any blood tests, drugs or invasive procedures.

Pacemaker settings will be altered during the study but reset to the original values prior to you going home.

What if something goes wrong?

Imperial College London holds insurance policies which apply to this study. If you experience harm or injury as a result of taking part in this study, you will 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 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
Investigators (Dr Punam Pabari or Dr Darrel Francis 0207 594 1093).

The normal National Health Service complaint complaints mechanisms are also available to you. If you are still not satisfied with the response, you may contact the Imperial College Clinical Research Office.

What are the possible disadvantages and risks of taking part?

This research does not involve any blood tests, drugs, contrast dye or invasive procedures.

Some people find undergoing echocardiography uncomfortable as it requires lying on the bed in various positions. We will allow for breaks during the scanning periods to alleviate this.

Occasionally the blood pressure measuring device, the Finapres, can cause tingling. This is short-lived and stops once removed.

What are the possible benefits of taking part?

This research will not benefit you directly. However, it will contribute to our understanding of physiological changes which occur during optimisation of cardiac resynchronisation therapy and improve our selection of patients for these devices.

This may help with the development of treatments and management of heart failure in the future.

What if new information becomes available?

Sometimes during the course of a research project, new information becomes available about the treatment being studied which could influence your willingness to participate. If this happens, your research doctor will tell you about it and discuss whether you want to or should continue in the study. If you decide not to carry on, your research doctor will make arrangements for your care to continue. If you decide to continue in the study you will be asked to sign an updated consent form.

Also, on receiving new information your research doctor might consider it to be in your best interests to withdraw you from the study. He/she will explain the reasons and arrange for your care to continue.

If the study is stopped for any other reason, you will be told why and your continuing care will be arranged.

What happens when the research study stops?

By the end of this process we hope to understand more about the physiological changes which occur when cardiac resynchronisation devices are optimised, i.e. have the settings changed to maximise benefits, and a greater understanding of changes which occur for long term benefits.

What if something goes wrong?

We will take every care in the course of this study, and the risk involved is minimal. However, if through our negligence, you should suffer any harm then you will be compensated and a full investigation will be conducted.

There are no adverse effects expected.

Will my taking part in this study be kept confidential?

If you agree to take part in this study, data collected about you will be entered on to a computer. However, all data entered will be in an anonymous format and any information obtained from this investigation that can be identified will remain confidential.

What will happen to the results of the research study?

Scientific data from this study may be presented at meetings and published so that the information can be used to help others, but your participation in the study will not be made known and will be kept strictly confidential.

Who is organising and funding the research?

This research is supported by the British Heart Foundation.

Who has reviewed the study?

This study has been reviewed the Brompton, Harefield and NHLI Hospital NHS Trust Research Ethics committee.

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

Dr Punam Pabari on 0207 594 1093

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

10.3 Appendix 3: Consent form

St Mary's NHS Trust

St Mary's Hospital
International Centre for Circulatory Health,
59-61 North Wharf Road,
London
W2 1LA

Direct Dial: 020 7594 1093
Direct Fax: 0707 505 5876

Centre Number::
Study Number:
Patient Identification Number for this trial:

CONSENT FORM

Title of Project: Use of Non Invasive Haemodynamics to Probe Physiology of Echocardiographic Features of Cardiac Dyssynchrony in Chronic Heart Failure

Name of Researcher: Dr Punam Pabari **Please tick box**

I confirm that I have read and understand the information sheet dated 22/11/2007 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
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.	☐
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 St Mary's Hospital 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.	☐
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.	☐
I agree to my GP being informed of my participation in the study.	☐
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

10.4 Appendix 4: Determining the contraction factor

A series of measurements of a variable (x) in a particular patient (e.g. the i'th patient) will not always be identical, but rather be scattered randomly with a mean (μ_{x_i}) and an error component (e_{x_i}), i.e. $x_i = \mu_{x_i} + e_{x_i}$. The same applies for another variable, y_i = $\mu_{y_i} + e_{y_i}$. The statistical term "error" in this context amalgamates observer error, equipment error, operator error, and genuine biological variability, all of which contribute to the scatter. The combined error in x is described by its variance $\sigma^2 e_{x_i}$. The size of this error variance in comparison to the size of the variance in the true underlying values $\sigma^2 \mu_{x_i}$ is an important characteristic of any measured biological variable. This is described by the intraclass correlation coefficient:

$$ICC_{x_i} = \frac{\sigma^2 \mu_{x_i}}{\sigma^2 \mu_{x_i} + \sigma^2 e_{x_i}}$$

The same can be said for y. If one tries to estimate the correlation between x and y, then the observed coefficient of determination $R^2(x_i,y_i)$ will not be the R^2 between the underlying variables $R^2(\mu_{xi}, \mu_{yi})$ but lower:

$$R^2\ (x_i, y_i) = R^2\ (\mu_{x_i}, \mu_{y_i}) \times \frac{\sigma^2 \mu_{x_i}}{\sigma^2 \mu_{x_i} + \sigma^2 e_{x_i}} \times \frac{\sigma^2 \mu_{y_i}}{\sigma^2 \mu_{y_i} + \sigma^2 e_{y_i}}$$

$$= R^2\ (\mu_{x_i}, \mu_{y_i}) \times ICC_{x_i} \times ICC_{y_i}$$

The factor $ICC_{x_i} \times ICC_{y_i}$ is the maximum R^2 that can be observed between x_i and y_i even if there is a perfect correlation between μ_{x_i} and μ_{y_i}.

In this paper, x is a baseline dyssynchrony marker and y (for example ΔLVEF) is a change in response marker over time. In this field unfortunately ICC_{x_i} is not known

for most dyssynchrony markers because blinded test-retest studies have not been performed, or not been reported. ICC_{y_i} , in contrast, can be estimated from randomised controlled trials that performed blinded analysis. The variance of the changes over time in the control arm $\sigma^2(\Delta_{\text{control arm}})$ is $\frac{\sigma^2\mu_{y_i}}{\sigma^2\mu_{y_i}+\sigma^2e_{y_i}}$, the corresponding variance in the intervention arm $\sigma^2(\Delta_{\text{intervention arm}})$ is $\frac{\sigma^2\mu_{y_i}}{\sigma^2\mu_{y_i}+\sigma^2e_{y_i}}$. Although these are not the same individual patients, the randomised trial design ensures that they are comparable so that the difference between them, σ^2 ($\Delta_{\text{intervention arm}}$)- $\sigma^2(\Delta_{\text{control arm}})$ is an estimate of the elusive $\frac{\sigma^2\mu_{y_i}}{\sigma^2\mu_{y_i}+\sigma^2e_{y_i}}$. Thus from an individual trial, the estimate of ICC_{y_i} is 1- $(\sigma^2(\Delta_i(control\ arm)))/(\sigma^2(\Delta_i(intervention\ arm)))$.

Estimate of the R^2 contraction factor caused by natural variability in response variable $= ICC_{y_i} \approx 1-\left(\frac{\sigma(\Delta_{control\ arm})}{\sigma(\Delta_{intervention\ arm})}\right)^2$

That last expression relies only to the standard deviation of Δ in the control and device arms, and is therefore easy to calculate in trials that report the distributions of Δ. It is not sufficient to know the distribution of the initial and final LVEFs. Rather, the distribution *of the change*, i.e. the standard deviation of Δ, is needed.

From a single trial, the estimate of ICC_{y_i} may overestimate or underestimate the true value. It will not be reliable in any study where there is a breakdown in blinding. Although a breakdown in blinding is not likely to be admitted, it can become obvious

from the pattern of Δy_i. The characteristic features of breakdown in blinding in endpoint measurements of a trial are:

- Unusually narrow $\sigma^2 \Delta y_i$ in comparison to those of other trials with ostensibly similar methods;
- Markedly asymmetric Δ values - for variables whose increase is desirable, the upper limit of the 95% confidence interval much further from the median than the lower limit is; for variables whose decrease is desirable, the converse.

Averaging estimates across multiple trials, weighting by their $\frac{1}{\sqrt{sample\ size}}$ allows a combined estimate for ICC_{y_i} to be developed for each class of response variable y_i, such as ΔLVEF.

11 References

(1987) Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). The CONSENSUS Trial Study Group. *N.Engl.J.Med.*, 316, 1429-1435.

Abraham, T.P., Dimaano, V.L. & Liang, H.Y. (2007) Role of tissue Doppler and strain echocardiography in current clinical practice. *Circulation*, 116, 2597-2609.

Abraham, W.T., Fisher, W.G., Smith, A.L., DeLurgio, D.B., Leon, A.R., Loh, E., Kocovic, D.Z., Packer, M., Clavell, A.L., Hayes, D.L., Ellestad, M., Trupp, R.J., Underwood, J., Pickering, F., Truex, C., McAtee, P. & Messenger, J. (2002) Cardiac resynchronization in chronic heart failure. *N.Engl.J.Med.*, 346, 1845-1853.

