

Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning

A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of
Imperial College London

Konstantinos Nektarios Tzortzis

National Heart and Lung Institute
Imperial College London

July 2019

Declaration of originality

I declare that this submission is my own work and that it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Konstantinos Nektarios Tzortzis

July 2019

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC).

Under this license, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use this, or any derivative works, for a commercial purpose.

When reusing or sharing the work, ensure you make the license terms clear to others by naming the license and linking to the license text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this license or permitted under UK Copyright Law.

Abstract

The extracellular contact electrogram, which is the signature of the interaction of electrical activation and architecture of the local myocardium, is recorded clinically in contact with myocardium. The morphology of the signal could show relationships between the local electrogram and conduction abnormalities that influence the electrophysiology. In this thesis, I sought to address the hypothesis that the local electro-architecture, which is responsible for identifiable features of local action potentials, can be predicted from specific characteristics of electrogram recordings using supervised machine learning algorithms.

In addressing this hypothesis, I utilised *in vitro* multicellular preparations for obtaining unipolar electrogram data. The recordings were collected under a variety of experimental conditions, in order to investigate the effects of functional abnormalities, such as ion channel blockade and gap junction uncoupling, as well as structural determinants, such as increasing amounts of fibroblasts co-cultured with cardiac myocytes. A signal processing and feature extraction process was developed and applied on electrograms. The relationships between the abnormalities, which were introduced to experimental models, and specific electrogram characteristics were then investigated. Electrograms were then used inversely for the development of prediction models. To demonstrate the translational potential of these tools, they were tested on tissue slices derived from human end-stage heart failure hearts. It was found that EGM morphology was significantly modified due to the different heart failure phenotypes. These differences in morphology allowed accurate predictions. Paced data were also obtained from patients with a history of persistent AF.

The functional and structural determinants of unipolar electrogram morphology, which are also responsible for a variety of cardiac arrhythmias, can be predicted accurately using supervised machine learning. By better understanding the role of electro-architecture on electrogram morphology and utilising machine learning, we are provided with new insights that could contribute to a progress in diagnostics and treatment of cardiac diseases.

To my parents and beloved sister

Acknowledgements

The research presented in this thesis would not have been possible without the contribution from a number of people. I would like to thank them for their help and support.

Firstly, I would like to thank my supervisors Professor Nicholas Peters, Professor Spencer Sherwin, Dr Rasheda Chowdhury and Dr Chris Cantwell for their support and guidance over the past four years. I am especially grateful to Rasheda and Chris for their fantastic supervision throughout this period. Their expertise and guidance were essential at each stage of the project.

I am also grateful to the British Heart Foundation Centre of Research Excellence at Imperial College London for their financial support and training programme. It has been a privilege to be part of this multi-disciplinary programme.

I would also like to thank my undergraduate BSc students Kai Wang, Efthymoulos Sokratous, Vignesh Dhileepan, Alec Saunders and Angeliki Iakovou, as well as my summer students Binita Pandya and Beth Stella, who were willing and excited to work with me and helped to perform some of the experiments in this thesis.

I am indebted to the many people who have helped me with specific components of this work. Pravina Patel taught me the immunofluorescence and Stephen Rothery helped with the microscopy techniques. Filippo Perbellini and Samuel Watson taught me how to handle heart transplants and produce tissue slices. I am incredibly grateful to Norman for his clinical insights and careful collection of clinical data, and Emman and Fu for their many discussions and constructive comments. I would like to extend my appreciation to the computational group of the ElectroCardioMaths programme and especially Caroline, Yumnah, Rheeda, Stef and Kishan for their support and for making me appreciate the computational work and machine learning. Finally, I would like to thank the members of the Cardiac Function Section in NHLI for their contribution in cardiac myocyte isolation.

In particular, I would like to thank my London family, Nick and Eleni, for their years of warm friendship, their patience, support and encouragement. I cannot thank you enough for being my best friends, companions and daily support to this adventurous journey.

Finally, and most importantly, I would like to thank my family. My parents have always encouraged me to follow my dreams and try for the best. I greatly appreciate their unwavering love and support. The work produced in this thesis is a testament of all the sacrifice and support given by my parents. I owe you everything.

Konstantinos Nektarios Tzortzis



Publications and abstracts related to work undertaken

Publications

1. R.A. Chowdhury, **K.N. Tzortzis**, E. Dupont, S. Selvadurai, F. Perbellini, C.D. Cantwell, F.S. Ng, A. R. Simon, C.M. Terracciano, N.S. Peters. “Concurrent micro- to macro- cardiac electrophysiology in myocyte cultures and human heart slices”. *Sci. Rep.* 2018; 8(1):6947
2. C.D. Cantwell, Y. Mohamied, **K.N. Tzortzis**, S. Garasto, C. Houston, R.A. Chowdhury, F.S. Ng, A.A. Bharath, N.S. Peters. “Rethinking multiscale cardiac electrophysiology with machine learning and predictive modelling”. *Comput. Biol. Med.* 2019; 104:339-351

Abstracts

1. **K.N. Tzortzis**, R.A. Chowdhury, E. Dupont, S. Selvadurai, F. Perbellini, C.D. Cantwell, F.S. Ng, A.R. Simon, C.M. Terracciano, N.S. Peters. “Simultaneous optical mapping of transmembrane potential and electrogram recordings in human heart slices”. *BioMedEng18*, pp74, ISBN 978-1-9996465-0-9
 - BioMedEng18, London, UK, September 2018
2. **K.N. Tzortzis**, A.R. Simon, F. Perbellini, C.M. Terracciano, R.A. Chowdhury, N.S. Peters, C.D. Cantwell. “Supervised machine learning predicts fibrosis in *ex vivo* human cardiac slices using electrogram morphology”. *BioMedEng18*, pp391, ISBN 978-1-9996465-0-9
 - BioMedEng18, London, UK, September 2018
3. **K.N. Tzortzis**, A.R. Simon, F. Perbellini, C.M. Terracciano, C.D. Cantwell, N.S. Peters, R.A. Chowdhury. “Machine learning predicts fibrosis in *ex vivo* human cardiac slices using electrogram features but not voltage and fractionation”. *JICE 2018:51(Suppl 1):S56*
 - 14th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2018
4. **K.N. Tzortzis**, A.R. Simon, F. Perbellini, C.M. Terracciano, R.A. Chowdhury, N.S. Peters, C.D. Cantwell. “Heart failure phenotypes can be predicted in *ex vivo* human cardiac tissue slices using electrogram morphology and machine learning algorithms”. *JICE 2018:51(Suppl 1):S79*
 - 14th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2018
5. V. Dhileepan, **K.N. Tzortzis**, I. Mann, N.A. Qureshi, E. Lim, P. Kanagaratnam, C.D. Cantwell, N.S. Peters, R.A. Chowdhury. “Altering inter-electrode distance, but not orientation, affects bipolar electrogram morphology”. *JICE 2018:51(Suppl 1):S79*
 - 14th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2018

6. E. Sokratous, **K.N. Tzortzis**, N.A. Qureshi, P. Kanagaratnam, C.D. Cantwell, R.A. Chowdhury, N.S. Peters. “Multiple time- and frequency-domain electrogram features change with addition of fibroblasts but not myofibroblasts”. *JICE 2018:51(Suppl 1):S86*
 - 14th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2018
7. Y. Mohamied, **K.N. Tzortzis**, R.A. Chowdhury, C.D. Cantwell, A.A. Bharath, N.S. Peters. “Machine learning unravels the cellular determinants of electrogram morphology using convolutional neural networks”. *JICE 2018:51(Suppl 1):S85*
 - 14th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2018
8. **K.N. Tzortzis**. “Understanding the origins of the electrogram using machine learning”.
 - Joint Symposium of BHF Centre of Research Excellence/EPSRC Centre for Mathematics of precision Healthcare, London, UK, June 2017
9. **K.N. Tzortzis**, A.R. Simon, P. Kanagaratnam, S.J. Sherwin, C.M. Terracciano, N.A. Qureshi, F.S. Ng, C.D. Cantwell, R.A. Chowdhury, N.S. Peters. “Supervised machine learning predicts high fibrotic areas in ex vivo human left ventricular slices using additional features alongside voltage and fractionation”.
 - 5th Annual BHF 4-year PhD student conference, April 2017, London, UK
10. **K.N. Tzortzis**, Y.K. Wang, C.D. Cantwell, S.J. Sherwin, R.A. Chowdhury, N.S. Peters. “Correlations between electrogram morphology and cellular component of scar are lost when occurring in combination with functional modulations”. *JICE 2017:48(Suppl 1):S55*
 - 13th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Rome, Italy, April 2017
11. **K.N. Tzortzis**, Y.K. Wang, R.A. Chowdhury, S.J. Sherwin, C.D. Cantwell, N.S. Peters. “Identification of specific cellular abnormalities from electrogram morphology using supervised machine learning”. *JICE 2017:48(Suppl 1):S76*
 - 13th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Rome, Italy, April 2017
12. **K.N. Tzortzis**, C.D. Cantwell, S.J. Sherwin, R.A. Chowdhury, N.S. Peters. “Electrogram abnormalities can be directly correlated to specific ion channel and action potential changes”. *JICE 2016:45(Suppl 1):S293*
 - 12th Annual Scientific Congress of the European Cardiac Arrhythmia Society, Paris, France, April 2016

Table of Contents

DECLARATION OF ORIGINALITY	2
COPYRIGHT DECLARATION	3
ABSTRACT	4
ACKNOWLEDGEMENTS	6
PUBLICATIONS AND ABSTRACTS RELATED TO WORK UNDERTAKEN	8
TABLE OF CONTENTS	10
LIST OF FIGURES	14
LIST OF TABLES	19
LIST OF ABBREVIATIONS	21
CHAPTER 1: INTRODUCTION	24
1.1 OVERVIEW.....	25
1.2 ELECTRICAL RECORDINGS	26
1.2.1 Cardiac Action Potential	26
1.2.2 Unipolar electrogram	28
1.2.3 Bipolar electrogram.....	29
1.2.4 Electrocardiogram.....	29
1.3 CIRCUMSTANCES THAT CAUSE COMPLEX ELECTROGRAM MORPHOLOGY	30
1.3.1 Structural determinants influencing electrogram morphology	30
1.3.2 Functional determinants influencing electrogram morphology	32
1.4 ATRIAL ARRHYTHMIAS	35
1.4.1 Overview	35
1.4.2 Mechanisms of atrial fibrillation.....	35
1.4.3 Electrical remodelling in AF.....	36
1.4.4 Dynamics of intracellular Ca ²⁺ cycling.....	36
1.4.5 Structural remodelling in AF	37
1.5 VENTRICULAR ARRHYTHMIAS	37
1.5.1 Ventricular arrhythmia mechanisms	38
1.6 CONTEMPORARY METHODS FOR EVALUATION OF THE ATRIAL SUBSTRATE IN PERSISTENT AF	40
1.6.1 Voltage mapping	40
1.6.2 Complex Fractionated Atrial Electrograms	41
1.6.3 Dominant Frequency.....	43
1.6.4 Shannon entropy	44
1.6.5 Late-gadolinium enhanced cardiac magnetic resonance imaging.....	44
1.7 MACHINE LEARNING	44
1.7.1 Supervised Machine Learning	45
1.7.2 Unsupervised Machine Learning and Reinforcement Learning	47
1.7.3 Machine learning applications in cardiac electrophysiology.....	47
1.8 OUTLINE OF THESIS	49
CHAPTER 2: MATERIALS & METHODS	52
2.1 INTRODUCTION.....	53
2.2 MATERIALS	54
2.2.1 Cell Culture.....	54
2.2.2 Collection and preparation of tissue samples.....	55
2.2.3 Ion channel and Gap junction modulators	58
2.3 Microelectrode array recordings	58
2.4 Clinical recordings	60
2.5 OPTICAL MAPPING	62
2.5.1 Principles of optical mapping of transmembrane voltage (V _m)	62
2.5.1.1 Voltage-sensitive dyes	62
2.5.1.2 Motion artefact.....	62
2.5.1.3 Optics	63

2.5.1.4 Detectors	64
2.5.2. Optical mapping protocol.....	64
2.5.3 Optical mapping data analysis	66
2.6 IMMUNOFLUORESCENCE.....	66
2.6.1 Principles of immunofluorescence.....	66
2.6.2 Immunofluorescence protocol.....	67
2.6.3 Cell Tracking System.....	68
2.6.4 Analysis of target-protein distribution	68
2.7 SECOND-HARMONIC GENERATION MICROSCOPY	69
2.7.1 Principles of second-harmonic generation microscopy	69
2.7.2 Second-harmonic generation microscopy protocol.....	69
2.8 STATISTICAL ANALYSIS	70
2.9 MACHINE LEARNING TECHNIQUES.....	70
2.9.1 Supervised machine learning	70
2.9.2 Feature selection	77
2.9.3 Assessment of predictions by cross-validation	78
2.9.4 Measurement of classification performance	79
CHAPTER 3: DEVELOPMENT OF AUTOMATED ELECTROGRAM MORPHOLOGY ANALYSIS 81	
3.1 INTRODUCTION.....	82
3.1.1 Digital signal processing in cardiac electrophysiology.....	82
3.1.2 Electrogram morphology analysis	83
3.2 METHODS	85
3.2.1 Electrogram features: An overview	85
3.2.2 Stimulus artefact removal	85
3.2.3 Time-domain feature extraction.....	86
3.2.4 Frequency-domain EGM features	90
3.2.5 Time-frequency EGM features	94
3.2.6 Features based on information theory	96
3.2.7 Computational time for morphology analysis.....	98
3.3 RESULTS.....	99
3.3.1 Quantification of electrogram features.....	99
3.3.2 Verification of automated feature analysis	101
3.3.3 The T-wave detection.....	102
3.4 DISCUSSION.....	104
3.4.1 Limitations	106
3.5 CONCLUSIONS	106
CHAPTER 4: CHARACTERISATION OF EGM MORPHOLOGY MODIFICATIONS DUE TO ION CURRENT MODULATIONS 107	
4.1 INTRODUCTION.....	108
4.1.1 Simultaneous electrogram and action potential recordings	108
4.1.2 AP and EGM morphology affected during electrical remodelling	109
4.2 METHODS	110
4.2.1 In vitro model of cellular component of scar	110
4.2.2 Ion channel blockade	111
4.2.3 Measurement of conduction velocity across each MEA.....	111
4.2.4 Synchronisation of dual modality concurrent electrogram and action potential recordings.....	112
4.2.5 Statistical analysis	113
4.3 RESULTS.....	114
4.3.1 The effects of Na ⁺ channel blockade on EGM and action potential morphology	114
4.3.2 EGM and action potential morphology manifestations of transient outward K ⁺ current blockade ...	118
4.3.4 I _{Kr} -related EGM and action potential morphology modifications	124
4.3.5 The effects of I _{Ks} blockade on EGM and action potential morphology	126
4.3.7 EGM and action potential morphology manifestations of double ion channel blockade	131
4.3.8 Comparative analysis of EGM features across different in vitro models	134
4.3.9 EGM morphology across different types of K ⁺ current modulation	147
4.4 DISCUSSION.....	149
4.4.1 Na ⁺ channel blockade effects on EGM morphology	150
4.4.2 Transient outward K ⁺ current effects	151
4.4.3 L-type Ca ²⁺ channel blockade effects	152