Abraham, W.T., Fonarow, G.C., Albert, N.M., Stough, W.G., Gheorghiade, M., Greenberg, B.H., O'Connor, C.M., Sun, J.L., Yancy, C.W. & Young, J.B. (2008) Predictors of in-hospital mortality in patients hospitalized for heart failure: insights from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *J.Am.Coll.Cardiol.*, 52, 347-356.

Abraham, W.T., Young, J.B., Leon, A.R., Adler, S., Bank, A.J., Hall, S.A., Lieberman, R., Liem, L.B., O'Connell, J.B., Schroeder, J.S. & Wheelan, K.R. (2004) Effects of cardiac resynchronization on disease progression in patients with left ventricular systolic dysfunction, an indication for an implantable cardioverter-defibrillator, and mildly symptomatic chronic heart failure. *Circulation*, 110, 2864-2868.

Adabag, S., Roukoz, H., Anand, I.S. & Moss, A.J. (2011) Cardiac resynchronization therapy in patients with minimal heart failure a systematic review and meta-analysis. *J.Am.Coll.Cardiol.*, 58, 935-941.

Adams, K.F., Jr., Fonarow, G.C., Emerman, C.L., LeJemtel, T.H., Costanzo, M.R., Abraham, W.T., Berkowitz, R.L., Galvao, M. & Horton, D.P. (2005) Characteristics and outcomes of patients hospitalized for heart failure in the United States: rationale, design, and preliminary observations from the first 100,000 cases in the Acute Decompensated Heart Failure National Registry (ADHERE). *Am.Heart J.*, 149, 209-216.

Agarwal, R., Beshai, J.F. & Lang, R.M. (2009) Can real-time three-dimensional echocardiography be used reliably for the assessment of left ventricular dyssynchrony? *Arch.Cardiovasc.Dis.*, 102, 469-472.

Anselmino, M., Antolini, M., Amellone, C., Piovano, E., Massa, R. & Trevi, G. (2009) Optimization of cardiac resynchronization therapy: echocardiographic vs semiautomatic device algorithms. *Congest.Heart Fail.*, 15, 14-18.

Auricchio, A. & Prinzen, F.W. (2011) Non-responders to cardiac resynchronization therapy: the magnitude of the problem and the issues. *Circ.J.*, 75, 521-527.

Auricchio, A., Stellbrink, C., Block, M., Sack, S., Vogt, J., Bakker, P., Klein, H., Kramer, A., Ding, J., Salo, R., Tockman, B., Pochet, T. & Spinelli, J. (1999a) Effect of pacing chamber and atrioventricular delay on acute systolic function of paced patients with congestive heart failure. The Pacing Therapies for Congestive Heart Failure Study Group. The Guidant Congestive Heart Failure Research Group. *Circulation*, 99, 2993-3001.

Auricchio, A., Stellbrink, C., Sack, S., Block, M., Vogt, J., Bakker, P., Mortensen, P. & Klein, H. (1999b) The Pacing Therapies for Congestive Heart Failure (PATH-CHF) study: rationale, design, and endpoints of a prospective randomized multicenter study. *Am.J.Cardiol.*, 83, 130D-135D.

Baldasseroni, S., Opasich, C., Gorini, M., Lucci, D., Marchionni, N., Marini, M., Campana, C., Perini, G., Deorsola, A., Masotti, G., Tavazzi, L. & Maggioni, A.P. (2002) Left bundle-branch block is associated with increased 1-year sudden and total

mortality rate in 5517 outpatients with congestive heart failure: a report from the Italian network on congestive heart failure. *Am.Heart J.*, 143, 398-405.

Bank, A.J., Kaufman, C.L., Kelly, A.S., Burns, K.V., Adler, S.W., Rector, T.S., Goldsmith, S.R., Olivari, M.T., Tang, C., Nelson, L. & Metzig, A. (2009) Results of the Prospective Minnesota Study of ECHO/TDI in Cardiac Resynchronization Therapy (PROMISE-CRT) study. *J.Card Fail.*, 15, 401-409.

Barnett, D., Phillips, S. & Longson, C. (2007) Cardiac resynchronisation therapy for the treatment of heart failure: NICE technology appraisal guidance. *Heart*, 93, 1134-1135.

Barold, S.S., Ilercil, A. & Herweg, B. (2008) Echocardiographic optimization of the atrioventricular and interventricular intervals during cardiac resynchronization. *Europace.*, 10 Suppl 3, iii88-iii95.

Bax, J.J. & Gorcsan, J. (2009) Echocardiography and noninvasive imaging in cardiac resynchronization therapy: results of the PROSPECT (Predictors of Response to Cardiac Resynchronization Therapy) study in perspective. *J.Am.Coll.Cardiol.*, 53, 1933-1943.

Bax, J.J., Marwick, T.H., Molhoek, S.G., Bleeker, G.B., van, E.L., Boersma, E., Steendijk, P., van der Wall, E.E. & Schalij, M.J. (2003a) Left ventricular dyssynchrony predicts benefit of cardiac resynchronization therapy in patients with end-stage heart failure before pacemaker implantation. *Am.J.Cardiol.*, 92, 1238-1240.

Bax, J.J., Molhoek, S.G., van, E.L., Voogd, P.J., Somer, S., Boersma, E., Steendijk, P., Schalij, M.J. & van der Wall, E.E. (2003b) Usefulness of myocardial tissue Doppler echocardiography to evaluate left ventricular dyssynchrony before and after biventricular pacing in patients with idiopathic dilated cardiomyopathy. *Am.J.Cardiol.*, 91, 94-97.

Bertini, M., Valzania, C., Biffi, M., Martignani, C., Ziacchi, M., Pedri, S., Domenichini, G., Diemberger, I., Saporito, D., Rocchi, G., Rapezzi, C., Branzi, A. & Boriani, G. (2010) Interventricular delay optimization: a comparison among three different echocardiographic methods. *Echocardiography.*, 27, 38-43.

Bertini, M., Ziacchi, M., Biffi, M., Martignani, C., Saporito, D., Valzania, C., Diemberger, I., Cervi, E., Frisoni, J., Sangiorgi, D., Branzi, A. & Boriani, G. (2008) Interventricular delay interval optimization in cardiac resynchronization therapy guided by echocardiography versus guided by electrocardiographic QRS interval width. *Am.J.Cardiol.*, 102, 1373-1377.

Beshai, J.F., Grimm, R.A., Nagueh, S.F., Baker, J.H., Beau, S.L., Greenberg, S.M., Pires, L.A. & Tchou, P.J. (2007) Cardiac-resynchronization therapy in heart failure with narrow QRS complexes. *N.Engl.J.Med.*, 357, 2461-2471.

Bleeker, G.B., Mollema, S.A., Holman, E.R., Van, d., V, Ypenburg, C., Boersma, E., van der Wall, E.E., Schalij, M.J. & Bax, J.J. (2007a) Left ventricular resynchronization is mandatory for response to cardiac resynchronization therapy: analysis in patients with echocardiographic evidence of left ventricular dyssynchrony at baseline. *Circulation*, 116, 1440-1448.

Bleeker, G.B., Schalij, M.J., Boersma, E., Holman, E.R., Steendijk, P., van der Wall, E.E. & Bax, J.J. (2007b) Relative merits of M-mode echocardiography and tissue Doppler imaging for prediction of response to cardiac resynchronization therapy in patients with heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. *Am.J.Cardiol.*, 99, 68-74.

Bleeker, G.B., Yu, C.M., Nihoyannopoulos, P., De, S.J., Van, d., V, Holman, E.R., Schalij, M.J., van der Wall, E.E. & Bax, J.J. (2007c) Optimal use of echocardiography in cardiac resynchronisation therapy. *Heart*, 93, 1339-1350.

Bogert, L.W. & van Lieshout, J.J. (2005) Non-invasive pulsatile arterial pressure and stroke volume changes from the human finger. *Exp.Physiol*, 90, 437-446.

Bordachar, P., Lafitte, S., Reant, P., Reuter, S., Clementy, J., Mletzko, R.U., Siegel, R.M., Goscinska-Bis, K., Bowes, R., Morgan, J., Benard, S. & Leclercq, C. (2010) Low value of simple echocardiographic indices of ventricular dyssynchrony in predicting the response to cardiac resynchronization therapy. *Eur J.Heart Fail.*, 12, 588-592.

Boriani, G., Muller, C.P., Seidl, K.H., Grove, R., Vogt, J., Danschel, W., Schuchert, A., Djiane, P., Biffi, M., Becker, T., Bailleul, C. & Trappe, H.J. (2006) Randomized comparison of simultaneous biventricular stimulation versus optimized interventricular delay in cardiac resynchronization therapy. The Resynchronization for the HemodYnamic Treatment for Heart Failure Management II implantable cardioverter defibrillator (RHYTHM II ICD) study. *Am.Heart J.*, 151, 1050-1058.

Bristow, M.R., Gilbert, E.M., Abraham, W.T., Adams, K.F., Fowler, M.B., Hershberger, R.E., Kubo, S.H., Narahara, K.A., Ingersoll, H., Krueger, S., Young, S. & Shusterman, N. (1996) Carvedilol produces dose-related improvements in left ventricular function and survival in subjects with chronic heart failure. MOCHA Investigators. *Circulation*, 94, 2807-2816.

Bristow, M.R., Saxon, L.A., Boehmer, J., Krueger, S., Kass, D.A., De, M.T., Carson, P., DiCarlo, L., DeMets, D., White, B.G., DeVries, D.W. & Feldman, A.M. (2004) Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N.Engl.J.Med.*, 350, 2140-2150.

Cappola, T.P., Harsch, M.R., Jessup, M., Abraham, W.T., Young, J.B., Petersen-Stejskal, S., Plappert, T. & St John, S.M. (2006) Predictors of remodeling in the CRT era: influence of mitral regurgitation, BNP, and gender. *J.Card Fail.*, 12, 182-188.

Carluccio, E., Biagioli, P., Alunni, G., Murrone, A., Pantano, P., Biscottini, E., Zuchi, C., Zingarini, G., Cavallini, C. & Ambrosio, G. (2011) Presence of Extensive LV Remodeling Limits the Benefits of CRT in Patients With Intraventricular Dyssynchrony. *JACC.Cardiovasc.Imaging*, 4, 1067-1076.

Cazeau, S., Bordachar, P., Jauvert, G., Lazarus, A., Alonso, C., Vandrell, M.C., Mugica, J. & Ritter, P. (2003) Echocardiographic modeling of cardiac dyssynchrony before and during multisite stimulation: a prospective study. *Pacing Clin.Electrophysiol.*, 26, 137-143.

Cazeau, S., Leclercq, C., Lavergne, T., Walker, S., Varma, C., Linde, C., Garrigue, S., Kappenberger, L., Haywood, G.A., Santini, M., Bailleul, C. & Daubert, J.C. (2001) Effects of multisite biventricular pacing in patients with heart failure and intraventricular conduction delay. *N.Engl.J.Med.*, 344, 873-880.

Cazeau, S., Ritter, P., Bakdach, S., Lazarus, A., Limousin, M., Henao, L., Mundler, O., Daubert, J.C. & Mugica, J. (1994) Four chamber pacing in dilated cardiomyopathy. *Pacing Clin.Electrophysiol.*, 17, 1974-1979.

Cazeau, S.J., Daubert, J.C., Tavazzi, L., Frohlig, G. & Paul, V. (2008) Responders to cardiac resynchronization therapy with narrow or intermediate QRS complexes identified by simple echocardiographic indices of dyssynchrony: the DESIRE study. *Eur.J.Heart Fail.*, 10, 273-280.

Chung, E.S., Leon, A.R., Tavazzi, L., Sun, J.P., Nihoyannopoulos, P., Merlino, J., Abraham, W.T., Ghio, S., Leclercq, C., Bax, J.J., Yu, C.M., Gorcsan, J., III, St John, S.M., De, S.J. & Murillo, J. (2008a) Results of the Predictors of Response to CRT (PROSPECT) trial. *Circulation*, 117, 2608-2616.

Chung, R., Sutton, R. & Henein, M.Y. (2008b) Beyond dyssynchrony in cardiac resynchronisation therapy. *Heart*, 94, 991-994.

Cleland, J., Freemantle, N., Ghio, S., Fruhwald, F., Shankar, A., Marijanowski, M., Verboven, Y. & Tavazzi, L. (2008) Predicting the long-term effects of cardiac resynchronization therapy on mortality from baseline variables and the early response a report from the CARE-HF (Cardiac Resynchronization in Heart Failure) Trial. *J.Am.Coll.Cardiol.*, 52, 438-445.

Cleland, J.G., Daubert, J.C., Erdmann, E., Freemantle, N., Gras, D., Kappenberger, L., Klein, W. & Tavazzi, L. (2001) The CARE-HF study (CArdiac REsynchronisation in Heart Failure study): rationale, design and end-points. *Eur.J.Heart Fail.*, 3, 481-489.

Cleland, J.G., Daubert, J.C., Erdmann, E., Freemantle, N., Gras, D., Kappenberger, L. & Tavazzi, L. (2005) The effect of cardiac resynchronization on morbidity and mortality in heart failure. *N.Engl.J.Med.*, 352, 1539-1549.

Colocousis, J.S., Huntsman, L.L. & Curreri, P.W. (1977) Estimation of stroke volume changes by ultrasonic doppler. *Circulation*, 56, 914-917.

Conca, C., Faletra, F.F., Miyazaki, C., Oh, J., Mantovani, A., Klersy, C., Sorgente, A., Pedrazzini, G.B., Pasotti, E., Moccetti, T. & Auricchio, A. (2009) Echocardiographic parameters of mechanical synchrony in healthy individuals. *Am.J.Cardiol.*, 103, 136-142.

Corsi, C., Lang, R.M., Veronesi, F., Weinert, L., Caiani, E.G., MacEneaney, P., Lamberti, C. & Mor-Avi, V. (2005) Volumetric quantification of global and regional left ventricular function from real-time three-dimensional echocardiographic images. *Circulation*, 112, 1161-1170.

Cowie, M.R., Wood, D.A., Coats, A.J., Thompson, S.G., Poole-Wilson, P.A., Suresh, V. & Sutton, G.C. (1999) Incidence and aetiology of heart failure; a population-based study. *Eur Heart J.*, 20, 421-428.

Cunningham D, Charles R, Cunningham M & de Lange A. (2011) Heart Rhythm Devices: UK National Clinical Audit 2009. *Network Devices Survey Group: Heart Rhythm National Audit Reports*.

Davies, L.C., Francis, D., Jurak, P., Kara, T., Piepoli, M. & Coats, A.J. (1999) Reproducibility of methods for assessing baroreflex sensitivity in normal controls and in patients with chronic heart failure. *Clin.Sci.(Lond)*, 97, 515-522.

De Boeck, B.W., Meine, M., Leenders, G.E., Teske, A.J., van, W.H., Kirkels, J.H., Prinzen, F.W., Doevendans, P.A. & Cramer, M.J. (2008) Practical and conceptual limitations of tissue Doppler imaging to predict reverse remodelling in cardiac resynchronisation therapy. *Eur J.Heart Fail.*, 10, 281-290.

Delgado, V., Ypenburg, C., van Bommel, R.J., Tops, L.F., Mollema, S.A., Marsan, N.A., Bleeker, G.B., Schalij, M.J. & Bax, J.J. (2008) Assessment of left ventricular dyssynchrony by speckle tracking strain imaging comparison between longitudinal, circumferential, and radial strain in cardiac resynchronization therapy. *J.Am.Coll.Cardiol.*, 51, 1944-1952.

Deplagne, A., Bordachar, P., Reant, P., Montaudon, M., Reuter, S., Laborderie, J., Dos, S.P., Roudaut, R., Jais, P., Haissaguerre, M., Laurent, F., Clementy, J. & Lafitte, S. (2009) Additional value of three-dimensional echocardiography in patients with cardiac resynchronization therapy. *Arch.Cardiovasc.Dis.*, 102, 497-508.

Diab, I.G., Hunter, R.J., Kamdar, R., Berriman, T., Duncan, E., Richmond, L., Baker, V., Abrams, D., Earley, M.J., Sporton, S. & Schilling, R.J. (2011) Does ventricular dyssynchrony on echocardiography predict response to cardiac resynchronisation therapy? A randomised controlled study. *Heart*, 97, 1410-1416.

Diaz-Infante, E., Sitges, M., Vidal, B., Mont, L., Delgado, V., Marigliano, A., Macias, A., Tolosana, J.M., Tamborero, D., Azqueta, M., Roig, E., Pare, C. & Brugada, J. (2007) Usefulness of ventricular dyssynchrony measured using M-mode echocardiography to predict response to resynchronization therapy. *Am.J.Cardiol.*, 100, 84-89.

Dickstein, K., Cohen-Solal, A., Filippatos, G., McMurray, J.J., Ponikowski, P., Poole-Wilson, P.A., Stromberg, A., van Veldhuisen, D.J., Atar, D., Hoes, A.W., Keren, A.,

Mebazaa, A., Nieminen, M., Priori, S.G., Swedberg, K., Vahanian, A., Camm, J., De, C.R., Dean, V., Dickstein, K., Filippatos, G., Funck-Brentano, C., Hellemans, I., Kristensen, S.D., McGregor, K., Sechtem, U., Silber, S., Tendera, M., Widimsky, P. & Zamorano, J.L. (2008) ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *Eur.Heart J.*, 29, 2388-2442.

Dorlas, J.C. & Nijboer, J.A. (1985) Photo-electric plethysmography as a monitoring device in anaesthesia. Application and interpretation. *Br.J.Anaesth.*, 57, 524-530.

Duncan, A.M., Lim, E., Clague, J., Gibson, D.G. & Henein, M.Y. (2006) Comparison of segmental and global markers of dyssynchrony in predicting clinical response to cardiac resynchronization. *Eur Heart J.*, 27, 2426-2432.

Duvall, W.L., Hansalia, R., Wijetunga, M.N., Buckley, S. & Fischer, A. (2010) Advantage of optimizing V-V timing in cardiac resynchronization therapy devices. *Pacing Clin.Electrophysiol.*, 33, 1161-1168.