4.4.4 The influence of rapidly activating delayed rectifier K ⁺ current on EGM morphology	152
4.4.5 Slowly activating delayed rectifier K ⁺ current related changes	153
4.4.6 K ⁺ current enhancement alters EGM morphology	154
4.4.7 Investigating the effects of double current blockade	155
4.5 LIMITATIONS	156
4.6 CONCLUSIONS	157
CHAPTER 5: EGM MORPHOLOGY EFFECTS OF GAP JUNCTIONAL BLOCK IN MULTICELLULAR PREPARATIONS	158
5.1 INTRODUCTION	159
5.2 RESULTS	161
5.2.1 The effects of gap junction uncoupling on cellular EGM morphology	161
5.2.2 Effects on upstroke heterogeneity and fractionation at the single cell level.....	168
5.2.3 The effects of gap junction uncoupling on human ventricular slice-derived EGM morphology	169
5.3 DISCUSSION	174
5.3.1 In vitro gap junction uncoupling influence on EGM morphology.....	174
5.3.2 Ex vivo gap junctional block effects.....	175
5.4 CONCLUSIONS	177
CHAPTER 6: EVALUATION OF EGM MORPHOLOGY MODIFICATIONS DUE TO STRUCTURAL REMODELLING.....	178
6.1 INTRODUCTION	179
6.1.1 Structural remodelling affects cardiac electrophysiology.....	179
6.1.2 The role of myofibroblasts.....	181
6.2 RESULTS	182
6.2.1 In vitro effects of cellular scar on EGM morphology	182
6.2.2 EGM morphology changes compared across modulations.....	185
6.2.3 Investigating EGM morphology discrepancies.....	187
6.2.4 Myofibroblasts influence EGM morphology	188
6.2.5 Ex vivo effects of fibrosis on AP and EGM morphology.....	190
6.2.6 In vivo effects of scarred areas on EGM morphology	195
6.3 DISCUSSION	197
6.3.1 Fibroblast effects of EGM morphology	197
6.3.2 The role of myofibroblasts.....	199
6.3.3 Characterisation of EGM feature modifications in adult human ventricular slices	200
6.3.4 The in vivo relationship between EGM morphology and atrial fibrosis.....	201
6.4 LIMITATIONS	203
6.5 CONCLUSION	203
CHAPTER 7: PREDICTING FUNCTIONAL AND STRUCTURAL ABNORMALITIES FROM CELLULAR EGM MORPHOLOGY	205
7.1 INTRODUCTION	206
7.2 METHODS	207
7.3 RESULTS	208
7.3.1 Data used for supervised machine learning	208
7.3.2 Feature selection for predicting functional and structural modulations.....	211
7.3.3 Identification of the optimal classifier	211
7.3.4 Predicting functional and structural modulations using cellular EGMs	214
7.3.5 Validation of prediction model	221
7.3.6 Predictive modelling differences between functional and structural abnormalities	223
7.4 DISCUSSION	227
7.4.1 Features describing the datasets and data reduction	227
7.4.2 Classification training for in vitro cellular level abnormalities	229
7.5 LIMITATIONS	231
7.6 CONCLUSIONS	232
CHAPTER 8: PREDICTING FIBROSIS FROM EX VIVO AND IN VIVO OBTAINED EGM MORPHOLOGY	233
8.1 INTRODUCTION	234
8.1.1 Applications of supervised machine learning in cardiac electrophysiology.....	234
8.1.2 Clinical relevance of classifying cardiac arrhythmias	234

8.1.3	Experimental models for the development of prediction models	235
8.2	RESULTS	237
8.2.1	Collagen coverage: the fibrosis measure on tissue slices	237
8.2.2	Feature selection using ex vivo EGMs	238
8.2.3	Optimising the classification training process	239
8.2.4	Predicting fibrosis levels using ex vivo data	241
8.2.5	Including conventional scar markers at prediction model	243
8.2.6	Predicting different degrees of collagen coverage	245
8.2.7	Feature selection using in vivo EGMs	249
8.2.8	Classification training using in vivo EGMs	249
8.2.9	Validation of predictions based on in vivo EGM data	252
8.3	DISCUSSION	253
8.3.1	Ex vivo structural determinants as a limitation to EGM signal recording	253
8.3.2	Are voltage and fractionation useful predictors of the structural fibrotic substrate?	254
8.3.3	Important features for identifying fibrotic tissue	255
8.4	LIMITATIONS	258
8.5	CONCLUSIONS	258
CHAPTER 9: PREDICTING HEART FAILURE PHENOTYPES FROM EX VIVO ELECTROGRAMS USING SUPERVISED MACHINE LEARNING		259
9.1	INTRODUCTION	260
9.1.1	Heart failure	260
9.1.2	Ischaemic Heart Disease	261
9.1.3	Dilated cardiomyopathy	262
9.1.4	Hypertrophic cardiomyopathy	262
9.1.5	Lymphocytic myocarditis	263
9.1.6	Electrophysiological changes in heart failure	263
9.2	RESULTS	265
9.2.1	Quantification of collagen coverage	265
9.2.2	Measurement of conduction velocity	266
9.2.3	Heart failure phenotypes and EGM morphology	267
9.2.3	Feature selection	272
9.2.4	Training an accurate heart failure phenotype classification algorithm	273
9.2.5	Trained algorithm reliably classifies EGMs according to heart failure phenotype	275
9.3	DISCUSSION	275
9.3.1	The effects of heart failure in EGM morphology	276
9.3.2	Classification of heart failure	277
9.4	LIMITATIONS	278
9.5	CONCLUSION	278
CHAPTER 10: CONCLUSIONS		279
10.1	SUMMARY OF KEY FINDINGS	280
10.1.1	Electrogram morphology effects due to functional and structural changes	282
10.1.2	Supervised machine learning for predicting functional and structural modifications	285
10.2	CLINICAL RELEVANCE OF FINDINGS	286
10.3	FUTURE DIRECTIONS	287
10.4	CONCLUDING REMARKS	289
REFERENCES		290
APPENDIX A: ANALYSIS OF ELECTROGRAM MORPHOLOGY MODIFIED BY INDIVIDUAL FUNCTIONAL AND STRUCTURAL MODULATIONS		312
I.	The effects of Na ⁺ channel blockade on EGM morphology	313
II.	EGM morphology manifestations of transient outward K ⁺ current blockade	318
III.	The effects of L-type Ca ²⁺ current blockade on EGM morphology	322
IV.	I _{Kr} -related EGM morphology modifications	327
V.	The effects of I _{Ks} blockade on EGM morphology	332
VI.	EGM morphology modifications following K _{ATP} channel opening	335
VII.	EGM morphology manifestations of double ion channel blockade	336
APPENDIX B: FEATURE SUBSETS SELECTED FOR CLASSIFICATION TRAINING		339
APPENDIX C: PERMISSIONS		343

List of Figures

Figure 1.1 Action potential waveforms in adult human ventricular and atrial myocytes	28
Figure 1.2 Types of human cardiac fibrosis in explanted hearts, as visualised with light microscopy	32
Figure 1.3 Determinants of EGM morphology	32
Figure 1.4 Summary of typical connexin expression patterns in mammalian heart	33
Figure 1.5 Example of unipolar electrogram recording collected from mouse Langendorff-perfused hearts with VT and VF	38
Figure 1.6 Voltage gradient maps in sinus rhythm and atrial fibrillation	41
Figure 1.7 Example of CFAEs with continuous prolonged activation complex over the posterior septal areas	42
Figure 1.8 Bipolar recording of CFAEs and the corresponding power spectra following Fast Fourier Transformation from an LA site	43
Figure 2.1 Overview of the methodology process followed in this thesis	53
Figure 2.2 Tissue slice preparation	56
Figure 2.3 MEA dish and panel electrogram recordings	59
Figure 2.4 The 7F Inquiry AFocusII catheter with 4F double loop used for acquiring clinical data ...	61
Figure 2.5 Optical mapping system	61
Figure 2.6 Simultaneous optical and EGM data	65
Figure 2.7 k-Nearest Neighbour illustration	72
Figure 2.8 An illustration of Support Vector Machine (SVM) for binary classification.....	73
Figure 2.9 Illustration of linear and quadratic discriminant analysis.....	74
Figure 2.10 Example of a decision tree illustration with its root at the top.....	75
Figure 2.11 Schematic representation of the Random Forests algorithm based on the Bagging method	76
Figure 3.1 Representative stimulus artefacts in EGM recordings obtained from NRVM monolayers	87
Figure 3.2 Representative resulting signals and corresponding power spectra after each step of analysis during the filtering protocol, in order to compute the dominant frequency.....	92
Figure 3.3 Example of an EGM obtained from a NRVM monolayer (top) with the subsequent decomposition of the CWT time series.....	99
Figure 3.4 Examples of unipolar EGM recordings obtained from different settings	100
Figure 3.5 Time-domain features, including EGM amplitude, EGM duration, fractionation index, R-/S-peak width, were measured both manually and using the automated feature extraction method ..	101
Figure 3.6 The process to identify a T-wave in EGM signal obtained from NRVM monolayers.....	103
Figure 4.1 Conduction velocity measurement across MEA	112
Figure 4.2 Effects of Na ⁺ channel blockade using lidocaine on action potential upstroke, EGM duration and conduction velocity in NRVM monolayers.	115
Figure 4.3 Dose response curves generated after treating NC cultures and co-cultures of NRVMs with low or high amounts of fibroblasts with a range of lidocaine concentrations	117
Figure 4.4 Data from concurrent optical mapping and EGM recordings following the administration of 4-AP in myocyte only NRVM cultures	119
Figure 4.5 Data from concurrent optical mapping and EGM recordings following the administration of 4-AP in NC NRVM cultures.....	120
Figure 4.6 Data from concurrent optical mapping and EGM recordings following the administration of nifedipine in myocyte only NRVM cultures	122

Figure 4.7 Data from concurrent optical mapping and EGM recordings following the administration of nifedipine in NC NRVM cultures.....	123
Figure 4.8 Effects of I_{K_r} current blockade using 303.3nM E-4031 on action potential duration and FPD.	125
Figure 4.9 Data from concurrent optical mapping and EGM recordings following the administration of E-4031 in NC NRVM cultures	126
Figure 4.10 Correlation between APD_{90} and FPD in NRVM cultures after I_{K_s} blockade using HMR-1556.....	128
Figure 4.11 Data from concurrent optical mapping and EGM recordings following the administration of HMR-1556 in NC NRVM cultures.....	129
Figure 4.12 Effect of K^+ current enhancement using pinacidil on AP duration and FPD.	130
Figure 4.13 Data from concurrent optical mapping and EGM recordings following the simultaneous blockade of Na^+ and L-type Ca^{2+} channels in myocyte only NRVM cultures	132
Figure 4.14 Data from concurrent optical mapping and EGM recordings following the simultaneous blockade of Na^+ and L-type Ca^{2+} channels in NC NRVM cultures.....	133
Figure 4.15 Bar charts presenting the modification of the EGM amplitude, the average EGM amplitude and the S.D. EGM amplitude due to ion channel blockade in different types of NRVM cell culture.....	136
Figure 4.16 Bar charts presenting changes in RS interval, QS interval and the RS interval gradient due to ion channel blockade in different types of NRVM cell culture	137
Figure 4.17 Modification of the EGM onset to R-peak gradient, S-peak to EGM offset gradient and R-peak width, due to ion channel blockade in different types of NRVM cell culture.....	138
Figure 4.18 Bar charts presenting the modification of the S-peak width, R-/S-width ratio and EGM duration due to ion channel blockade in different types of NRVM cell culture.....	139
Figure 4.19 Bar charts presenting the R-width:EGM duration ratio, R-/S-peak ratio and the fractionation index due to ion channel blockade in different types of NRVM cell culture	140
Figure 4.20 Bar charts presenting the modification of the dominant frequency, Shannon entropy and the logarithmic energy entropy due to ion channel blockade in different types of NRVM cell culture	141
Figure 4.21 Bar charts presenting the modification of the percent of energy, maximum modulus and variance of energy due to ion channel blockade in different types of NRVM cell culture.....	142
Figure 4.22 Bar charts presenting the modification of the frequency of maximum energy, frequency of minimum energy and the scale of maximum energy due to ion channel blockade in different types of NRVM cell culture.....	143
Figure 4.23 Bar charts presenting the modification of the scale of minimum energy, mean frequency of PSD estimates and the median frequency of PSD estimates due to ion channel blockade in different types of NRVM cell culture	144
Figure 4.24 Bar charts presenting the modification of the maximum PSD estimates, the skewness of PSD estimates and the S.D. PSD estimates due to ion channel blockade in different types of NRVM cell culture.....	145
Figure 4.25 Bar charts presenting the modification of the kurtosis of PSD estimates and the S.D. autocorrelation due to ion channel blockade in different types of NRVM cell culture	146
Figure 4.26 Features changed by K^+ current enhancement with pinacidil, but not by K^+ current blockade by 4-AP, E-4031 or HMR-1556	147
Figure 4.27 Features mainly changed in an indiscriminate manner following drug exposure	148
Figure 4.28 Features changed in an indiscriminate manner following drug exposure	149
Figure 5.1 Data from concurrent optical mapping and EGM recordings following the gap junction uncoupling using CBX in myocyte only NRVM cultures	162