Ellenbogen, K.A., Gold, M.R., Meyer, T.E., Fernndez, L., I, Mittal, S., Waggoner, A.D., Lemke, B., Singh, J.P., Spinale, F.G., Van Eyk, J.E., Whitehill, J., Weiner, S., Bedi, M., Rapkin, J. & Stein, K.M. (2010) Primary results from the SmartDelay determined AV optimization: a comparison to other AV delay methods used in cardiac resynchronization therapy (SMART-AV) trial: a randomized trial comparing empirical, echocardiography-guided, and algorithmic atrioventricular delay programming in cardiac resynchronization therapy. *Circulation*, 122, 2660-2668.

Faletra, F.F., Conca, C., Klersy, C., Klimusina, J., Regoli, F., Mantovani, A., Pasotti, E., Pedrazzini, G.B., De, C.S., Moccetti, T. & Auricchio, A. (2009) Comparison of eight echocardiographic methods for determining the prevalence of mechanical

dyssynchrony and site of latest mechanical contraction in patients scheduled for cardiac resynchronization therapy. *Am.J.Cardiol.*, 103, 1746-1752.

Finegold, J.A., Manisty, C.H., Cecaro, F., Sutaria, N., Mayet, J. & Francis, D.P. (2012) Choosing between velocity-time-integral ratio and peak velocity ratio for calculation of the dimensionless index (or aortic valve area) in serial follow-up of aortic stenosis. *Int.J.Cardiol.*

Fischer, A., Hansalia, R., Buckley, S., Goldberg, R., Goldman, M., Muntner, P., Mehta, D. & Duvall, W.L. (2009) Lack of clinical predictors of optimal V-V delay in patients with cardiac resynchronization devices. *J.Interv.Card Electrophysiol.*, 25, 153-158.

Flachskampf, F.A. & Daniel, W.G. (2010) Cardiac imaging in the patient with chest pain: echocardiography. *Heart*, 96, 1063-1072.

Flather, M.D., Yusuf, S., Kober, L., Pfeffer, M., Hall, A., Murray, G., Torp-Pedersen, C., Ball, S., Pogue, J., Moye, L. & Braunwald, E. (2000) Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: a systematic overview of data from individual patients. ACE-Inhibitor Myocardial Infarction Collaborative Group. *Lancet*, 355, 1575-1581.

Foley, P.W., Patel, K., Irwin, N., Sanderson, J.E., Frenneaux, M.P., Smith, R.E., Stegemann, B. & Leyva, F. (2011) Cardiac resynchronisation therapy in patients with heart failure and a normal QRS duration: the RESPOND study. *Heart*, 97, 1041-1047.

Francis, D.P. (2011) Precision of a parabolic optimum calculated from noisy biological data, and implications for quantitative optimization of biventricular pacemakers (cardiac resynchronization therapy) . *Applied Mathematics*.

Francis, D.P., Coats, A.J. & Gibson, D.G. (1999) How high can a correlation coefficient be? Effects of limited reproducibility of common cardiological measures. *Int.J.Cardiol.*, 69, 185-189.

Geibel, A., Gornandt, L., Kasper, W. & Bubenheimer, P. (1991) Reproducibility of Doppler echocardiographic quantification of aortic and mitral valve stenoses: comparison between two echocardiography centers. *Am.J.Cardiol.*, 67, 1013-1021.

Ghio, S., Constantin, C., Klersy, C., Serio, A., Fontana, A., Campana, C. & Tavazzi, L. (2004) Interventricular and intraventricular dyssynchrony are common in heart failure patients, regardless of QRS duration. *Eur.Heart J.*, 25, 571-578.

Ghio, S., Freemantle, N., Scelsi, L., Serio, A., Magrini, G., Pasotti, M., Shankar, A., Cleland, J.G. & Tavazzi, L. (2009) Long-term left ventricular reverse remodelling with cardiac resynchronization therapy: results from the CARE-HF trial. *Eur J.Heart Fail.*, 11, 480-488.

Ghio, S., Freemantle, N., Serio, A., Magrini, G., Scelsi, L., Pasotti, M., Cleland, J.G. & Tavazzi, L. (2006) Baseline echocardiographic characteristics of heart failure patients enrolled in a large European multicentre trial (CArdiac REsynchronisation Heart Failure study). *Eur J.Echocardiogr.*, 7, 373-378.

Gimenes, V.M., Vieira, M.L., Andrade, M.M., Pinheiro, J., Jr., Hotta, V.T. & Mathias, W., Jr. (2008) Standard values for real-time transthoracic three-dimensional echocardiographic dyssynchrony indexes in a normal population. *J.Am.Soc.Echocardiogr.*, 21, 1229-1235.

Gold, M.R., Niazi, I., Giudici, M., Leman, R.B., Sturdivant, J.L., Kim, M.H., Yu, Y., Ding, J. & Waggoner, A.D. (2007) A prospective comparison of AV delay programming methods for hemodynamic optimization during cardiac resynchronization therapy. *J.Cardiovasc.Electrophysiol.*, 18, 490-496.

Gorcsan, J., Abraham, T., Agler, D.A., Bax, J.J., Derumeaux, G., Grimm, R.A., Martin, R., Steinberg, J.S., Sutton, M.S. & Yu, C.M. (2008) Echocardiography for cardiac resynchronization therapy: recommendations for performance and reporting--a report from the American Society of Echocardiography Dyssynchrony Writing Group endorsed by the Heart Rhythm Society. *J.Am.Soc.Echocardiogr.*, 21, 191-213.

Gorcsan, J., Kanzaki, H., Bazaz, R., Dohi, K. & Schwartzman, D. (2004a) Usefulness of echocardiographic tissue synchronization imaging to predict acute response to cardiac resynchronization therapy. *Am.J.Cardiol.*, 93, 1178-1181.

Gorcsan, J., Lang, R.M., Picard, M.H., Zoghbi, W.A., Frommelt, P.C. & Gillam, L.D. (2004b) Meeting highlights of the 15th annual Scientific Sessions of the American Society of Echocardiography: June 26 to 30, 2004. *J.Am.Coll.Cardiol.*, 44, 2111-2116.

Gorcsan, J., Tanabe, M., Bleeker, G.B., Suffoletto, M.S., Thomas, N.C., Saba, S., Tops, L.F., Schalij, M.J. & Bax, J.J. (2007) Combined longitudinal and radial dyssynchrony predicts ventricular response after resynchronization therapy. *J.Am.Coll.Cardiol.*, 50, 1476-1483.

Gras, D., Leclercq, C., Tang, A.S., Bucknall, C., Luttikhuis, H.O. & Kirstein-Pedersen, A. (2002) Cardiac resynchronization therapy in advanced heart failure the multicenter InSync clinical study. *Eur.J.Heart Fail.*, 4, 311-320.

Hawkins, N.M., Petrie, M.C., Burgess, M.I. & McMurray, J.J. (2009) Selecting patients for cardiac resynchronization therapy: the fallacy of echocardiographic dyssynchrony. *J.Am.Coll.Cardiol.*, 53, 1944-1959.

Hawkins, N.M., Wang, D., McMurray, J.J., Pfeffer, M.A., Swedberg, K., Granger, C.B., Yusuf, S., Pocock, S.J., Ostergren, J., Michelson, E.L. & Dunn, F.G. (2007) Prevalence and prognostic impact of bundle branch block in patients with heart failure: evidence from the CHARM programme. *Eur.J.Heart Fail.*, 9, 510-517.

Higgins, S.L., Hummel, J.D., Niazi, I.K., Giudici, M.C., Worley, S.J., Saxon, L.A., Boehmer, J.P., Higginbotham, M.B., De, M.T., Foster, E. & Yong, P.G. (2003) Cardiac resynchronization therapy for the treatment of heart failure in patients with intraventricular conduction delay and malignant ventricular tachyarrhythmias. *J.Am.Coll.Cardiol.*, 42, 1454-1459.

Hochleitner, M., Hortnagl, H., Ng, C.K., Hortnagl, H., Gschnitzer, F. & Zechmann, W. (1990) Usefulness of physiologic dual-chamber pacing in drug-resistant idiopathic dilated cardiomyopathy. *Am.J.Cardiol.*, 66, 198-202.

Holloway, C.J., Edwards, L.M., Rider, O.J., Fast, A., Clarke, K., Francis, J.M., Myerson, S.G. & Neubauer, S. (2011) A comparison of visual and quantitative assessment of left ventricular ejection fraction by cardiac magnetic resonance. *Int.J.Cardiovasc.Imaging*, 27, 563-569.

Hutcheon, J.A., Chiolero, A. & Hanley, J.A. (2010) Random measurement error and regression dilution bias. *BMJ*, 340, c2289.

Imholz, B.P., Wieling, W., van Montfrans, G.A. & Wesseling, K.H. (1998) Fifteen years experience with finger arterial pressure monitoring: assessment of the technology. *Cardiovasc.Res.*, 38, 605-616.

Inoue, N., Ishikawa, T., Sumita, S., Nakagawa, T., Kobayashi, T., Matsushita, K., Matsumoto, K., Ohkusu, Y., Taima, M., Kosuge, M., Uchino, K., Kimura, K. & Umemura, S. (2005a) Long-term follow-up of atrioventricular delay optimization in patients with biventricular pacing. *Circ.J.*, 69, 201-204.

Inoue, N., Ishikawa, T., Sumita, S., Nakagawa, T., Kobayashi, T., Matsushita, K., Matsumoto, K., Ohkusu, Y., Taima, M., Kosuge, M., Uchino, K., Kimura, K. & Umemura, S. (2005b) Long-term follow-up of atrioventricular delay optimization in patients with biventricular pacing. *Circ.J.*, 69, 201-204.

Jansen, A.H., Bracke, F.A., van Dantzig, J.M., Meijer, A., van der Voort, P.H., Aarnoudse, W., van Gelder, B.M. & Peels, K.H. (2006a) Correlation of echo-Doppler optimization of atrioventricular delay in cardiac resynchronization therapy with invasive hemodynamics in patients with heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. *Am.J.Cardiol.*, 97, 552-557.

Jansen, A.H., Bracke, F.A., van Dantzig, J.M., Meijer, A., van der Voort, P.H., Aarnoudse, W., van Gelder, B.M. & Peels, K.H. (2006b) Correlation of echo-Doppler optimization of atrioventricular delay in cardiac resynchronization therapy with invasive hemodynamics in patients with heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. *Am.J.Cardiol.*, 97, 552-557.

Jenkins, C., Moir, S., Chan, J., Rakhit, D., Haluska, B. & Marwick, T.H. (2009) Left ventricular volume measurement with echocardiography: a comparison of left ventricular opacification, three-dimensional echocardiography, or both with magnetic resonance imaging. *Eur Heart J.*, 30, 98-106.

Kapetanakis, S., Bhan, A., Murgatroyd, F., Kearney, M.T., Gall, N., Zhang, Q., Yu, C.M. & Monaghan, M.J. (2011) Real-time 3D echo in patient selection for cardiac resynchronization therapy. *JACC.Cardiovasc.Imaging*, 4, 16-26.

Kapetanakis, S., Kearney, M.T., Siva, A., Gall, N., Cooklin, M. & Monaghan, M.J. (2005) Real-time three-dimensional echocardiography: a novel technique to quantify global left ventricular mechanical dyssynchrony. *Circulation*, 112, 992-1000.

Kass, D.A., Chen, C.H., Curry, C., Talbot, M., Berger, R., Fetics, B. & Nevo, E. (1999) Improved left ventricular mechanics from acute VDD pacing in patients with dilated cardiomyopathy and ventricular conduction delay. *Circulation*, 99, 1567-1573.

Kaufman, C.L., Kaiser, D.R., Burns, K.V., Kelly, A.S. & Bank, A.J. (2010) Multi-plane mechanical dyssynchrony in cardiac resynchronization therapy. *Clin.Cardiol.*, 33, E31-E38.

Kuhl, H.P., Schreckenberg, M., Rulands, D., Katoh, M., Schafer, W., Schummers, G., Bucker, A., Hanrath, P. & Franke, A. (2004) High-resolution transthoracic real-time three-dimensional echocardiography: quantitation of cardiac volumes and function using semi-automatic border detection and comparison with cardiac magnetic resonance imaging. *J.Am.Coll.Cardiol.*, 43, 2083-2090.

Kyriacou, A., Pabari, P.A. & Francis, D.P. (2011) Cardiac resynchronization therapy is certainly cardiac therapy, but how much resynchronization and how much atrioventricular delay optimization? *Heart Fail.Rev.*

Lane, R.E., Chow, A.W., Chin, D. & Mayet, J. (2004) Selection and optimisation of biventricular pacing: the role of echocardiography. *Heart*, 90 Suppl 6, vi10-vi16.

Lane, R.E., Mayet, J., Peters, N.S., Davies, D.W. & Chow, A.W. (2008) Comparison of temporary bifocal right ventricular pacing and biventricular pacing for heart failure: evaluation by tissue Doppler imaging. *Heart*, 94, 53-58.

Lansdorp, B., Ouweneel, D., de, K.A., van der Hoeven, J.G., Lemson, J. & Pickkers, P. (2011) Non-invasive measurement of pulse pressure variation and systolic pressure variation using a finger cuff corresponds with intra-arterial measurement. *Br.J.Anaesth.*

Le Cam L. (1986) The Central Limit Theorem Around 1935. *Statistical Science*, 1, 78-91.

Lehmann, K.G. & Platt, M.S. (1999) Improved accuracy and precision of thermodilution cardiac output measurement using a dual thermistor catheter system. *J.Am.Coll.Cardiol.*, 33, 883-891.

Linde, C., Abraham, W.T., Gold, M.R. & Daubert, C. (2010) Cardiac resynchronization therapy in asymptomatic or mildly symptomatic heart failure patients in relation to etiology: results from the REVERSE (REsynchronization

reVErses Remodeling in Systolic Left vEntricular Dysfunction) study. *J.Am.Coll.Cardiol.*, 56, 1826-1831.

Linde, C., Abraham, W.T., Gold, M.R., St John, S.M., Ghio, S. & Daubert, C. (2008) Randomized trial of cardiac resynchronization in mildly symptomatic heart failure patients and in asymptomatic patients with left ventricular dysfunction and previous heart failure symptoms. *J.Am.Coll.Cardiol.*, 52, 1834-1843.

Linde, C., Leclercq, C., Rex, S., Garrigue, S., Lavergne, T., Cazeau, S., McKenna, W., Fitzgerald, M., Deharo, J.C., Alonso, C., Walker, S., Braunschweig, F., Bailleul, C. & Daubert, J.C. (2002) Long-term benefits of biventricular pacing in congestive heart failure: results from the MUltisite STimulation in cardiomyopathy (MUSTIC) study. *J.Am.Coll.Cardiol.*, 40, 111-118.

Liodakis, E., Sharef, O.A., Dawson, D. & Nihoyannopoulos, P. (2009) The use of real-time three-dimensional echocardiography for assessing mechanical synchronicity. *Heart*, 95, 1865-1871.

Lopez-Candales, A. & Edelman, K. (2011) Ratio of right to left ventricular ejection: a pilot study using Doppler to detect interventricular dyssynchrony. *Clin.Cardiol.*, 34, 366-371.

Lubitz, S.A., Leong-Sit, P., Fine, N., Kramer, D.B., Singh, J. & Ellinor, P.T. (2010) Effectiveness of cardiac resynchronization therapy in mild congestive heart failure: systematic review and meta-analysis of randomized trials. *Eur J.Heart Fail.*, 12, 360-366.

MacRoberts MH & MacRoberts BR. (1984) The Negontional Reference: Or the Art of Dissembling. *Social Studies of Science*, 14, 91-94.

Marcus, G.M., Rose, E., Viloria, E.M., Schafer, J., De, M.T., Saxon, L.A. & Foster, E. (2005) Septal to posterior wall motion delay fails to predict reverse remodeling or

clinical improvement in patients undergoing cardiac resynchronization therapy. *J.Am.Coll.Cardiol.*, 46, 2208-2214.

Marsan, N.A., Bleeker, G.B., Ypenburg, C., Ghio, S., Van de Veire, N.R., Holman, E.R., van der Wall, E.E., Tavazzi, L., Schalij, M.J. & Bax, J.J. (2008) Real-time three-dimensional echocardiography permits quantification of left ventricular mechanical dyssynchrony and predicts acute response to cardiac resynchronization therapy. *J.Cardiovasc.Electrophysiol.*, 19, 392-399.

McGowan, J.H. & Cleland, J.G. (2003) Reliability of reporting left ventricular systolic function by echocardiography: a systematic review of 3 methods. *Am.Heart J.*, 146, 388-397.

Mele, D., Pasanisi, G., Capasso, F., De, S.A., Morales, M.A., Poggio, D., Capucci, A., Tabacchi, G., Sallusti, L. & Ferrari, R. (2006) Left intraventricular myocardial deformation dyssynchrony identifies responders to cardiac resynchronization therapy in patients with heart failure. *Eur Heart J.*, 27, 1070-1078.

Minners, J., Allgeier, M., Gohlke-Baerwolf, C., Kienzle, R.P., Neumann, F.J. & Jander, N. (2008) Inconsistencies of echocardiographic criteria for the grading of aortic valve stenosis. *Eur.Heart J.*, 29, 1043-1048.

Miyazaki, C., Redfield, M.M., Powell, B.D., Lin, G.M., Herges, R.M., Hodge, D.O., Olson, L.J., Hayes, D.L., Espinosa, R.E., Rea, R.F., Bruce, C.J., Nelson, S.M., Miller, F.A. & Oh, J.K. (2010) Dyssynchrony indices to predict response to cardiac resynchronization therapy: a comprehensive prospective single-center study. *Circ.Heart Fail.*, 3, 565-573.