Figure 5.2 EGM morphology modifications due to gap junction uncoupling when 40µM CBX were administered on myocyte only cultures	164
Figure 5.3 Data from concurrent optical mapping and EGM recordings following the gap junction uncoupling using CBX in NC and +20%FB NRVM cultures	165
Figure 5.4 Modifications on EGM morphology due to gap junction uncoupling when 40µM CBX were administered on NC cultures.....	166
Figure 5.5 EGM morphology modifications due to gap junction uncoupling when 3µM CBX were administered on NRVM/+20% fibroblast co-cultures	167
Figure 5.6 Single cell level correlation of action potential dispersion and EGM fractionation.....	169
Figure 5.7 Quantification of collagen coverage on slices derived from human left ventricle.....	170
Figure 5.8 Data from concurrent optical mapping and EGM recordings following the gap junction uncoupling using CBX in human ventricular tissue slices	172
Figure 5.9 EGM morphology modifications due to gap junction uncoupling when human left ventricle tissue slices were perfused with 15.8µM CBX	173
Figure 6.1 Immunostaining for fibroblast quantification.....	183
Figure 6.2 Effects of fibroblast burden on APD ₉₀ and FPD.	184
Figure 6.3 EGM properties showing significant differences among cell groups with various amounts of fibroblasts	185
Figure 6.4 Heatmap presenting EGM feature changes per modulation.....	187
Figure 6.5 Evaluation of the EGM morphology effects with the absence of normal gap junctions... ..	188
Figure 6.6 EGM morphology modifications showing a significant difference in the myocyte only group compared to either the fibroblast or the myofibroblast group or both.....	189
Figure 6.7 Human left ventricle tissue slices visualised in three different ways.	191
Figure 6.8 Representative examples of EGMs obtained from tissue slices with different collagen coverage	192
Figure 6.9 Data from concurrent optical mapping and EGM recordings in human ventricular tissue slices.....	193
Figure 6.10 EGM properties modified due to differences in the size of collagen coverage, and subsequently the extent of fibrosis, in human left ventricle tissue slices	194
Figure 6.11 Representative examples of 2.5msec clinical recordings obtained from the non-scar and scar regions of the LA posterior wall.....	195
Figure 6.12 EGM amplitude was no different between control and scarred regions	196
Figure 6.13 EGM morphology modifications due to the presence of scar, as measured on 3Hz data collected from control and scarred regions of the LA endocardium.....	196
Figure 7.1 Classification performance achieved using in vitro data.....	213
Figure 7.2 Optimisation of the Ensemble Bagging method.....	213
Figure 7.3 Classification performance achieved on the EGM signal dataset by Bagging ensemble classification.....	215
Figure 7.4 ROC curves obtained after classification training using the Ensemble Bagging method.	216
Figure 7.5 Bar graphs presenting how the SFS features used during classification training changed across a range of classes which resulted to high false negative rates, and consequently low sensitivity, in relation to the control (MO baseline) class	218
Figure 7.6 Bar graph analysis presenting no significant differences between the baseline state of NC cell cultures and the rest of classes with which a high false negative rate was found comparing the SFS feature changes.	219
Figure 7.7 High false negative rates for the baseline class of High % FB cell cultures can be explained by the absence of significantly different SFS features between that class and a group of others.....	220
Figure 7.8 Classification performance following training for the functional modifications.	221

Figure 7.9 Classification performance following training and using an EGM dataset consisting of data obtained from the four different types of cell monolayer used in this thesis at the baseline state.....	224
Figure 7.10 Classification performance following training and using EGM dataset consisting of data obtained from the four different types of cell monolayer used in this thesis at the baseline state.....	226
Figure 8.1 Frequency distribution graph of tissue slices used in this thesis based on the percentage of the area of each slice covered by collagen.....	238
Figure 8.2 EGM properties modified due to differences in the size of collagen coverage, as measured on data obtained from human left ventricle slices	240
Figure 8.3 Optimisation of the Ensemble Bagging method.....	240
Figure 8.4 Classification training performance using instances classified as obtained from tissue slices with collagen coverage below or above 15% of the total slice area	241
Figure 8.5 Classification performance achieved using <i>ex vivo</i> data classified as obtained from tissue slices with <15% or >15% collagen coverage	242
Figure 8.6 Confusion matrices extracted from classification training with the same training dataset but using different features each time	244
Figure 8.7 Panel of confusion matrices presenting the classification training performance for each dataset. EGMs were classified based on a certain percentage of the slice area covered by collagen and that ranged between 15-23%.....	247
Figure 8.8 Assessment of classification training using the Ensemble Bagging method by presenting the results on a confusion matrix	251
Figure 8.9 Confusion matrices presenting the classification performance when the predictors were the SFS features in addition to voltage	252
Figure 9.1 Collagen coverage of tissue slices derived from different heart failure phenotypes.....	266
Figure 9.2 Average conduction velocity and conduction velocity anisotropy across tissue slices.....	267
Figure 9.3 Time-domain features analysed using <i>ex vivo</i> data, which were affected by HF phenotypes	270
Figure 9.4 Morphological descriptors modified due to HF phenotypes	271
Figure 9.5 Linear regression analysis for the relationship between the conduction velocity, measured as the average conduction velocity from two perpendicular sides of a tissue slice, and EGM duration	272
Figure 9.6 Classification training results. (A) Relative classification errors as a function of number of grown decision trees.....	274
Figure 10.1 Schematic representation of a Langendorff perfusion system.....	288
Figure A.1 EGM morphology modifications due to Na ⁺ channel blockade when 50µM lidocaine were administered on NRVM cultures.	314
Figure A.2 EGM morphology modifications in NC cell cultures, due to the Na ⁺ channel blockade using 1µM lidocaine.....	315
Figure A.3 The effects of Na ⁺ channel blockade using 1µM lidocaine on EGM morphology recorded in co-cultures of NRVM and 20% cultured fibroblasts.	316
Figure A.4 EGM morphology modifications due to Na ⁺ channel blockade when 10µM lidocaine were administered on co-cultures of NRVM with 40% cultured fibroblasts	317
Figure A.5 EGM morphology modifications due to the I _{to} blockade when 750µM 4-AP were administered on NRVM cultures	319
Figure A.6 EGM morphology modifications due to the I _{to} blockade when 750µM 4-AP were administered on NC cell monolayers.	320
Figure A.7 EGM morphology modifications due to the I _{to} blockade when 300µM 4-AP were administered on co-cultures of NRVM and 20% cultured fibroblasts.....	321
Figure A.8 EGM morphology modifications due to I _{CaL} blockade on MO cell monolayers	323

Figure A.9 EGM morphology modifications due to the I_{CaL} blockade on NC monolayers.....	324
Figure A.10 EGM morphology modifications due to I_{CaL} blockade on co-cultures of NRVM and 20% cultured fibroblasts.....	325
Figure A.11 Modifications on EGM morphology due to the I_{CaL} blockade on co-cultures of NRVM and 40% cultured fibroblasts.....	326
Figure A.12 EGM morphology changes due to the I_{Kr} blockade on NRVM monolayers.	328
Figure A.13 EGM morphology modifications due to the I_{Kr} blockade on NC cell monolayers.....	329
Figure A.14 EGM morphology modifications due to the I_{Kr} blockade on NRVM/20% fibroblast co-cultures	339
Figure A.15 Modifications on EGM morphology due to the I_{Kr} blockade on co-cultures of NRVM and 40% cultured fibroblasts	331
Figure A.16 EGM morphology modifications due to the I_{Ks} blockade on MO monolayers	332
Figure A.17 Modifications on EGM morphology due to the I_{Ks} blockade on NC cultures	333
Figure A.18 EGM morphology modifications due to the I_{Ks} blockade on NRVM/20% fibroblast co-cultures	334
Figure A.19 EGM morphology modifications due to the KATP channel opening on MO cultures ..	335
Figure A.20 Modifications on EGM morphology due to the double ion channel blockade on MO monolayers	337
Figure A.21 Modifications on EGM morphology due to the double ion channel blockade on NC cell cultures	338

List of Tables

Table 2.1 Heart failure patient characteristics	56
Table 2.2 Characteristics of the donor of rejected heart transplant.....	56
Table 2.3 Equations for assessing the multi-class classification of a dataset into l classes	77
Table 3.1 List of time-domain features measured in this thesis.....	86
Table 3.2 List of frequency-domain features measured by using the algorithms presented in section 3.2.4.....	91
Table 3.3 List of time-frequency features computed in this thesis by using the algorithms presented in section 3.2.5.	93
Table 3.4 List of features which are based on information theory and non-linear dynamics and which are computed using the algorithms presented in section 3.2.6.....	95
Table 3.5 Manual and automated measurement of representative time-domain features.....	99
Table 4.1 Combination of lidocaine/nifedipine concentrations used for double ion channel blockade treatment of myocyte only NRVM cultures.....	125
Table 4.2 Combination of lidocaine/nifedipine concentrations used for double ion channel blockade treatment of NC NRVM cultures	127
Table 7.1 Total number of electrograms collected under a variety of experimental conditions from cell cultures, which were analysed and further used for the development of training and test datasets....	202
Table 7.2 Number of EGMs collected per experimental condition used for classification training and evaluation of predictive models	203
Table 7.3 Classification performance for each of the classification methods considered	205
Table 7.4 Feature subset extracted by forward SFS using the Ensemble Bagging method during feature selection process listed in order of significance from the most to the least important	205
Table 7.5 Assessment of classification training using training dataset.....	209
Table 7.6 Evaluation performance of prediction model using a test dataset	214
Table 7.7 The 10-fold cross-validation classification training performance was assessed.....	217
Table 7.8 Evaluation performance of the model developed for predicting functional abnormalities using a test dataset with 9 classes	217
Table 7.9 Assessment of the classification training performance using EGMs obtained from cell cultures with different amounts of cellular scar	218
Table 7.10 Evaluation of the prediction model based on the different types of cell culture used in this thesis.....	219
Table 8.1 Classification performance presenting the accuracy after applying 20 classification training methods which are available on MATLAB	231
Table 8.2 Feature subset extracted by forward SFS using the Ensemble Bagging method during feature selection process listed in order of significance	233
Table 8.3 Assessment of the 13-fold cross-validation classification training performance	235
Table 8.4 Evaluation performance of the model developed for predicting the collagen coverage on tissue slices.....	235

Table 8.5 Assessment of classification training and evaluation of subsequent prediction models using the training dataset and different feature subsets on each case.....	236
Table 8.6 Table presenting the number of tissue slices used per class (above or below each % collagen coverage), the number of EGMs which were recorded and analysed in each case and the size of training and test datasets.	237
Table 8.7 Assessment of classification training using datasets based on a range of collagen coverage percentages, in order to investigate the impact of different classification systems on the same dataset	240
Table 8.8 Classification performance presenting the accuracy after applying 20 classification training methods which are available on MATLAB	242
Table 8.9 Assessment of the 10-fold cross-validation classification training. Calculation of indices for the overall and per class performance	243
Table 8.10 Assessment of the classification training using different EGM features as predictors. Classification indices computed for the overall and per class performance	244
Table 8.11 Evaluation performance of the model developed for predicting the presence of scar on the substrate where the EGMs were collected from	245
Table 9.1 Number of tissue slices used for EGM recordings and the total number of EGMs on which feature extraction was carried out	260
Table 9.2 Classification accuracy of prediction methods using the same training dataset. The classification training was carried out with 10-fold cross-validation	265
Table 9.3 Systematic analysis of classification performance with Bagging Decision Trees.....	266
Table 9.4 Evaluation performance of prediction model using a test dataset	267

List of Abbreviations

2D	Two-dimensional
3D	Three-dimensional
4-AP	4-aminopyridine
ACF	Autocorrelation function
AF	Atrial Fibrillation
AI	Artificial Intelligence
ANOVA	Analysis of variance
AP	Action potential
APD	Action potential duration
ARI	Activation recovery interval
ATPase	Adenosine triphosphatase
AUC	Area under curve
BDM	2,3-butanedione monoxime
bpm	Beats per minute
BrdU	5'-bromo-2'deoxyuridine
BSA	Bovine Serum Albumin
CBX	Carbenoxolone
CCD	Charge-coupled device
CFAE	Complex Fractionated Atrial Electrograms
CMOS	Complementary metal-oxide semiconductor
CO ₂	Carbon dioxide
CONFIRM	Conventional Ablation with or without FIRM
CWT	Continuous wavelet transform
Cx	connexin
DAD	Delayed afterdepolarisation
DCM	Dilated cardiomyopathy
DF	Dominant frequency
DMSO	Dimethyl sulfoxide
DSP	Digital signal processing
EAD	Early afterdepolarisation
ECG	electrocardiogram
ECM	Extracellular matrix

EDTA	Ethylenediaminetetraacetic acid
EGM	electrogram
EP	electrophysiology
FB	Fibroblasts
FBS	Foetal Bovine Serum
FFT	Fast Fourier Transform
FPD	Field potential duration
HBSS	Hank's buffered saline solution
Hz	Hertz
ICL	Interval confidence level
IHD	Ischaemic heart disease
IPs	Inflection points
kDa	Kilodalton
kNN	k-Nearest Neighbours
LA	Left atrium
LDA	Linear discriminant analysis
LED	Light-emitting diode
LGE-CMRI	Late-gadolinium enhancement cardiac magnetic resonance imaging
LogEn	Logarithmic energy entropy
LQTS	Long QT Syndrome
LVAD	Left ventricular assist device
MEA	Micro-electrode array
Mg	Milligram
ml	Millilitre
mM	Millimolar
MO	Myocyte only cell culture
ms	Millisecond
n	Sample size
NC	Natural composition
NCX	Na ⁺ -Ca ²⁺ exchanger
NPV	Negative predictive value
NRVM	Neonatal rat ventricular myocytes
ns	Not significant

OCS	Organ care system
OOB	Out-of-bag
p	Statistical p value
PBS	Phosphate buffered saline
PDA	Photodiode array
PDGF	Platelet-derived growth factor
PPV	Positive predictive value
PSD	Power spectral density
PVI	Pulmonary vein isolation
QDA	Quadratic discriminant analysis
RI	Regularity index
ROC	Receiver operating characteristic
RUSBoost	Random Under-Sampling Boosting
s	Second
SD	Standard deviations
Se	Sensitivity
SFS	Sequential Forward Selection
ShanEn	Shannon entropy
SHG	Second Harmonic Generation
Sp	Specificity
SVM	Support Vector Machines
TGF- β 1	Transforming growth factor-beta 1
UV	ultraviolet
VF	Ventricular Fibrillation
VT	Ventricular Tachycardia
α SMA	α -smooth muscle actin
Δ T	Time difference
μ M	Micromolar

Chapter 1

Introduction

1.1 Overview

The extracellular contact electrogram (EGM) is the signature of the interaction of electrical activation and architecture of the local myocardium and is the primary diagnostic tool in the cardiac catheter laboratory. An EGM is the summation of the electric fields of ions in the vicinity of the electrode. Both the ion flux across cell membranes, which produces the action potential, and the propagation of the action potential lead to variations in electric field strength at the electrode and the subsequent EGM morphology (1). Healthy myocardium gives rise to a simple EGM, but pathological conditions can lead to changes in EGM characteristics and as a result the morphology becomes more complex (2). EGMs are currently categorised clinically by largely binary descriptors, such as simple or complex, early or late, which do not provide significant information about the content of the signal (2,3). In order to make more effective use of EGMs for arrhythmia diagnosis, it is therefore necessary to systematically characterise cellular electrophysiology and the factors that mediate action potential propagation and understand their impact on the resulting EGM morphology. The EGM could then be used inversely for the identification of specific cellular and tissue characteristics and potentially provide a novel clinical diagnostic tool.

There are a number of studies which have attempted to correlate cell structure and tissue architecture with function and have identified specific factors that can be responsible for the production of abnormal EGMs (4). EGM morphology can be modified due to conduction slowing, gap junction uncoupling, presence of fibrotic areas or electrical remodelling (4,5). These factors have been studied individually for their role in cardiac electrophysiology, but it is unclear how they are directly related to specific changes in EGM morphology and how different combinations of these structural and functional factors are implicated on arrhythmogenesis. For example, it has been suggested that EGM fractionation observed in atrial fibrillation (AF) is correlated to sites responsible for clinical AF maintenance (6). It is assumed that fractionation relates to an increase of fibrotic regions and thus, translates to reduced EGM voltage. However, it has been consistently reported (7) that there may be an inverse and counter-intuitive relationship between voltage and fractionation in AF patients. In addition, emerging clinical evidence from persistent AF treatment by focal ablation has not shown any specific electro-architectural features to indicate these sites.