Monaghan, M.J. (2006) Role of real time 3D echocardiography in evaluating the left ventricle. *Heart*, 92, 131-136.

Mortensen, P.T., Sogaard, P., Mansour, H., Ponsonaille, J., Gras, D., Lazarus, A., Reiser, W., Alonso, C., Linde, C.M., Lunati, M., Kramm, B. & Harrison, E.M. (2004) Sequential biventricular pacing: evaluation of safety and efficacy. *Pacing Clin.Electrophysiol.*, 27, 339-345.

Moss, A.J., Hall, W.J., Cannom, D.S., Klein, H., Brown, M.W., Daubert, J.P., Estes, N.A., III, Foster, E., Greenberg, H., Higgins, S.L., Pfeffer, M.A., Solomon, S.D., Wilber, D. & Zareba, W. (2009a) Cardiac-resynchronization therapy for the prevention of heart-failure events. *N.Engl.J.Med.*, 361, 1329-1338.

Moss, A.J., Hall, W.J., Cannom, D.S., Klein, H., Brown, M.W., Daubert, J.P., Estes, N.A., III, Foster, E., Greenberg, H., Higgins, S.L., Pfeffer, M.A., Solomon, S.D., Wilber, D. & Zareba, W. (2009b) Cardiac-resynchronization therapy for the prevention of heart-failure events. *N.Engl.J.Med.*, 361, 1329-1338.

Muller, R. & Buttner, P. (1994) A critical discussion of intraclass correlation coefficients. *Stat.Med.*, 13, 2465-2476.

Nelson, G.S., Berger, R.D., Fetics, B.J., Talbot, M., Spinelli, J.C., Hare, J.M. & Kass, D.A. (2000) Left ventricular or biventricular pacing improves cardiac function at diminished energy cost in patients with dilated cardiomyopathy and left bundle-branch block. *Circulation*, 102, 3053-3059.

Nishimura, R.A., Hayes, D.L., Holmes, D.R., Jr. & Tajik, A.J. (1995) Mechanism of hemodynamic improvement by dual-chamber pacing for severe left ventricular dysfunction: an acute Doppler and catheterization hemodynamic study. *J.Am.Coll.Cardiol.*, 25, 281-288.

Norisada, K., Kawai, H., Tanaka, H., Tatsumi, K., Onishi, T., Fukuzawa, K., Yoshida, A. & Hirata, K. (2010) Myocardial contractile function in the region of the left ventricular pacing lead predicts the response to cardiac resynchronization therapy assessed by two-dimensional speckle tracking echocardiography. *J.Am.Soc.Echocardiogr.*, 23, 181-189.

Notabartolo, D., Merlino, J.D., Smith, A.L., DeLurgio, D.B., Vera, F.V., Easley, K.A., Martin, R.P. & Leon, A.R. (2004) Usefulness of the peak velocity difference by tissue Doppler imaging technique as an effective predictor of response to cardiac resynchronization therapy. *Am.J.Cardiol.*, 94, 817-820.

Otterstad, J.E., Froeland, G., St John, S.M. & Holme, I. (1997) Accuracy and reproducibility of biplane two-dimensional echocardiographic measurements of left ventricular dimensions and function. *Eur Heart J.*, 18, 507-513.

Pabari PA, Moraldo M, Mayet J, Kyriacou A, Francis DP & Manisty CH. (2012) Validating markers of mechanical dyssynchrony by experimental manipulation of interventricular timings: what is needed to make them a reasonable prospect for cardiac resynchronisation therapy selection? *Heart*, 98.

Pabari, P.A., Willson, K., Stegemann, B., van Geldorp, I.E., Kyriacou, A., Moraldo, M., Mayet, J., Hughes, A.D. & Francis, D.P. (2011) When is an optimization not an optimization? Evaluation of clinical implications of information content (signal-to-noise ratio) in optimization of cardiac resynchronization therapy, and how to measure and maximize it. *Heart Fail.Rev.*, 16, 277-290.

Palmieri, V., Russo, C., Buonomo, A., Cimmino, G., Tartaglione, D., Pezzullo, S. & Celentano, A. (2010) Test-re-test reproducibility of Doppler echocardiography for assessment of electromechanical dyssynchrony: implications for heart failure clinic. *J.Cardiol.*, 56, 271-279.

Park, H.E., Chang, S.A., Kim, H.K., Shin, D.H., Kim, J.H., Seo, M.K., Kim, Y.J., Cho, G.Y., Sohn, D.W., Oh, B.H. & Park, Y.B. (2010) Impact of loading condition on the 2D speckle tracking-derived left ventricular dyssynchrony index in nonischemic dilated cardiomyopathy. *Circ.Cardiovasc.Imaging*, 3, 272-281.

Parreira, L., Santos, J.F., Madeira, J., Mendes, L., Seixo, F., Caetano, F., Lopes, C., Venancio, J., Mateus, A., Ines, J.L. & Mendes, M. (2005) Cardiac resynchronization

therapy with sequential biventricular pacing: impact of echocardiography guided VV delay optimization on acute results. *Rev.Port.Cardiol.*, 24, 1355-1365.

Penaz, J. (1973) Photoelectric measurement of blood pressure, volume and flow in the finger. *Digest of the international Conference on Medicine and Biological Engineering*, 104.

Penicka, M., Bartunek, J., De, B.B., Vanderheyden, M., Goethals, M., De, Z.M., Brugada, P. & Geelen, P. (2004) Improvement of left ventricular function after cardiac resynchronization therapy is predicted by tissue Doppler imaging echocardiography. *Circulation*, 109, 978-983.

Penzel, T., Mayer, J., Peter, J.H., Podszus, T., Voigt, K.H. & von, W.P. (1992) Continuous non-invasive blood pressure monitoring in patients with sleep disorders. *Physiol Res.*, 41, 11-17.

Perego, G.B., Chianca, R., Facchini, M., Frattola, A., Balla, E., Zucchi, S., Cavaglia, S., Vicini, I., Negretto, M. & Osculati, G. (2003) Simultaneous vs. sequential biventricular pacing in dilated cardiomyopathy: an acute hemodynamic study. *Eur.J.Heart Fail.*, 5, 305-313.

Pitt, B., Poole-Wilson, P.A., Segal, R., Martinez, F.A., Dickstein, K., Camm, A.J., Konstam, M.A., Riegger, G., Klinger, G.H., Neaton, J., Sharma, D. & Thiyagarajan, B. (2000) Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial--the Losartan Heart Failure Survival Study ELITE II. *Lancet*, 355, 1582-1587.

Pitt, B., Williams, G., Remme, W., Martinez, F., Lopez-Sendon, J., Zannad, F., Neaton, J., Roniker, B., Hurley, S., Burns, D., Bittman, R. & Kleiman, J. (2001) The EPHESUS trial: eplerenone in patients with heart failure due to systolic dysfunction complicating acute myocardial infarction. Eplerenone Post-AMI Heart Failure Efficacy and Survival Study. *Cardiovasc.Drugs Ther.*, 15, 79-87.

Pitt, B., Zannad, F., Remme, W.J., Cody, R., Castaigne, A., Perez, A., Palensky, J. & Wittes, J. (1999) The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N.Engl.J.Med.*, 341, 709-717.

Pitzalis, M.V., Iacoviello, M., Romito, R., Guida, P., De, T.E., Luzzi, G., Anaclerio, M., Forleo, C. & Rizzon, P. (2005) Ventricular asynchrony predicts a better outcome in patients with chronic heart failure receiving cardiac resynchronization therapy. *J.Am.Coll.Cardiol.*, 45, 65-69.

Pitzalis, M.V., Iacoviello, M., Romito, R., Massari, F., Rizzon, B., Luzzi, G., Guida, P., Andriani, A., Mastropasqua, F. & Rizzon, P. (2002) Cardiac resynchronization therapy tailored by echocardiographic evaluation of ventricular asynchrony. *J.Am.Coll.Cardiol.*, 40, 1615-1622.

Pope, B.B. (2002) How to perform 3- or 5-lead monitoring. *Nursing.*, 32, 50-52.

Porciani, M.C., Dondina, C., Macioce, R., Demarchi, G., Cappelli, F., Lilli, A., Pappone, A., Ricciardi, G., Colombo, P.C., Padeletti, M., Jelic, S. & Padeletti, L. (2006a) Temporal variation in optimal atrioventricular and interventricular delay during cardiac resynchronization therapy. *J.Card Fail.*, 12, 715-719.

Porciani, M.C., Lilli, A., Macioce, R., Cappelli, F., Demarchi, G., Pappone, A., Ricciardi, G. & Padeletti, L. (2006b) Utility of a new left ventricular asynchrony index as a predictor of reverse remodelling after cardiac resynchronization therapy. *Eur Heart J.*, 27, 1818-1823.

Raphael CE., Kyriacou, A., Pabari, P.A., Cole G, Baruah R, Hughes, A.D. & Francis, D.P. (2011) Multinational evaluation of the interpretability of the iterative method of optimization of AV delay for CRT. *ESC.*

Reuter, S., Garrigue, S., Barold, S.S., Jais, P., Hocini, M., Haissaguerre, M. & Clementy, J. (2002) Comparison of characteristics in responders versus nonresponders with biventricular pacing for drug-resistant congestive heart failure. *Am.J.Cardiol.*, 89, 346-350.

Ritter, P., Padeletti, L., Gillio-Meina, L. & Gaggini, G. (1999) Determination of the optimal atrioventricular delay in DDD pacing. Comparison between echo and peak endocardial acceleration measurements. *Europace.*, 1, 126-130.

Sassone, B., Capecchi, A., Boggian, G., Gabrieli, L., Sacca, S., Vandelli, R., Petracci, E. & Mele, D. (2007) Value of baseline left lateral wall postsystolic displacement assessed by M-mode to predict reverse remodeling by cardiac resynchronization therapy. *Am.J.Cardiol.*, 100, 470-475.

Sawhney, N.S., Waggoner, A.D., Garhwal, S., Chawla, M.K., Osborn, J. & Faddis, M.N. (2004) Randomized prospective trial of atrioventricular delay programming for cardiac resynchronization therapy. *Heart Rhythm.*, 1, 562-567.

Scharf, C., Li, P., Muntwyler, J., Chugh, A., Oral, H., Pelosi, F., Morady, F. & Armstrong, W.F. (2005) Rate-dependent AV delay optimization in cardiac resynchronization therapy. *Pacing Clin.Electrophysiol.*, 28, 279-284.

Schutte, A.E., Huisman, H.W., van Rooyen, J.M., Malan, N.T. & Schutte, R. (2004) Validation of the Finometer device for measurement of blood pressure in black women. *J.Hum.Hypertens.*, 18, 79-84.

Shamim, W., Francis, D.P., Yousufuddin, M., Varney, S., Pieopli, M.F., Anker, S.D. & Coats, A.J. (1999) Intraventricular conduction delay: a prognostic marker in chronic heart failure. *Int.J.Cardiol.*, 70, 171-178.

Smith, N.T., Wesseling, K.H. & de, W.B. (1985) Evaluation of two prototype devices producing noninvasive, pulsatile, calibrated blood pressure measurement from a finger. *J.Clin.Monit.*, 1, 17-29.

Sogaard, P., Egeblad, H., Pedersen, A.K., Kim, W.Y., Kristensen, B.O., Hansen, P.S. & Mortensen, P.T. (2002) Sequential versus simultaneous biventricular resynchronization for severe heart failure: evaluation by tissue Doppler imaging. *Circulation*, 106, 2078-2084.

Soliman, O.I., Geleijnse, M.L., Theuns, D.A., van Dalen, B.M., Vletter, W.B., Jordaens, L.J., Metawei, A.K., Al-Amin, A.M. & Ten Cate, F.J. (2009) Usefulness of left ventricular systolic dyssynchrony by real-time three-dimensional echocardiography to predict long-term response to cardiac resynchronization therapy. *Am.J.Cardiol.*, 103, 1586-1591.

Soliman, O.I., Krenning, B.J., Geleijnse, M.L., Nemes, A., Bosch, J.G., van Geuns, R.J., Kirschbaum, S.W., Anwar, A.M., Galema, T.W., Vletter, W.B. & Ten Cate, F.J. (2007a) Quantification of left ventricular volumes and function in patients with cardiomyopathies by real-time three-dimensional echocardiography: a head-to-head comparison between two different semiautomated endocardial border detection algorithms. *J.Am.Soc.Echocardiogr.*, 20, 1042-1049.

Soliman, O.I., Theuns, D.A., Geleijnse, M.L., Anwar, A.M., Nemes, A., Caliskan, K., Vletter, W.B., Jordaens, L.J. & Cate, F.J. (2007b) Spectral pulsed-wave tissue Doppler imaging lateral-to-septal delay fails to predict clinical or echocardiographic outcome after cardiac resynchronization therapy. *Europace.*, 9, 113-118.

Solomon, S.D., Foster, E., Bourgoun, M., Shah, A., Viloria, E., Brown, M.W., Hall, W.J., Pfeffer, M.A. & Moss, A.J. (2010) Effect of cardiac resynchronization therapy on reverse remodeling and relation to outcome: multicenter automatic defibrillator implantation trial: cardiac resynchronization therapy. *Circulation*, 122, 985-992.

St John Sutton, M.G., Plappert, T., Abraham, W.T., Smith, A.L., DeLurgio, D.B., Leon, A.R., Loh, E., Kocovic, D.Z., Fisher, W.G., Ellestad, M., Messenger, J., Kruger, K., Hilpisch, K.E. & Hill, M.R. (2003) Effect of cardiac resynchronization therapy on left ventricular size and function in chronic heart failure. *Circulation*, 107, 1985-1990.

Stewart, R.A., Young, A.A., Anderson, C., Teo, K.K., Jennings, G. & Cowan, B.R. (2011) Relationship between QRS duration and left ventricular mass and volume in patients at high cardiovascular risk. *Heart*.

Suffoletto, M.S., Dohi, K., Cannesson, M., Saba, S. & Gorcsan, J., III. (2006) Novel speckle-tracking radial strain from routine black-and-white echocardiographic images to quantify dyssynchrony and predict response to cardiac resynchronization therapy. *Circulation*, 113, 960-968.

Sutton, M.G., Plappert, T., Hilpisch, K.E., Abraham, W.T., Hayes, D.L. & Chinchoy, E. (2006a) Sustained reverse left ventricular structural remodeling with cardiac resynchronization at one year is a function of etiology: quantitative Doppler echocardiographic evidence from the Multicenter InSync Randomized Clinical Evaluation (MIRACLE). *Circulation*, 113, 266-272.

Sutton, M.G., Plappert, T., Hilpisch, K.E., Abraham, W.T., Hayes, D.L. & Chinchoy, E. (2006b) Sustained reverse left ventricular structural remodeling with cardiac resynchronization at one year is a function of etiology: quantitative Doppler echocardiographic evidence from the Multicenter InSync Randomized Clinical Evaluation (MIRACLE). *Circulation*, 113, 266-272.

Tamborero, D., Vidal, B., Tolosana, J.M., Sitges, M., Berruezo, A., Silva, E., Castel, M., Matas, M., Arbelo, E., Rios, J., Villacastin, J., Brugada, J. & Mont, L. (2011) Electrocardiographic versus Echocardiographic Optimization of the Interventricular Pacing Delay in Patients Undergoing Cardiac Resynchronization Therapy. *J.Cardiovasc.Electrophysiol.*

Thomas, D.E., Yousef, Z.R. & Fraser, A.G. (2009) A critical comparison of echocardiographic measurements used for optimizing cardiac resynchronization therapy: stroke distance is best. *Eur.J.Heart Fail.*, 11, 779-788.

Tribouilloy, C., Shen, W.F., Rey, J.L., Adam, M.C. & Lesbre, J.P. (1994) Mitral to aortic velocity-time integral ratio. A non-geometric pulsed-Doppler regurgitant index in isolated pure mitral regurgitation. *Eur.Heart J.*, 15, 1335-1339.

Turcott, R.G. & Pavek, T.J. (2008) Hemodynamic sensing using subcutaneous photoplethysmography. *Am.J.Physiol Heart Circ.Physiol*, 295, H2560-H2572.

Turcott, R.G., Witteles, R.M., Wang, P.J., Vagelos, R.H., Fowler, M.B. & Ashley, E.A. (2010) Measurement precision in the optimization of cardiac resynchronization therapy. *Circ.Heart Fail.*, 3, 395-404.

Valzania, C., Eriksson, M.J., Boriani, G. & Gadler, F. (2008) Cardiac resynchronization therapy during rest and exercise: comparison of two optimization methods. *Europace.*, 10, 1161-1169.

van Bommel, R.J., Gorcsan, J., III, Chung, E.S., Abraham, W.T., Gjestvang, F.T., Leclercq, C., Monaghan, M.J., Nihoyannopoulos, P., Peraldo, C., Yu, C.M., Demas, M., Gerritse, B. & Bax, J.J. (2010a) Effects of cardiac resynchronisation therapy in patients with heart failure having a narrow QRS Complex enrolled in PROSPECT. *Heart*, 96, 1107-1113.

van Bommel, R.J., Ypenburg, C., Borleffs, C.J., Delgado, V., Marsan, N.A., Bertini, M., Holman, E.R., Schalij, M.J. & Bax, J.J. (2010b) Value of tissue Doppler echocardiography in predicting response to cardiac resynchronization therapy in patients with heart failure. *Am.J.Cardiol.*, 105, 1153-1158.