The information contained in the electrical signal could be used for the improvement of diagnosis and treatment of cardiac arrhythmias. A wide variety of algorithms have been developed based on morphological, spectral or complexity parameters extracted from electrophysiology recordings (8). The combination of signal characteristics with machine learning techniques has been suggested to be a

useful approach for more efficient diagnosis. Machine learning, the scientific discipline that focuses on training computer models from data, uses statistical algorithms in order to group or find patterns in a dataset (9). Supervised machine learning is one of the main fields of machine learning and its goal is training of algorithms using predictors, in order to make predictions afterwards from new data (9,10). A number of efforts have been made over the past two decades, in order to classify electrocardiogram (ECG) recordings based on tissue architecture or pathological characteristics. There has been recent progress, in terms of algorithm development mainly focusing on ECG morphology analysis (11,12). Signal analysis techniques can be used to generate the features which are necessary during classification training in supervised machine learning.

The main clinical EGM recording system in the clinic is the cardiac catheter. This is also used to deliver catheter ablation targeted in the area of atrial pulmonary veins for the treatment of abnormal action potential propagation. However, the success rate of catheter ablation is lower than expected (71). This is a motivation for better understanding in more detail the factors underlying abnormal EGMs. In addition, our understanding of the changes that occur during functional and structural remodelling in cardiac diseases, and the subsequent alterations of excitation patterns, are essential for improved diagnosis, prevention and treatment of arrhythmic and heart failure conditions. Hence an objective of this thesis is investigating how specific EGM morphology features, recorded from different experimental models, are modified by functional and structural determinants of cellular electrophysiology. The primary aim though is developing the capability to get more information out of the EGM than a simple binarisation. This knowledge is aimed to be used subsequently for predicting the local electro-architecture from a decomposition of the EGM morphology, which is also a prominent focus of the work presented in this thesis.

1.2 Electrical recordings

Electrical activation can be measured at different scales ranging from the single cell to the body surface. This section presents the most common modalities for measuring electrophysiological activity.

1.2.1 Cardiac Action Potential

The *action potential* (AP) of a cardiac cell describes the changes in potential difference between the intracellular and extracellular spaces (transmembrane potential) and arises from the movement of ions across the cell membrane through ion-specific channels (Figure 1.1). The AP morphology varies not

only between species, but also between the different cell types in the myocardium. In general, the AP consists of five phases (13):

- a. **Phase 0 (rapid depolarisation)**: the membrane potential increases to a threshold value ($\sim -60\text{mV}$), due to an inward current transmitted either from an applied stimulus or an adjacent cell. This results in the opening of Na^+ channels leading to a sudden influx of Na^+ ions and the rapid depolarisation of the membrane until the membrane potential reaches a positive overshoot of $20\text{-}30\text{mV}$ at which the Na^+ current, I_{Na} , is quickly inactivated.
- b. **Phase 1 (rapid repolarisation)**: Na^+ channels close and the K^+ transient outward current (I_{to}) is activated, as K^+ ions leave the cell. The membrane partially repolarises to $-20 - 0\text{mV}$. This part of the action potential morphology, called the notch, is only present in cells that express the transient outward current.
- c. **Phase 2 (plateau)**: L-type Ca^{2+} channels (I_{CaL}) open to provide depolarising current against the repolarising K^+ currents (I_{Kr} and I_{Ks}) resulting to a plateau formation in the potential.
- d. **Phase 3 (repolarisation)**: L-type Ca^{2+} channels close and the membrane repolarises mainly due to the inward rectifier K^+ current (I_{K1}). The rate of repolarisation is slower than the rate of depolarisation during phase 0.
- e. **Phase 4 (resting phase)**: the resting membrane potential is $\sim -80\text{mV}$, as a result of the high intracellular K^+ concentration compared to the extracellular K^+ concentration.

The time period following depolarisation during which the cell cannot generate a full AP is called the *refractory period*. This consists of an initial time period when the cell is completely unexcitable (*absolute refractory period*) and a small time period during repolarisation when a sufficient stimulus may cause a partial AP (*relative refractory period*). The time duration of the AP, measured at a particular level of repolarisation, e.g. 90% of repolarisation, is the *action potential duration* (APD). The time interval between action potentials is called *diastolic interval* (13).

Optical mapping is the most prominent technique for capturing the local activation and repolarisation in multicellular preparations using fluorescent-voltage dyes. However, a significant drawback of this method is the difficulty on directly applying it on patients. Even though optical imaging would probably be the most informative technique for diagnosing abnormalities, technological advances have been made up to the point where optical mapping can be applied on intact large animal hearts in a Langendorff setup (471,472), but also in human hearts (14). Thus, it is necessary to use alternative methods for the clinical recording of the electrophysiological activity directly from cardiac arrhythmia patients.

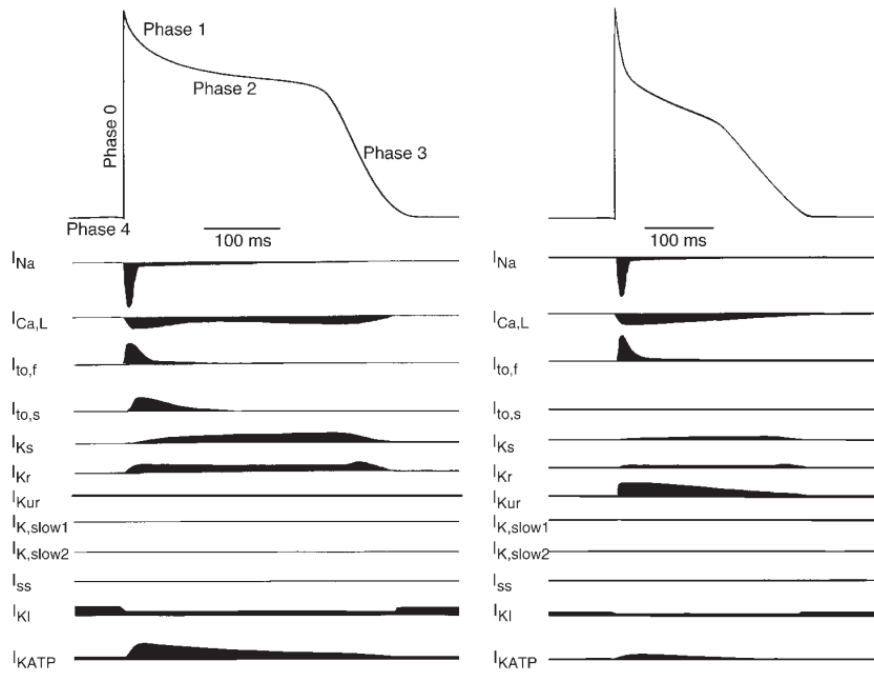


Figure 1.1: Action potential waveforms in adult human ventricular (left) and atrial (right) myocytes. The active ion currents at each phase of the action potential are also presented. Reproduced from Nerbonne and Kass (13) with permission of the rights holder, The American Physiological Society.

1.2.2 Unipolar electrogram

Unipolar and bipolar EGMs represent the summation of electrical activity from all cells within the field of view of the intracardiac electrode catheter. Unipolar EGMs represent the difference between a recording electrode placed on the tissue surface and a reference (indifferent) electrode placed a large distance away from the area of interest. The sensing electrode is connected to the positive input of a differential amplifier and the indifferent electrode is connected to the negative input, in order to measure the signal (15). While bipolar EGMs are mainly used to identify the local activation time, the unipolar signal is used to define the activation time more accurately (16).

The relationship between the slope of the AP, the slope of the extracellular EGM and the time of peak Na^+ current has been investigated in a uniform cable model and it was found that the maximum positive slope of the AP, the maximum negative slope of the unipolar EGM and the peak Na^+ current coincide in time (17). Since the peak Na^+ current corresponds to the maximum rate of increase of the depolarising current, this can be defined to be the activation time. This leads to the standard definition of the AP activation time which is the time point of maximum positive slope. The extracellular unipolar EGM equivalent of the activation time is the point of maximum negative slope.

The main advantage of using unipolar EGMs is that they give an accurate measure of the activation time of the underlying tissue. Moreover, their morphology is easily interpretable as the passing of a wavefront across an electrode. Unipolar EGMs can be also used for identifying the repolarisation of the tissue. The activation-recovery interval, which is the time duration between the maximum negative slope of the clinically recorded EGM and the maximum positive derivative occurring close to the repolarisation wave peak, has been suggested to correlate with the duration of the cellular AP (18,19). A similar time period recorded in an *in vitro* research setting and occurring between the minimum EGM deflection and the peak of the T-wave, known as field potential duration (FPD), has been suggested to be equivalent to APD₉₀ (20). However, a disadvantage of unipolar EGMs is that they include far-field electrical activity, making the analysis of unipolar EGM morphology more difficult. In particular, when analysing atrial signals, far-field signals from the larger ventricular depolarisation is often present (21).

1.2.3 Bipolar electrogram

In order to obtain bipolar EGMs, two electrodes are placed close to each other, usually within a few millimetres distance. This approach removes the far-field signal because this is similar at the two electrodes, leaving the local activity. The bipolar EGM approximates the first derivative of the unipolar EGM (15). Activation time in bipolar EGMs is typically defined as the time of maximum amplitude, although there are other definitions in the literature (22).

The main disadvantage of bipolar EGMs is that the signal is dependent on the direction of the wavefront. In the case that the wavefront is parallel to the bipolar electrodes, the wavefront cannot be detected on the resulting bipolar EGM. In addition, it is more difficult to determine the activation time and to interpret the morphology of the signal. Bipolar EGMs are also sensitive to inter-electrode distance (23).

1.2.4 Electrocardiogram

The ECG is a non-invasive measurement of the electrical activity of the myocardium using electrodes placed on the body surface. The output from each pair of electrodes is known as a lead, with a total of twelve leads being used. The angle of view of the lead depends on the electrode location. For example, lead I is measured from the left arm to the right arm and is defined as having an angle of view of 0° (24). The main components of an ECG are the P wave which corresponds to the atrial depolarisation, the QRS complex which is the ventricular depolarisation, and the T wave which represents the ventricular repolarisation. Atrial repolarisation is hidden within the larger ventricular depolarisation and

therefore it cannot be seen on the ECG leads (25). Some of the benefits of using ECGs, compared to EGMs, in the clinic are that it is an inexpensive and non-invasive method, it is relatively easy to apply, it is highly informative and can provide additional electrical information, which may correspond to conditions such as the chest pain and shortness of breath, and it can detect any silent cardiac conditions as a result of a surgery. However, its main disadvantage is that it provides a static picture captured from a relatively long distance from the heart surface and it may not reflect severe or localised underlying cardiac issues (26).

1.3 Circumstances that cause complex electrogram morphology

Extracellular EGMs, which are invasively recorded directly from the heart, provide information about the functional state of the myocardium. The basic shape of an extracellular unipolar EGM is simple in the healthy myocardium. However, EGM morphology may be much more complex under pathological conditions, often consisting of multiple components and long duration. These are characteristics of abnormal conduction and arrhythmias in the underlying myocardium (4). Therefore, changes to the EGM morphology can be due to either functional abnormalities or the presence of non-conductive areas (27).

1.3.1 Structural determinants influencing electrogram morphology

A tissue-structure related parameter influencing EGM morphology is the anisotropy. Cardiac tissue is composed of fibres making it anatomically and electrically anisotropic (4). This is caused by the combination of cell morphology and the electric coupling between cells, mediated by gap junctions. The result of anisotropy is faster propagation in the direction of fibres, as opposed to the much slower propagation in the perpendicular direction of the fibres. The unipolar extracellular morphology at sites where the activation front runs in parallel to the fibre orientation is a biphasic deflection at the site where the activation passes (4). However, EGM morphology is more complex when a wavefront is recorded at sites where activation propagates orthogonal to the fibre orientation (28). Sudden changes in fibre orientation, leading to discontinuous conduction and complex EGM configuration, have been also observed in fibre disarray (29).

In addition, complex EGMs may arise in regions with tissue discontinuities. These are usually areas with a sudden change in conduction velocity. The abrupt increase in the diameter of a myocardial bundle leads to a source-sink mismatch at the connection site. The current delivered by a small region may be

insufficient to trigger homogeneous activation of the thicker bundle resulting to conduction delay and a complex EGM (30). Such tissue discontinuities may be also present in regions where tissue types with different excitability or coupling characteristics are connected. An example of this case is the interface between myocytes and fibroblasts. Fibroblasts in between strands of myocytes are able to propagate the AP only after a large conduction delay (31).

Another structural element known for affecting the electrophysiology, and subsequently the EGM morphology, is fibrosis. Fibrosis is the formation of elevated levels of fibrous connective tissue, such as during reparative or reactive processes, and it is quantified through the presence of collagen. Collagen is the main component of the acellular part of connective tissue (32). There are different types of fibrosis with variable degrees of density, from focal and compact, which is the case of scars, to patchy and diffuse (Figure 1.2) (32). This leads to separation of strands of myocardium which impacts on action potential propagation, forcing excitation waves to take anisotropic, circuitous paths (33) that may be responsible for the generation of re-entrant activation (34) (Figure 1.3). Fibrosis creates areas of conduction block and defines anchor locations for re-entrant activation (35,36). However, it is still unclear how fibrosis is involved in the generation or the maintenance of abnormal electrical conduction (32). Despite this fact, scar locations have been suggested to be responsible for complex fractionated atrial electrograms (CFAE) (37). The CFAE-targeted catheter ablation, which was considered to also target scar areas, has limited success though and it is a debatable method of treatment (32).

Structural remodelling may occur in cardiac diseases and it involves changes in cell size, enhancement of collagen deposition and myocardial fibre disarray (4). Collagen deposition, and subsequently increased fibrosis, are a major component in the remodelling during a cardiac disease (38,39). The increased collagen deposition is linked to different forms of fibrosis (32), the electrical separation of myocytes and the reduced ability to support coherent wave fronts.

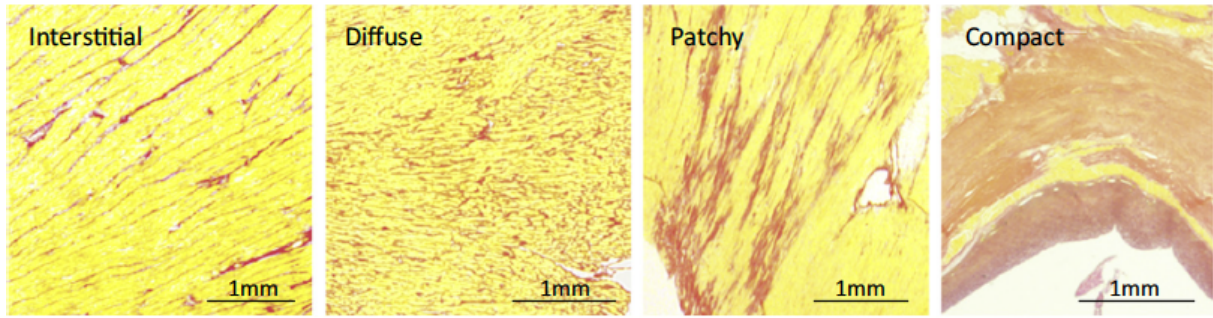


Figure 1.2: Types of human cardiac fibrosis in explanted hearts, as visualised with light microscopy. Collagen-dense areas are shown with red and these are stained with picroserious red. Interstitial fibrosis is the accumulation of collagen among groups of myocytes. Diffuse fibrosis is characterised by interspersed collagen among myocardial fibres. Patchy fibrosis separates myocardial strands over relatively long distances between them. Compact fibrosis involves the formation of large collagen-dense areas completely devoid of myocytes. Adapted from Rog-Zielinska *et al.* (32) with permission of the rights holder, Elsevier.

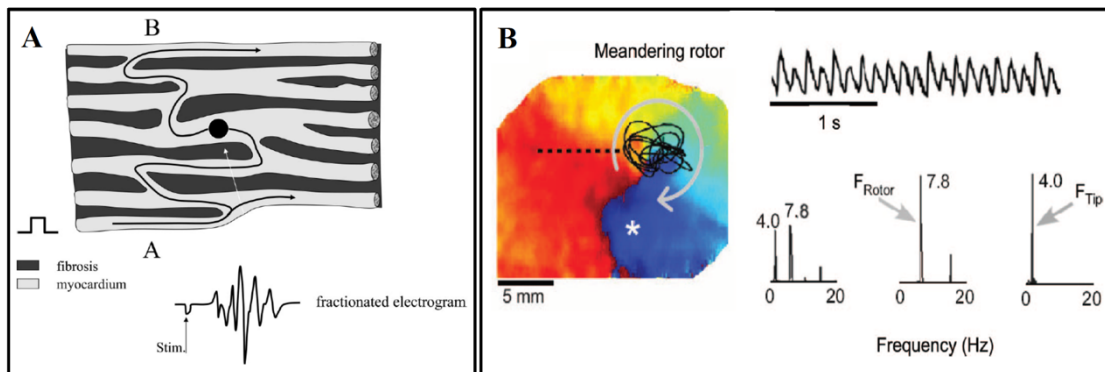


Figure 1.3: Determinants of EGM morphology. (A) Activation induced at site A of an infarcted myocardium propagates following the indicated zig-zag route due to the complex distribution of myocardial bundles. (B) The presence of re-entrant activity in the LA leads to continuous activation patterns. This leads to a fractionated EGM. (Top right) Optical signal recorded at the asterisk. (Bottom right) Spectra derived from the optical signal (left), the periodic constituent (middle) – reflects the periodicity of the rotor - and the residual constituent (right) – reflects the tip meandering component of the rotor. Adapted from de Bakker *et al.* (4) with permission of the rights holder, Wolters Kluwer Health, Inc.

1.3.2 Functional determinants influencing electrogram morphology

Functional changes leading to a modified EGM morphology are ion channel and gap junction abnormalities (4,40). Gap junctions play a key role in the AP propagation by electrically connecting neighbouring cells (41,42). They are protein complexes that function as mediators of cell-to-cell coupling, they allow ion exchange between cells and the transfer of small molecules of up to 1kDa by directly linking the cytoplasm of neighbouring cells (43–45). The term “gap junctions” indicates the gap between adjacent cells, as there is a 2-3nm distance between them, proved by electron microscopy images.

Two connexon hemichannels form a gap junction and each connexon consists of six connexin subunits. Connexins are cell membrane proteins with four transmembrane domains, two extracellular loops and one intracellular loop. The main connexins expressed in the heart are connexin 43 (Cx43), Cx40 and Cx45. The expression of each connexin is region-specific in the mammalian heart, as showed in Figure 1.4. Each connexon can be composed of combinations of six Cx43, Cx40 or Cx45 subunits and each gap junction channel may consist of combinations of two connexons. Thus, a connexon can be homomeric, when all connexins are the same, or heteromeric, when two or more connexins are different (46). Due to the differential Cx expression, the resultant action potential propagation is also varied in different parts of the heart, according to the localised physiological needs, i.e. differences in atrial and ventricular areas.

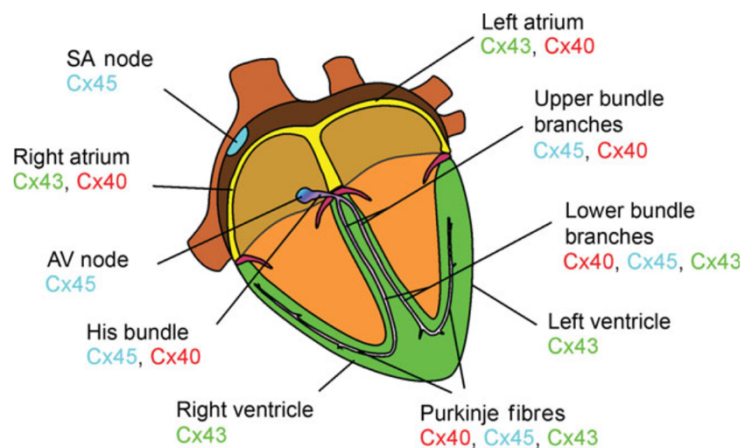


Figure 1.4: Summary of typical connexin expression patterns in mammalian heart. Cx43, Cx40 and Cx45 are the main cardiac connexins, but their distribution varies throughout the heart. Adapted from Severs *et al.* (43) with permission of the rights holder, Oxford University Press.

Gap junctions also reduce repolarisation heterogeneities between adjacent myocytes and between different regions of the myocardium (47). Gap junction function alterations or remodelling of gap junctions are known to occur in almost all arrhythmogenic cardiac pathologies, including acute ischaemia (48,49), the mature, healed infarct (50) and the failing myocardium (51). Alterations in gap junctions lead to conduction slowing and prolonged APD, both of which increase the likelihood for arrhythmogenesis and are correlated with increased fractionation in unipolar extracellular EGMs (52).

Ion channels are essential for the normal electrical and mechanical function of the heart (53). For example, loss-of-function mutations of K^+ channel genes or gain-of-function mutations of the SCN5A gene, which is responsible for the expression of the α -subunit of the Na^+ channel, are related to the long QT syndrome (LQTS) (53). This gene has also been correlated with other cardiac disorders, including

AF, Brugada syndrome and idiopathic ventricular fibrillation (54). Na^+ channels play a key role during phase 0 of the action potential, because when cardiac myocytes are excited by electric stimuli from the conducting cells or neighbouring cells, the membrane potential will rapidly depolarise and then Na^+ channels become activated (53).

The depolarisation (phase 0) is followed by phase 1 of the AP, due to the closing or inactivation of Na^+ channels but also because of the transient outward K^+ current (I_{to}) and outward Cl^- current. There are two identified classes of I_{to} : a) I_{to1} , which is independent of intracellular Ca^{2+} and sensitive to 4-aminopyridine; and b) I_{to2} , which is Ca^{2+} -dependent and selective to Cl^- (55). The I_{to} channels are formed by four α -subunits of the K_v4 family together with the K_v subunit KChIP2C (55). Variations of I_{to} current density are region- and species-specific. These variations lead to profound changes in the magnitude and time course of I_{CaL} resulting in corresponding changes in sarcoplasmic reticulum Ca^{2+} loading, release and contractility (56).

The plateau (phase 2) of the action potential arises from the balance of the inward and outward current crossing the cell membrane. The inward current is induced by the opening of the slow-activated Ca^{2+} channels and the outward currents carried by Cl^- channels and various K^+ channels (53). The L-type Ca^{2+} channel family has four members which are expressed by the CACN gene family. Mutations in these genes have been found to promote early and delayed afterdepolarisations and the LQT syndrome. It has also been reported that mutations in the L-type Ca^{2+} gene family are responsible for shortening of the QT interval in families characterised by sudden cardiac death, AF and Brugada type I ECG pattern. Therefore, the clinical condition known as short QT syndrome has been introduced, in order to include these phenotypes (57). Moreover, reduced I_{CaL} leading to loss of the AP plateau phase and reduced APD is known to be related with AF (58).

The inactivation of slow inward Ca^{2+} and Na^+ channels and the opening of various slow outward K^+ channels lead to the plateau phase (phase 3) of the action potential. The K^+ currents include the inward rectified K^+ current (I_{K1}), the ultra-rapid delayed rectifier K^+ current (I_{Kur}), rapid and slow delayed rectifier K^+ currents (I_{Kr} and I_{Ks}), acetylcholine-regulated K^+ current (I_{KACh}) and ATP-sensitive K^+ current (I_{KATP}) (59). Malfunction of K^+ channels, due to either congenital-encoded gene mutations or drug blockade has been found to alter not only the cardiomyocyte excitability, but also the electrical balance of depolarisation and repolarisation. Thus, these abnormalities usually cause a long QT interval and underlie a variety of cardiac arrhythmias. For this reason, K^+ channels are important targets for anti-arrhythmic drugs (59).

1.4 Atrial Arrhythmias

1.4.1 Overview

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia in humans with a prevalence of around 20% in people over the age of 85 (60). It is major risk factor for stroke and is correlated to an increased risk of morbidity and mortality (61). Current anti-arrhythmic drug therapies have limited efficacy for clinical suppression of AF. Many of them have been proved to be less specific against atrial electrical activity leading to serious pro-arrhythmic effects (62). Catheter ablation is a superior method of treatment of AF with the long-term success varying between 50-80%. However, this depends on various factors: a) the type of AF, as the efficiency is lower for persistent AF than paroxysmal, b) the ablation strategy, i.e. pulmonary vein isolation gives better outcomes, and c) the number of procedures, as the AF freedom rate is around 80% for patients treated with multiple AF ablations, compared to 50-60% following a single procedure (62). Therefore, a better understanding of the AF mechanisms that are involved in the initiation and maintenance of the disease is necessary and it may help the development of new therapies or the improvement of the existing ones. There are multiple hypotheses for the mechanistic basis of AF and different mechanisms involved in arrhythmogenesis and AF sustainability have been suggested.

1.4.2 Mechanisms of atrial fibrillation

The mother rotor hypothesis is one of the mechanisms believed to be involved in AF. According to this theory, the electrical activity during AF is characterised by rotational activity, termed *rotors*. These are defined as the organising source of functional re-entrant activity (63). AF is not entirely random, but there are hierarchical periodic rotors that drive the AF and act as sources of high frequency wavefronts (64). However, the presence of rotors in humans is not thoroughly proved and their existence is a current topic of debate (65), even though they have been detected in animal experimental models (66). Narayan *et al.* (67) though presented a clinical computational mapping tool, which is reported to identify rotors and focal sources in the human atria, which were then ablated. Their studies use the simultaneous collection of data from the majority of the atrial surface using a 64-pole catheter basket, while most of the other centres use only sequential mapping technology. Another study though suggested that the use of the ablation strategy developed by Narayan *et al.*, which provides information about the localisation of rotors and is supposed to enable targeting rotors during catheter ablation, showed poor efficiency (68). Therefore, this raises questions about the limitations of applying such technologies during catheter ablation, and even the usefulness of rotors.

Another mechanism suggested to be responsible for the AF maintenance is the presence of focal sources. These have been also detected in patients and targeted during catheter ablation in the CONFIRM (Conventional Ablation with or without FIRM) trial (67). A non-invasive method including a vest of electrodes, known as ECG imaging, has been used for the detection of focal sources, as well as electrical rotors (69).

A third theory for AF maintenance involves the presence of multiple random wavelets. Under this hypothesis, AF is considered anarchical and disorganised and the activity is self-sustaining and independent of the initiating events (70). The disadvantage of this theory though is that it cannot explain the origins of the activity that causes the generation of the wavelets. If there were a number of wavelets, then these should join together and obliterate AF (64).

It is clear that there are conflicting hypotheses about the mechanism of arrhythmogenesis. Studies using low-resolution global data acquisition report hierarchical rotors and focal sources, while studies that use high-resolution mapping catheters tend to suggest disorganised multiple wavelets.

1.4.3 Electrical remodelling in AF

Electrical remodelling occurs during AF and is one of the multiple factors affecting arrhythmia susceptibility, making AF a complicated disease. Multiple ion currents are altered during AF and these include I_{K1} , I_{KACh} , I_{Na} , I_{CaL} , I_{to} and I_{Kur} . I_{CaL} is usually reduced in chronic AF. This change affects the AF plateau phase, but this effect is smaller compared to the one due to I_{K1} (71). I_{K1} is the main determinant of the resting potential and it has been shown to increase in chronic AF. The increase of I_{K1} leads to a reduction of spiral wave meander and re-entry stabilisation, based on a computational modelling study (71). The same ion current change also increases dominant frequency, which occurs due to changes in AF and its rate dependence (72).

1.4.4 Dynamics of intracellular Ca^{2+} cycling

Intracellular Ca^{2+} cycling is another factor implicated in the complexity of AF with regional variations, which may lead to discordant alternans implicated in arrhythmia. It has been suggested that APD alternans and Ca^{2+} alternans usually happen simultaneously and the myocytes that were more prone to alternans suffered from disturbances in intracellular Ca^{2+} cycling (73). A combined clinical and modelling study showed that the differences between control and heart failure patients in the rate dependence of AP voltage can be predicted using a mathematical model with abnormal Ca^{2+} handling

due to reduced dynamics in the sarcoplasmic reticulum (74). Regional heterogeneities in the expression of Ca^{2+} cycling proteins and repolarising currents is an explanation suggested for the progression from concordant to discordant alternans. Another study, based on guinea pigs, found that transmural differences in sarcoplasmic reticulum Ca^{2+} cycling proteins take place and that could be the molecular basis for spatially heterogeneous susceptibility to AP alternans (75).

1.4.5 Structural remodelling in AF

Atrial myocytes are modified during AF and these changes include an increase in cell size, myolysis, perinuclear accumulation of glycogen, alterations in connexin expression, fragmentation of sarcoplasmic reticulum and changes in mitochondrial shape (76). These modifications are induced by prolonged rapid atrial pacing during AF. In addition, atrial fibrosis has been well documented in patients with lone AF involving severe hypertrophy (77). Electrical uncoupling and slow conduction can be provoked by structural remodelling, enhancing AF continuation without inducing changes in atrial action potential properties. The changes that take place during structural remodelling are less reversible and they persist even after the re-establishment of sinus rhythm (78).

Fibrosis plays a key role in AF and its importance has been highlighted through quantifications that link the development of fibrosis on the atrial substrate with its effects on conduction slowing (79,80). The extent of fibrosis has been shown to increase the risk of re-entry and wave break (81). However, it is unclear from some of these studies whether structural remodelling is the cause or the effect of AF. A computational modelling study reported that structural discontinuities in a model of cardiac tissue increased the probability of alternans occurring at high pacing frequencies since the random barriers to excitation cause large spatial differences in activation time distribution (82). It was suggested in the same study that this phenomenon provided a substrate for re-entry formation.