Van de Veire, N.R., Bleeker, G.B., De, S.J., Ypenburg, C., Holman, E.R., van der Wall, E.E., Schalij, M.J. & Bax, J.J. (2007) Tissue synchronisation imaging

accurately measures left ventricular dyssynchrony and predicts response to cardiac resynchronisation therapy. *Heart*, 93, 1034-1039.

van Gelder, B.M., Bracke, F.A., Meijer, A., Lakerveld, L.J. & Pijls, N.H. (2004) Effect of optimizing the VV interval on left ventricular contractility in cardiac resynchronization therapy. *Am.J.Cardiol.*, 93, 1500-1503.

van Geldorp, I.E., Delhaas, T., Hermans, B., Vernooy, K., Broers, B., Klimusina, J., Regoli, F., Faletra, F.F., Moccetti, T., Gerritse, B., Cornelussen, R., Settels, J.J., Crijns, H.J., Auricchio, A. & Prinzen, F.W. (2011) Comparison of a non-invasive arterial pulse contour technique and echo Doppler aorta velocity-time integral on stroke volume changes in optimization of cardiac resynchronization therapy. *Europace.*, 13, 87-95.

van, E.J., Hasenbos, M. & Crul, J.F. (1985) Invasive v. non-invasive measurement of arterial pressure. Comparison of two automatic methods and simultaneously measured direct intra-arterial pressure. *Br.J.Anaesth.*, 57, 434-444.

Vanderheyden, M., De, B.T., Rivero-Ayerza, M., Geelen, P., Bartunek, J., Verstreken, S., De, Z.M. & Goethals, M. (2005) Tailored echocardiographic interventricular delay programming further optimizes left ventricular performance after cardiac resynchronization therapy. *Heart Rhythm.*, 2, 1066-1072.

Vesely, M.R., Li, S., Kop, W.J., Reese, A., Marshall, J., Shorofsky, S.R., Gottlieb, S.S., Mehra, M.R. & Gottdiener, J.S. (2008) Test-retest reliability of assessment for intraventricular dyssynchrony by tissue Doppler imaging echocardiography. *Am.J.Cardiol.*, 101, 645-650.

Vidal, B., Tamborero, D., Mont, L., Sitges, M., Delgado, V., Berruezo, A., Diaz-Infante, E., Tolosana, J.M., Pare, C. & Brugada, J. (2007) Electrocardiographic optimization of interventricular delay in cardiac resynchronization therapy: a simple method to optimize the device. *J.Cardiovasc.Electrophysiol.*, 18, 1252-1257.

Wang, C.L., Wu, C.T., Yeh, Y.H., Wu, L.S., Chang, C.J., Ho, W.J., Hsu, L.A., Luqman, N. & Kuo, C.T. (2010) Recoordination rather than resynchronization predicts reverse remodeling after cardiac resynchronization therapy. *J.Am.Soc.Echocardiogr.*, 23, 611-620.

Wesseling, K.H., Settels, J.J., van der Hoeven, G.M., Nijboer, J.A., Butijn, M.W. & Dorlas, J.C. (1985) Effects of peripheral vasoconstriction on the measurement of blood pressure in a finger. *Cardiovasc.Res.*, 19, 139-145.

Whinnett, Z.I., Davies, J.E., Nott, G., Willson, K., Manisty, C.H., Peters, N.S., Kanagaratnam, P., Davies, D.W., Hughes, A.D., Mayet, J. & Francis, D.P. (2008) Efficiency, reproducibility and agreement of five different hemodynamic measures for optimization of cardiac resynchronization therapy. *Int.J.Cardiol.*, 129, 216-226.

Whinnett, Z.I., Davies, J.E., Willson, K., Chow, A.W., Foale, R.A., Davies, D.W., Hughes, A.D., Francis, D.P. & Mayet, J. (2006a) Determination of optimal atrioventricular delay for cardiac resynchronization therapy using acute non-invasive blood pressure. *Europace.*, 8, 358-366.

Whinnett, Z.I., Davies, J.E., Willson, K., Manisty, C.H., Chow, A.W., Foale, R.A., Davies, D.W., Hughes, A.D., Mayet, J. & Francis, D.P. (2006b) Haemodynamic effects of changes in atrioventricular and interventricular delay in cardiac resynchronisation therapy show a consistent pattern: analysis of shape, magnitude and relative importance of atrioventricular and interventricular delay. *Heart*, 92, 1628-1634.

Whinnett, Z.I., Nott, G., Davies, J.E., Willson, K., Manisty, C.H., Kanagaratnam, P., Peters, N.S., Davies, D.W., Hughes, A.D., Mayet, J. & Francis, D.P. (2011) Maximizing efficiency of alternation algorithms for hemodynamic optimization of the AV delay of cardiac resynchronization therapy. *Pacing Clin.Electrophysiol.*, 34, 217-225.

Wikstrom, G., Blomstrom-Lundqvist, C., Andren, B., Lonnerholm, S., Blomstrom, P., Freemantle, N., Remp, T. & Cleland, J.G. (2009) The effects of aetiology on outcome in patients treated with cardiac resynchronization therapy in the CARE-HF trial. *Eur Heart J.*, 30, 782-788.

Xiao, H.B., Lee, C.H. & Gibson, D.G. (1991) Effect of left bundle branch block on diastolic function in dilated cardiomyopathy. *Br.Heart J.*, 66, 443-447.

Young, J.B., Abraham, W.T., Smith, A.L., Leon, A.R., Lieberman, R., Wilkoff, B., Canby, R.C., Schroeder, J.S., Liem, L.B., Hall, S. & Wheelan, K. (2003) Combined cardiac resynchronization and implantable cardioversion defibrillation in advanced chronic heart failure: the MIRACLE ICD Trial. *JAMA*, 289, 2685-2694.

Yu, C.M., Bax, J.J., Monaghan, M. & Nihoyannopoulos, P. (2004a) Echocardiographic evaluation of cardiac dyssynchrony for predicting a favourable response to cardiac resynchronisation therapy. *Heart*, 90 Suppl 6, vi17-vi22.

Yu, C.M., Bax, J.J., Monaghan, M. & Nihoyannopoulos, P. (2004b) Echocardiographic evaluation of cardiac dyssynchrony for predicting a favourable response to cardiac resynchronisation therapy. *Heart*, 90 Suppl 6, vi17-vi22.

Yu, C.M., Chau, E., Sanderson, J.E., Fan, K., Tang, M.O., Fung, W.H., Lin, H., Kong, S.L., Lam, Y.M., Hill, M.R. & Lau, C.P. (2002) Tissue Doppler echocardiographic evidence of reverse remodeling and improved synchronicity by simultaneously delaying regional contraction after biventricular pacing therapy in heart failure. *Circulation*, 105, 438-445.

Yu, C.M., Fung, J.W., Zhang, Q., Chan, C.K., Chan, Y.S., Lin, H., Kum, L.C., Kong, S.L., Zhang, Y. & Sanderson, J.E. (2004c) Tissue Doppler imaging is superior to strain rate imaging and postsystolic shortening on the prediction of reverse remodeling in both ischemic and nonischemic heart failure after cardiac resynchronization therapy. *Circulation*, 110, 66-73.

Yu, C.M., Fung, W.H., Lin, H., Zhang, Q., Sanderson, J.E. & Lau, C.P. (2003) Predictors of left ventricular reverse remodeling after cardiac resynchronization therapy for heart failure secondary to idiopathic dilated or ischemic cardiomyopathy. *Am.J.Cardiol.*, 91, 684-688.

Yu, C.M., Gorcsan, J., III, Bleeker, G.B., Zhang, Q., Schalij, M.J., Suffoletto, M.S., Fung, J.W., Schwartzman, D., Chan, Y.S., Tanabe, M. & Bax, J.J. (2007) Usefulness of tissue Doppler velocity and strain dyssynchrony for predicting left ventricular reverse remodeling response after cardiac resynchronization therapy. *Am.J.Cardiol.*, 100, 1263-1270.

Yu, C.M., Zhang, Q., Fung, J.W., Chan, H.C., Chan, Y.S., Yip, G.W., Kong, S.L., Lin, H., Zhang, Y. & Sanderson, J.E. (2005) A novel tool to assess systolic asynchrony and identify responders of cardiac resynchronization therapy by tissue synchronization imaging. *J.Am.Coll.Cardiol.*, 45, 677-684.

Yu, Y.M., Chen, M., Xiong, Y., Chau, M.M., Li, R.S. & Lau, T.K. (2009) Comparison of conventional and PureWave Crystal transducer in obstetric sonography. *J.Matern.Fetal Neonatal Med.*, 22, 616-621.

Zhang, Q., Fung, J.W., Chan, Y.S., Chan, H.C., Lin, H., Chan, S. & Yu, C.M. (2008) The role of repeating optimization of atrioventricular interval during interim and long-term follow-up after cardiac resynchronization therapy. *Int.J.Cardiol.*, 124, 211-217.

Zhang, Q., Liang, Y.J., Zhang, Q.H., Li, R.J., Chua, Y., Xie, J.M., Lee, P.W. & Yu, C.M. (2011) Impact of a Dedicated Training Program on the Reproducibility of Systolic Dyssynchrony Measures Using Tissue Doppler Imaging. *J.Am.Soc.Echocardiogr.*

Zuber, M., Toggweiler, S., Quinn-Tate, L., Brown, L., Amkieh, A. & Erne, P. (2008) A comparison of acoustic cardiography and echocardiography for optimizing pacemaker settings in cardiac resynchronization therapy. *Pacing Clin.Electrophysiol.*, 31, 802-811.