1.5 Ventricular Arrhythmias

Ventricular arrhythmias are cardiac rhythm disturbances that occur in cardiac ventricles. This term often refers to the two most common ventricular arrhythmias which are ventricular tachycardia (VT) and ventricular fibrillation (VF). It may also encompass arrhythmias such as premature ventricular beats and bigeminy (83). VT is defined in humans as three or more consecutive beats that originate from the ventricles and the rate is greater than 100 beats per minute (bpm). ECG complexes obtained from a VT patient usually have a consistent QRS morphology, which is the characteristic of monomorphic VT. However, there are cases in which the beat-to-beat morphology of the QRS complex changes and this is

what defines polymorphic VT (83). Another way to categorise VT is based on duration. Non-sustained VT self-terminates within 30 seconds while sustained VTs lasts longer than 30 seconds.

VF is defined as a rapid, irregular rhythm with more than 300 bpm. It is characterised by a variable QRS complex, amplitude and cycle length (83). An example of the variability of the signal is presented in Figure 1.5B. Ventricular tachycardia often degenerates into fibrillation leading to VF. VF is also haemodynamically unstable and is correlated to a cessation of cardiac output.

Ventricular arrhythmias are responsible for the majority of 100,000 and 300,000 sudden deaths per year in the UK and USA respectively (84,85). Among the pathologies which are associated with ventricular arrhythmias are dilated cardiomyopathy, hypertrophic cardiomyopathy, congenital heart disease, inherited cardiac arrhythmia syndromes and chronic coronary artery disease (86,87).

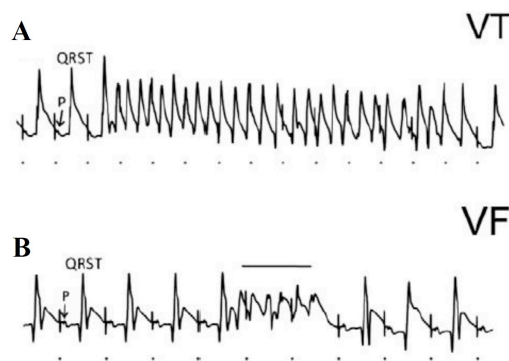


Figure 1.5: Example of unipolar EGM recording collected from mouse Langendorff-perfused hearts with VT (A) and VF (B). Hearts were paced at 10Hz from a bipolar electrode on the right atrium and the pacing artefacts (100ms apart) are marked with black dots below each EGM. P waves and QRST complexes are labelled for reference. The time scale is different between recordings. Adapted from Walker and Curtis (88) with permission of the rights holder, Oxford University Press.

1.5.1 Ventricular arrhythmia mechanisms

The three main mechanisms for ventricular tachyarrhythmias are the re-entry formation, the triggered activity and the enhanced automaticity (89–91).

Re-entrant activation occurring in the ventricles is the same phenomenon that takes place in the atria. It is the phenotype during which the propagating wavefront fails to terminate and instead re-excites cardiac myocytes that have recovered from their previous refractory state. A perpetual re-entrant circuit is set up due to this re-excitation through which cardiac excitation can continue to propagate. Re-entrants are categorised into anatomical, which take place around an anatomical obstacle, and functional re-entrants, where the anatomical obstacle is absent. Functional re-entrant activity may be formed based

on the leading circle theory, according to which the impulse propagation occurs around areas of refractory tissue, often referred to as areas of functional block (92). A second form of functional re-entry is the spiral wave. Spiral waves, also known as rotors, differ from the leading circle, because its core is excitable but not excited. They are implicated as the main cause both for tachycardia and fibrillation (42,93).

Triggered arrhythmias are the result of abnormal impulse generation, due to afterdepolarisations. Afterdepolarisations are extra depolarisations that occur following the initial depolarisation of a propagated action potential. If the afterdepolarisation is large enough to reach the threshold potential for activation of a regenerative inward current, then it may trigger further action potentials, which may be either single action potentials or a series of them (94,95). Afterdepolarisations are divided into early afterdepolarisations (EADs) and delayed afterdepolarisations (DADs) based on timing of the abnormal depolarisation. The abnormal depolarisation occurs during phase 2 or 3 of the action potential in EADs. EADs are often in the context of action potential prolongation or slowing, such as during bradycardia, LQTS or Brugada syndrome. AP prolongation leads to the time- and voltage-dependent recovery of the L-type Ca^{2+} channels which are inactivated following the initial depolarisation (94). The opening of a limited number of L-type Ca^{2+} channels leads to further inward currents that create the oscillations that are seen in phase 2 of the action potential during EAD. DADs occur during phase 4, after repolarisation is completed and before another action potential occurs. Cytosolic Ca^{2+} overload is a reason leading to DADs and this is often in pathological conditions such as heart failure, during acute reperfusion following ischaemia, during exposure to catecholamines or during digitalis intoxication. The increase of cytosolic Ca^{2+} concentration enhances the forward mode activity of the sarcolemmal Na^+ - Ca^{2+} exchanger (NCX), extruding one Ca^{2+} ion in exchange for three Na^+ moving into the cell (96).

Enhanced automaticity is the third mechanism which is responsible for tachyarrhythmias. Automaticity is a normal property of myocytes in the sinus node, the AV node and the His-Purkinje system. The basis of this is a spontaneous, gradual fall in the resting membrane potential during phase 4 of the action potential. This is the result of an interplay of several ionic currents, including the I_f current (97). Automation can be enhanced by alterations in the ion current activity, such as during sympathetic activation, leading to an increase in the rate of action potential generation.

1.6 Contemporary methods for evaluation of the atrial substrate in persistent AF

The identification of sites which are crucial to the maintenance of AF is essential to effectively treat AF using catheter ablation. The morphology of the contact EGM is the primary tool of the cardiac electrophysiology for identifying such sites. Its analysis leads to the detection of distinct atrial regions which can be targeted during catheter ablation. The identification of various functional and structural parameters, such as ion channel blockade and the presence of extended fibrotic areas respectively, defining the AF substrate is challenging, due to the fact that this work has been mainly carried out in humans. Thus, the “learning-by-burning” concept is increasingly used for investigating how EGM morphology is affected by the atrial substrate.

1.6.1 Voltage mapping

It has been suggested that atrial voltage mapping may capture the electrophysiological health of the underlying atrial substrate in AF. This is supported by the fact that the presence of low voltage areas can predict the outcomes of catheter ablation of AF (98). The voltage recorded in AF exhibits spatio-temporal variability highlighting the challenges of voltage mapping in AF (Figure 1.6). This led though to the adoption of an approach of mapping in sinus rhythm or during pacing, in order to study the underlying substrate more accurately. Low voltage areas are more often observed in permanent AF than in paroxysmal AF, as reported by studies that utilise voltage in paced rhythm (99) or in AF (100). Apart from the clinical variability described so far, there is also regional variation of low-voltage areas in AF patients. According to histological studies, it has been found that low-voltage regions are more often in the septum and posterior wall of the left atrium (LA) (101). However, the relationship between the anatomical location of scar and the AF drivers is yet to be elucidated.

Some of the discrepancies in the interpretation of endocardial voltage mapping could be explained by the differences between unipolar and bipolar EGMs where the electrode orientation and its interaction with the propagating wavefront are crucial factors for the EGM voltage and morphology (102). The majority of clinical studies use bipolar voltage mapping, and thus the interaction between the activation wavefront and the bipoles may be the reason for the variability in recorded values.

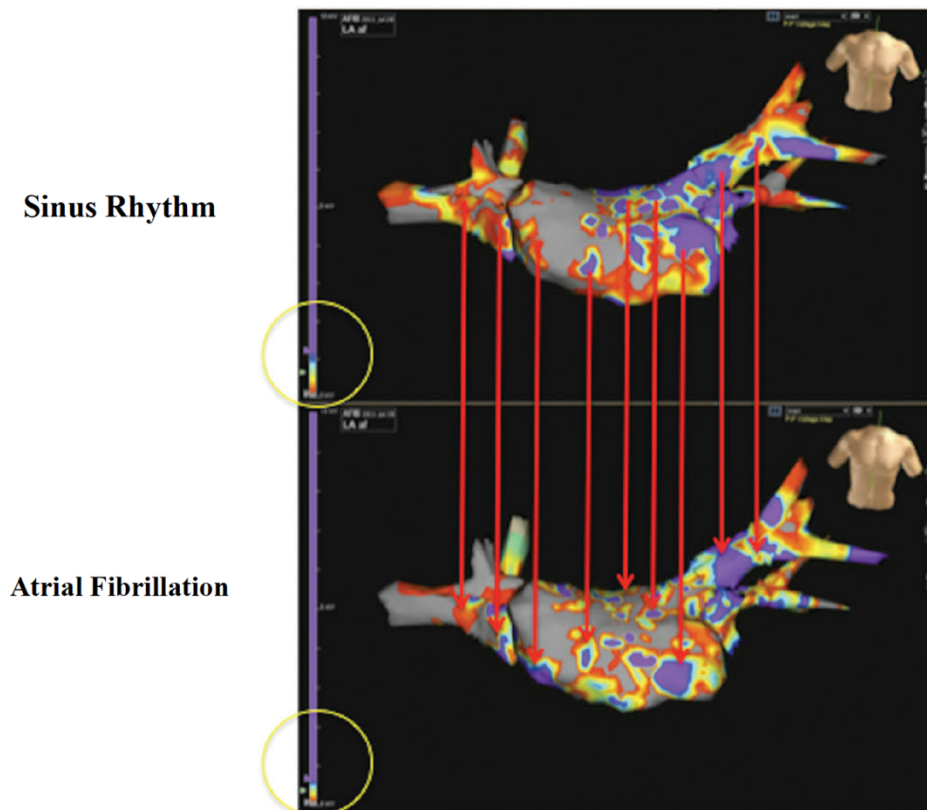


Figure 1.6: Voltage gradient maps in sinus rhythm and atrial fibrillation. The overall voltage was adjusted demonstrating lower voltage in atrial fibrillation. Low-voltage bridges are detected in similar locations in both cases. The variation in the voltage gradient maps is unavoidable, due to the variation in endocardial surface sampling. However, the large-scale structures remain similar. Red arrows connect voltage regions which are similar in sinus rhythm and atrial fibrillation. Adapted from Bailin *et al.* (103) under the terms of the Creative Commons Attribution License.

1.6.2 Complex Fractionated Atrial Electrograms

CFAEs, such as the ones presented in Figure 1.7, are one of the characteristic features of the underlying substrate in AF. They have been observed in intra-operative epicardial mapping of human AF and are suggested to co-exist with slow conduction areas and pivot points where the wavefront activation turns around at the end of arcs of conduction block (104). This has been confirmed by a study which suggested that the EGM morphology during AF reflects the occurrence of various specific patterns of conduction (105). It was also postulated in the same study that the EGM morphology might be useful for the identification of regions with conduction abnormalities, which are also responsible for the sustainability of AF, due to structural modulations (105).

However, there is a mixed clinical outcome following the use of CFAEs during catheter ablation. It was reported by Nademanee *et al.* that 70% of a cohort with equal numbers of paroxysmal and persistent

AF patients remained arrhythmia-free a year after the ablation procedure (106). However, many other clinical centres have not managed to reproduce these results. For instance, Oral *et al.* reported a success rate of 33% maintaining the sinus rhythm without medical therapy at a mean follow-up period of 14 months after the CFAE-targeted ablation in 100 persistent AF patients (107). CFAE-targeted ablation has also been used as part of a multifaceted ablation strategy, but still with mediocre success compared to the results published by Nademanee *et al* (108,109). It has also been shown through a meta-analysis of CFAE ablation in addition to pulmonary vein isolation (PVI), that CFAE ablation increases the success rate of AF in persistent AF patients, but without significant benefit for paroxysmal AF patients (110). The relationship between CFAE and AF sites is poorly defined, as shown by Narayan *et al* using FIRM mapping (111). They reported that CFAE sites lie remote from AF sources and are not suitable targets for catheter ablation. Moreover, other groups demonstrated that CFAE regions could be detected in healthy tissue regions, as determined by voltage mapping in sinus rhythm combined with late-gadolinium enhanced cardiac magnetic resonance imaging (LGE-CMRI) results (7,112).

It becomes clear that the usefulness and the role of CFAE ablation in modification of the substrate is controversial, due to discrepancies in clinical outcomes and the lack of reproducibility. This variability is enhanced by the fact that the targeted areas seem to not correlate, in terms of structural abnormalities of the substrate, with CFAEs. The stability of CFAE in the context of a spatio-temporally variable rhythm is also debated. It has been suggested that an EGM signal of longer than 5sec duration is required to accurately characterise an area as a CFAE site, while recordings less than this can be misleading in targeting the appropriate sites and are not correlated with acceptable success rates (113). Fractionation has been observed to be transient in recordings at ganglionic plexi sites and at the posterior LA wall, which is contrary to other studies (114).

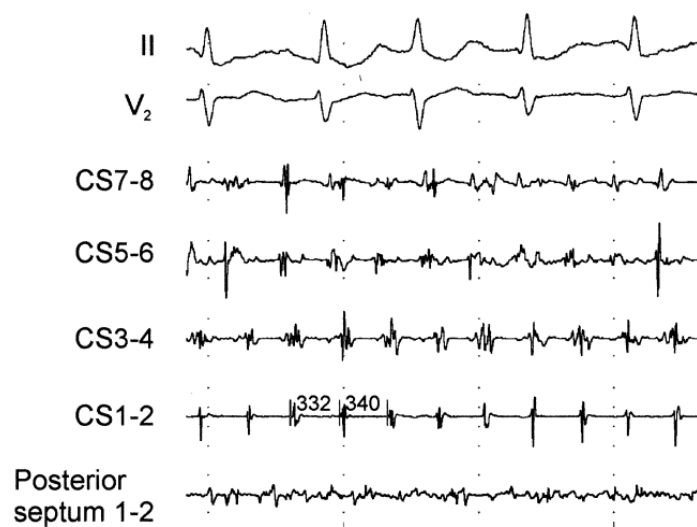


Figure 1.7: Example of CFAEs with continuous prolonged activation complex over the posterior septal areas. CS = coronary sinus. Adapted from Nademanee *et al.* (106) with permission of the rights holder, Elsevier.

1.6.3 Dominant Frequency

Dominant frequency refers to the frequency with the largest amplitude at the power spectrum of the EGM (115) (Figure 1.8). The spatial hierarchical organisation of local dominant frequency of the atria is believed to represent the fibrillatory conduction as a result of propagation waves emerging from focal sources which subsequently interact with anatomical and functional obstacles (116). Dominant frequency is pathophysiologically distinct from CFAE (116). It has been shown that when dominant frequency and CFAE are concomitantly mapped, areas corresponding to each one of them show limited overlap (117). Dominant frequency was used to overcome the limitations due to the chaotic nature of electrical activity during AF causing difficulties on observing the underlying patterns of depolarisation that may be important in the substrate for AF sustainability. This is due to the complex or noisy EGM obtained under these circumstances, such as the example of bipolar signal recording presented in Figure 1.8. Thus, dominant frequency is used to detect areas of high frequency activation with greater accuracy, through the analysis of EGMs in the frequency domain by Fast Fourier Transform (FFT) (118,119).

High dominant frequency regions have been detected in animal (120) and human (121) paroxysmal AF and especially at the pulmonary vein junctions and gradients toward the left and right atria. The distribution of high dominant frequency regions is more widespread in persistent AF. These are mainly detected outside the pulmonary vein regions (122). Dominant frequency analysis has been suggested to provide a mechanistic insight into understanding the persistence of AF and targeting these sites. The elimination of a left atrium-right atrium gradient is a marker for long-term freedom from both paroxysmal and persistent AF (122). However, these results are debatable, as other studies have demonstrated mixed outcomes with no additional benefit over pulmonary vein isolation alone in patients with paroxysmal or persistent AF (123).

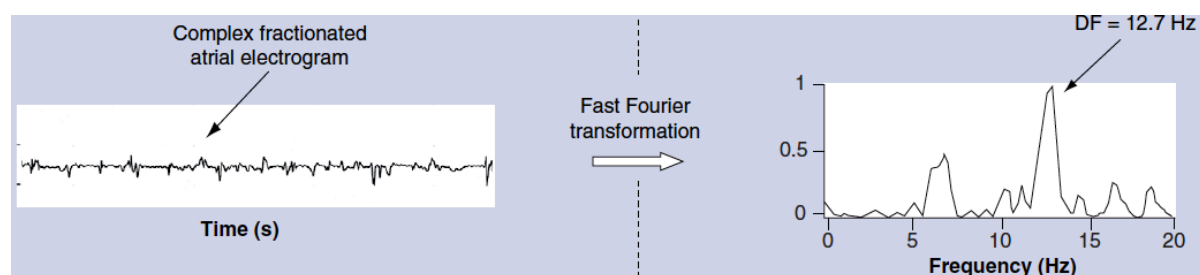


Figure 1.8: Bipolar recording of CFAEs (left) and the corresponding power spectra (right) following Fast Fourier Transformation from an LA site. The dominant frequency is shown in the power spectra and it was computed to be 12.7Hz. Adapted from Katritsis *et al.* (124) with permission of the rights holder, Taylor & Francis.

1.6.4 Shannon entropy

Shannon entropy is a statistical measure of information content based on the distribution of amplitude values within the signal histogram (125,126). This has been recently used for the measurement of the information uncertainty of bipolar EGMs which are related to the spatial uncertainty of activation at the pivot point of curved wavefronts in AF (125). This idea is reflected to the relationship between the bipole with maximum Shannon entropy and the pivot zone. Therefore, it is suggested that Shannon entropy represents the differences in bipolar EGMs between the pivot and periphery at sites of rotational activity. This is used as an alternative mapping method for pivot identification. It has been suggested that high Shannon entropy signals have similar appearance to CFAEs, as concluded from sheep AF data (127). This is an alternative mapping method assisting mapping of locally stable rotors in AF (125).

1.6.5 Late-gadolinium enhanced cardiac magnetic resonance imaging

LGE-CMRI appears to be a method for detecting atrial structural remodelling, even though the level of its accuracy is controversial. It is used for delineating left atrial fibrosis *in vivo* and evaluating the structural remodelled substrate in AF patients (128). This type of imaging has been applied on ventricular myocardium and it has proved to be an effective method to identify *in vivo* fibrotic non-viable myocardium in coronary artery disease and cardiomyopathies (128). This technique is based on the principle that gadolinium has a differential uptake and washout kinetics in blood, healthy myocardium and myocardial scar (129). Ablative scar sites are represented by bright scar tissue with LGE-CMRI, but native scar is less enhanced and therefore these areas are darker on the final image. LGE-CMRI has shown to correlate to clinical parameters which are used as surrogate markers of LA structural remodelling, such as left atrial size and the persistence of AF (130). It has been reported that the evaluation of collagen content with Masson trichrome staining correlated with structural remodelling on LGE-CMRI on 10 patients with AF history (131). However, the use of this method is only limited to few specialist centres, where its use has been mainly evidenced.

1.7 Machine Learning

Electrograms provide information about tissue activation. However, there is limited experience on thoroughly analysing EGM morphology and correlate this with the underlying pathology. This challenge can be overpassed by finding previously unknown relationships. To do this, the knowledge can be sought by learning directly from the data using machine learning algorithms.

Machine learning is the scientific discipline extracting patterns from data (132). It arises at the intersection of statistics, which seeks to learn relationships from data, and computer science (9). It is a field within Artificial Intelligence (AI) which has been very successful in developing software for voice recognition, computer vision, natural language processing and robot control among other applications. A variety of machine learning algorithms have been developed, in order to cover a wide variety of data and problem types. The large variety of machine learning algorithms is the result of the way they represent candidate programs and the necessity for different ways to search through the space of programmes (133). The main categories of machine learning methods are supervised machine learning, unsupervised machine learning and reinforcement learning. There are also modern blends across these categories. Such paradigms are the semi-supervised learning, which uses unlabelled data to handle labelled data in a supervised learning context, and discriminative training architectures developed for unsupervised learning with optimisation formulations that use labelled data (133).

1.7.1 Supervised Machine Learning

Supervised machine learning is the type of machine learning solely used in this thesis. The function approximation problem for these methods is the production of a prediction y^* in response to a query x^* using training data that take the form of a collection of (x,y) pairs. The inputs x are the training data and they may be numerical data or complex objects, i.e. documents, images, graphs. There has been much progress on the simple binary classification problem in which y takes on one of two values, even though there have been recent improvements on problems such as the multiclass classification (where y can take one of N labels) and general structured prediction problems (where y is a combinatorial object such as a graph, whose components are required to satisfy a set of constraints). The extracted model at the end of the training process can be applied to unknown data to predict the classification or outcome of each datum. This is expected to be done accurately once sufficient and suitable training data are used with an appropriate model (134).

Supervised learning is based on the formation of predictions via a learned mapping $f(x)$, which produces an output y for each input x . There are a variety of methods for this such as decision trees, decision forests, logistic regression, support vector machines, neural networks and Bayesian classifiers (135). There are a variety of algorithms for the estimation of these mappings and there are also generic procedures that combine the outputs of multiple learning algorithms, such as boosting and multiple kernel learning. The diversity of learning methods reflects the variety of their applications (133).

Features can be used to describe a signal recording and they are often represented by numbers which form feature vectors. Feature-based machine learning algorithms are an area of study within supervised machine learning and they use feature vectors both during the training process and the prediction stage. The selection of instructive features is critical for the effectiveness of a machine learning algorithm. However, using a large number of features may be counter-productive due to including less informative features in the training process or due to redundancy of information between features. Therefore, this issue can be eliminated by selecting a feature subset, which can enhance the ability of a training algorithm for accurate predictions (134).

There are a variety of feature-based supervised learning algorithms. There are linear and non-linear classifiers. The linear ones are usually faster and more efficient in time constraint applications compared to the non-linear classifiers (136). However, the linear classifiers are generally less accurate. All classifiers use a weighted sum, as shown in equation 1.1, to make predictions,

$$d(x) = f(\vec{w} \cdot \vec{x}) = f\left(\sum_i w_i x_i\right) \quad (1.1)$$

where d is a set of classes (the output of classification), \vec{w} is a real vector of weights, \vec{x} is the input feature vector to the classifier and f is a function that converts the dot product of the two vectors into the desired output. In other words, the output score equals to the function f that maps the result of the weighted sum onto the different classes with w_i , the weights learnt during training and x_i the training data. The weights can be computed using different approaches. Support vector machines search for the maximum margin between the hyperplane and the two data classes that are separated (137). Linear discriminant analysis computes weights that best separate inputs of different classes (138).

Modern approaches involve the combination or improvement of simpler methods, such as decision trees, in order to enhance their efficiency. They may also help to overcome the over-fitting problem, which is a problem that often occurs with decision trees, because they fail to generalise (134). Bootstrap aggregation (Bagging) is an example of one such method (139). Random forests extend the bootstrap aggregation of decision trees by selecting random feature subsets when deciding how to split at each node of a tree (140).

Also, a high-impact area of study within supervised machine learning is deep neural networks. This field involves multilayer networks of threshold units, each of which computes a simple parametrised function of its inputs (141,142). Deep networks use gradient-based optimisation algorithms to adjust

parameters throughout a multi-layered network based on errors at its output. The internal layers of deep learning networks can be considered as providing latent representations of the input data. There are efforts to develop deep learning algorithms that discover useful representations of the input without the need for labelled training data (143). The general problem in this case is known as unsupervised machine learning (144).

1.7.2 Unsupervised Machine Learning and Reinforcement Learning

Unsupervised machine learning involves the analysis of unlabelled data under assumptions about structural properties of the data. There are dimension reduction methods, such as principal component analysis, manifold learning and random projections (135,144), that make different specific assumptions regarding the underlying manifold.

Reinforcement learning is another major field of machine learning (145,146). This considers that the information of the training data is intermediate between supervised and unsupervised learning. The training data in reinforcement learning are assumed to provide only an indication as to whether an action towards learning is correct or not, instead of using training data that indicate the correct output for a given input. In case that an action is mistaken, the problem for the investigation of the correction action remains.

1.7.3 Machine learning applications in cardiac electrophysiology

Some of the various applications of machine learning techniques in the medical field include the detection or segmentation of anatomical structures (147), registration (148) and image recognition (149). Especially in the field of cardiology, machine learning offers the ability to make a variety of predictions. These are mainly related to improvements in the diagnosis and treatment of cardiac arrhythmia and cardiac disease events. A variety of data have been used for the development of predictive models, including 2D echocardiography (150), ECG (151,152) and EGM (153,154). The increasing need for automated high-throughput analysis of cardiac electrophysiology data from the cell membrane to the organ level is apparent, but remains unsolved (155).

Regarding the use of cardiac electrophysiology digital signals, there is poor experience from literature on the development of predictive models based on contact intracardiac EGM data (156). Some of the most recent studies include the prediction of AF sub-types using a variety of features extracted from *in silico* or clinical contact EGMs (153,157) and the automated location of *in silico* re-entrant drivers using EGMs (158).

However, ECGs are widely used for classification purposes, because they are non-invasive widely available recordings which have the ability to provide adequate information about cardiac function (156). A number of studies have been published presenting algorithms for the analysis of ECGs and the subsequent use of these data for classification into different types of AF (159). In most cases the ECG analysis includes de-noising and correction of baseline wander prior to feature extraction. Neural networks (160–162) and recurrent neural networks (160,163) are increasingly preferred for classification training, but accurate classification results have been also reported using ensembles of decision trees (random forests) (164), multi-level binary classifiers (165) and least-squares support vector machine classifiers (166). Most recent advances in this field include the combination of classification approaches for more accurate predictions (161), the real-time feature extraction and classification using ECG data (167) and the diagnosis of specific cardiac abnormalities (168).

Another application of machine learning in cardiac electrophysiology is relevant to the AF treatment. The combination of CFAE targeting with pulmonary vein isolation is a preferable strategy, as it is believed to reduce the long-term recurrence of AF (38). However, this approach is debatable due to the uncertainty on interpreting CFAE morphologies (36). For this reason, there are published methods such as the semi-supervised mapping (153) and the application of supervised learning algorithms that use features derived from combined extraction approaches (154), which can locate the distribution of different fractionated EGM patterns on the atria and improve the outcome of ablating drivers that maintain AF in persistent AF patients. Hajimolahoseini *et al.* (169) recently presented an algorithm to extract the active intervals of intracardiac bipolar EGMs during AF. The analysis included the characteristics of the signal waveform at its inflection points (IPs), which are the points at which the concavity of intracardiac EGMs changes from being convex to concave or vice versa. The signal features included the distance between the consecutive IPs, the slope of signal waveform at the IPs and the energy concentrated between them, which were all inspired by the way the electrophysiologists annotate the EGMs visually. The data were then used for an Expectation Maximization algorithm for Gaussian mixtures for automatic clustering. The important outcome of that study was the significant reduction of the mean computational time to 31ms per 1s recordings obtained from 10 channel EGMs which makes the algorithm able to be applied in real-time clinic applications.

In addition to AF, VF and VT are studied for improvement of their diagnosis applying machine learning techniques. Both VF and VT are dangerous arrhythmic events leading to sudden death if no defibrillation shock is applied to the subject within a few minutes (170). A number of algorithms have been proposed for the efficient VF/VT detection based on processing the ECG signal, mainly in the

time-domain (171). The reason for this is that correct detection and classification of these arrhythmic events is of extreme importance for an automatic external defibrillator and patient monitoring.

1.8 Outline of thesis

The central hypothesis of this work is that EGM morphology is determined by identifiable features of local action potentials and the architectural determinants of their propagation and, inversely, the local electro-architecture in cell monolayers, tissue slices and the whole heart may be predicted from a decomposition of the EGM morphology using supervised machine learning methods.

In order to address the hypotheses, there were a number of specific aims to the work described in this thesis:

1. Modification of the action potential morphology and conduction velocity using a variety of ion channel blockers on cell monolayers, in order to investigate subsequent changes in EGM morphology.
2. Induction of propagation discontinuities by introducing cellular scar, in order to record fractionated EGMs.
3. Modulation of inter-cellular coupling on cell monolayers and tissue slices, in order to cause discontinuous conduction.
4. Carrying out simultaneous optical mapping and EGM recordings, followed by histological analysis, using cell monolayers and myocardial slices, in order to correlate action potential and EGM morphology changes.
5. Correlation of scar areas with specific EGM feature modifications using *in vivo* data.
6. Development of unipolar EGM datasets obtained during experimental conditions with an increasing complexity (cell cultures – *in vitro*, tissue slices – *ex vivo*, human patients – *in vivo*), in order to investigate how the addition of an increasing number of factors may modify the EGM morphology. Also, datasets need to consist of an adequate size of data, in order to be able to proceed to machine learning and achieve high classification performance.
7. Development of algorithms for EGM feature extraction using data obtained from cell monolayers, tissue slices and patients.

8. Investigation of supervised machine learning techniques for predicting functional and structural abnormalities

Chapter 1 provides the background of the purpose of applying machine learning techniques, in order to seek a better understanding of the correlations between unipolar EGM morphology and pathological states. It details the electrophysiological substrate based on the contact EGM morphological analyses, such as fractionation, voltage and dominant frequency, and how this type of recording is compared to other types of data – bipolar EGM, ECG and LGE-MRI. Since the purpose of this thesis is the improvement of treatment against atrial and ventricular fibrillation, these cardiac arrhythmias are presented alongside our current knowledge about the EGM morphology that we obtain under each condition.

Chapter 2 presents the biological techniques applied in this thesis. These include the methodologies that were followed, in order to set up the experimental models. Neonatal rat ventricular myocytes with fibroblasts and tissue slices derived from human healthy and heart failure cardiac samples were the *in vitro* and *ex vivo* models respectively. Persistent AF patients were also used for paced recordings. In addition, this chapter presents the experimental conditions for multi-electrode arrays and optical mapping, the main methods for electrophysiological recordings in this thesis, and techniques for investigating the histology of multicellular preparations.

The algorithms for processing the unipolar EGM recordings that were obtained from different experimental models form the basis of Chapter 3. These algorithms involve the time-domain, time-frequency and time-frequency domain analysis of the EGM, as well as the analysis of information theory features. Novel algorithms for morphological analysis were validated for their accuracy and consistency against manual measurements.

The influence of ion channel blockade on EGM morphology was studied, and results were presented in Chapter 4. Chapter 5 presents the effects of gap junction uncoupling on EGM morphology recorded from multicellular preparations and Chapter 6 shows the effects of structural changes on EGM morphology, using variable amounts of fibroblasts in cell cultures, studying the degree of fibrosis in tissue slices and comparing scarred and non-scarred areas recorded from the left atrium of a patient with clinical history of persistent AF. Regarding all the above chapters, the EGM signal obtained from cell cultures was correlated with concurrently recorded action potentials, in order to confirm the above changes.

Chapter 7 presents the application of machine learning techniques on *in vitro* data for predicting the variety of functional and structural modifications that were characterised in previous chapters. A detailed analysis for the optimisation of the process is also presented, in order to achieve high classification accuracy. This included the investigation of multiple supervised learning algorithms and optimisation of specific parameters in the algorithm which seemed to perform the best. Chapter 8 presents the capability of predicting factors that are responsible for a modified EGM morphology using the *ex vivo* and clinical data. Voltage, but not fractionation, was included in the feature subset for classifying EGMs obtained from tissue slices. This analysis showed that these two features are weak predictors, despite the fact that they are used in the clinic as markers for ablating locations which maintain AF. Both chapters showed that a combination of features extracted from different types of analysis achieved the optimum performance.

The effects of different heart failure phenotypes on EGM morphology and how this is different from EGM obtained from rejected donor transplants are presented in Chapter 9. It was shown that each clinical phenotype resulted to a distinct morphology with phenotype-specific EGM modifications. These differences helped the development of a prediction model which can successfully predict these phenotypes from unipolar EGM morphology and using supervised learning algorithms.

Chapter 10 summarizes the major findings from this thesis and expands on future directions for research. It suggests that supervised machine learning can be applied for predicting functional and structural determinants of EGM morphology using a combination of different types of data analysis. It also advises that supervised learning can be used for predicting heart failure phenotypes.

Chapter 2

Materials & Methods

2.1 Introduction

This chapter provides a general description of the materials and biological techniques used in this thesis. Methods covered in this chapter include EGM acquisition from *in vitro* and *ex vivo* models before and after the administration of ion channel or gap junction modulators, *in vitro* and *ex vivo* optical mapping of transmembrane voltage (V_m), acquisition of clinical recordings, immunofluorescence and second harmonic generation microscopy. Machine learning methods are also presented. Further specific details on methodology will be provided in the relevant chapters.

The methodology, developed by combining the methods presented in this chapter, aimed to address the questions stated in Chapter 1. A schematic representation of this methodology is showed in Fig.2.1. It presents the route from obtaining the biological data, and especially cellular electrophysiology and imaging data, under a variety of experimental conditions for the understanding of the effects of these abnormalities on EGM morphology and followed by the application of supervised machine learning for developing models which can predict the same abnormalities.

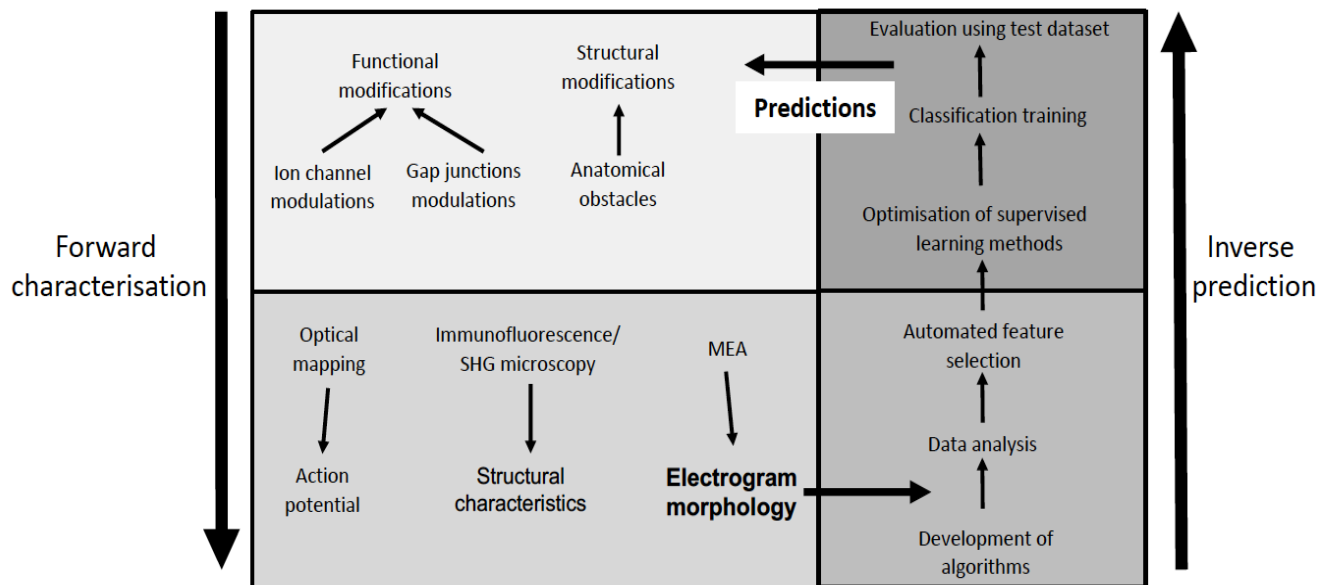


Figure 2.1: Overview of the methodology that was followed in this thesis towards the development of predictive models using supervised machine learning techniques.

2.2 Materials

2.2.1 Cell Culture

Neonatal rat ventricular myocytes (NRVMs) are preferred for studying the influence of functional and structural abnormalities on EGM morphology in this thesis. It is a well-established cell line that has been long used at cellular electrophysiology studies and for screening drug effects (40,172).

NRVMs were obtained from Sprague-Dawley rats (0 to 3 days old). All procedures were conducted according to the standards set by the EU Directive 2010/63/EU and were approved by the Imperial College London Ethical Committee. Anaesthesia by isoflurane was followed by euthanasia carried out by cervical dislocation and assessed by cessation of circulation. The heart was then removed and dissected immediately. The vessels and atria were carefully removed to isolate the ventricles which were further dissected into 1-2mm³ pieces. Enzymatic digestion of ventricles was carried out using the gentleMACS neonatal heart dissociation kit (Miltenyi Biotec GmbH). The tissue and enzyme mix was incubated three times at 37°C for 15 minutes each time before being attached to the gentleMACS dissociator for gentle agitation. The digested sample was resuspended in 10.5mL M199 10% cell culture medium. The cell medium consisted of 100mL M199 (catalogue no.: 22350-029; Gibco), 10mL newborn calf serum (catalogue no.: NB-112/500; Biosera), 10µM/mL penicillin-streptomycin (catalogue no.: P0781; Sigma-Aldrich), 0.68mM L-glutamine (catalogue no.: G7513; Sigma-Aldrich), 2µg/mL vitamin B₁₂. The cell suspension was then passed through a 70µm pre-separation filter and centrifuged at 1000 rpm for 5 min. The cell pellet was re-suspended in 20mL M199 10% culture medium. The natural composition (NC) cell population was isolated at this stage, including myocytes, fibroblasts and endothelial cells. Fibroblasts were removed by preplating NC cells for 1 hour at 37°C/1% CO₂. The remaining suspended cell population consisting of NRVM only was extracted after a final filtration through a 70µm pre-separation filter. Fibroblasts were incubated for 7 days at 37°C/1% CO₂ in T75 flasks and were provided with 10mL 10% M199.

In addition, HeLa cells (catalogue no.: 93021013; ECACC) were grown in MEM medium complemented with 1% non-essential amino acids (catalogue no.: 11140-035; Gibco), 10% Foetal Bovine Serum (FBS), 10µL/mL penicillin-streptomycin and 2mM L-glutamine. Cells were maintained at 37°C in a 5% CO₂-humidified incubator.

Microelectrode array (MEA) dishes (MultiChannel Systems MCS GmbH, Germany), which are described in more detail in section 2.3, were prepared in a class II microbiological safety cabinet. Plates were coated with 30µL 0.2mg/mL collagen over the electrode matrix and left for 2 minutes. The same

volume of double-distilled water from a Milli-Q Integral 3 water purification system (Merck) was used to wash each plate twice. The preparation was left in 1 mL double-distilled water for 1 hour and was then sterilised by UV light exposure for 1 further hour. A 30 μ L-cell medium drop with a total population of 200,000 cells was plated on each MEA dish. Cells were firstly rested for 30 minutes at 37°C/1% CO₂, then additional 10% M199 medium was used to bring total volume up to 1mL/plate. Cell cultures were incubated in the above conditions with 10% M199 for the first 24 hours and then a daily replacement of 1mL 5% M199 (prepared as above, but with 5mL neonate serum) until electrophysiological measurements were recorded.

2.2.2 Collection and preparation of tissue samples

Human left ventricular transmural tissue samples were prepared from end-stage heart failure samples and explanted healthy donor hearts. There were 11 heart failure patients who consented to take part in the study and one organ donor, this heart was unsuitable for transplant and it could be used for research purposes. The patients' demographics, indications for anti-arrhythmia treatment, the diagnosed heart failure subtype and the presence of a device implant are shown in Table 2.1. Table 2.2 also presents the reason for rejecting the donor sample as a transplant, alongside the donor's history for alcohol consumption, smoking and the demographics. Heart failure samples were immersed in ice-cold cardioplegia solution (Harefield Hospital formulation) immediately after explantation and transported to the laboratory within one hour. Donor transplants may arrive later, but always less than 3 hours after explantation. Samples with an approximately 6 x 6 mm surface area were obtained and mounted using tissue adhesive Histoacryl[®] (catalogue No.: 1050044; B.Braun Surgical S.A., Spain) with epicardium down onto the specimen holder of a high precision vibrating microtome (7000smz, Campden Instruments Ltd., UK). Samples were in cold (4°C) oxygenated (100% O₂) cutting Tyrode's solution (140 mM NaCl, 6 mM KCl, 10 mM glucose, 10 mM HEPES, 1 mM MgCl₂, 1.8 mM CaCl₂, 3g/L 2,3-butanedione monoxime (BDM), pH 7.4) during the preparation of 300 μ m-thick slices derived from the endocardial and epicardial sample areas (Figure 2.2). The advancement speed of the steel blade was 0.04mm/s, the amplitude 2 mm and the vibration frequency was 80Hz. Slices were incubated in oxygenated ice-cold cutting Tyrode solution for at least 30 minutes prior to being placed in the centre of a MEA plate and obtaining the electrophysiological recordings in Tyrodes solution (140mM NaCl, 4.5 mM KCl, 10 mM glucose, 10 mM HEPES, 1 mM MgCl₂, 1.8 mM CaCl₂, 1g/L BDM). This study was supported by the supply of human tissue samples from the Cardiovascular Research Centre Biobank at the Royal Brompton and Harefield NHS Foundation Trust (NRES ethics number for biobank samples: 09/H0504/104+5; Biobank approval number NP001-06-2015). Informed consent was obtained from each patient involved in this study. This study was also supported by the supply of human tissue samples from NHS Blood and Transplant (REC reference 16/LO/1568). Informed consent was

obtained from the donor's next of kin. All procedures described in this thesis were carried out in accordance to the Human Tissue Act 2004 (c30).

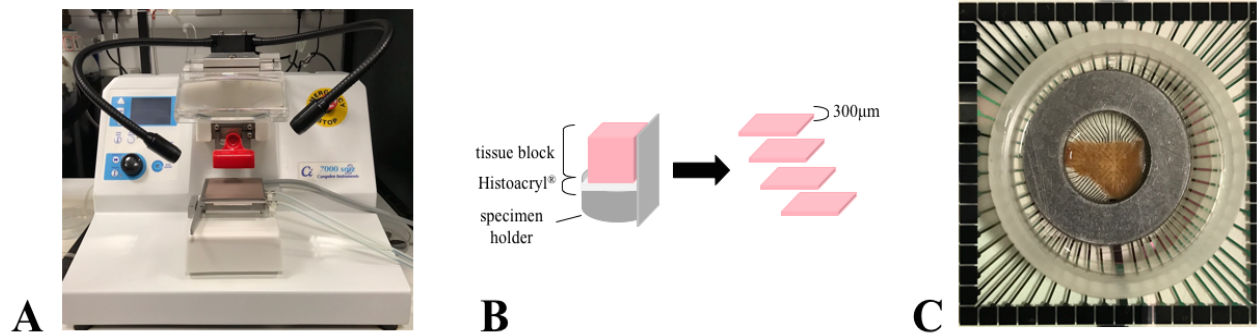


Figure 2.2: (A) The vibratome used for producing tissue slices. (B) A block of tissue is mounted on the specimen holder using Histoacryl® during slicing. The tissue is submerged in a waterbath full of cold and oxygenated Tyrodes solution. (C) A tissue slice placed on top of the electrode matrix of a micro-electrode array dish. A metal ring is placed on top of the slice, in order to keep it in contact with the electrodes while it is perfused.

Table 2.1: Heart failure patient characteristics. Slices derived from explants 1-9 could be stimulated, but not for slices derived from the rest. IHD: ischaemic heart disease; DCM: dilated cardiomyopathy; LVAD: left ventricular assist device

# of explanted heart	Age	Gender (M/F)	HF phenotype	Arrhythmia (Anti-arrhythmic drugs)	Ca ²⁺ antagonist	Device implantation
1	60	F	IHD	No	No	Yes (CRT-D)
2	57	F	DCM	Yes (amiodarone)	No	Yes (LVAD)
3	38	F	HCM	Unknown	Unknown	Yes (CRT-D)
4	65	M	IHD	No	No	No
5	28	M	DCM	No (but digoxin treatment)	No	No
6	32	M	DCM	Yes (digoxin)	No	No
7	67	M	IHD	No	No	No
8	23	M	Lymphocyte myocarditis	No	No	No
9	20	F	DCM	No	No	No (pre-LVAD patient)
10	50	M	DCM (sample was ischemic for 30min post explantation)	No	No	Yes
11	30	F	DCM	Yes (amiodarone)	No	Yes (LVAD + ICD)

Table 2.2: Characteristics of the donor of rejected heart transplant. OCS: organ care system

Age	Gender (M/F)	History of hypertension	History of cardiothoracic disease	History of smoking	Donor alcohol consumption	Note
48	M	No	No	No	Light drinker – 1-2u/day	Rejected due to degeneration on OCS.

2.2.3 Ion channel and Gap junction modulators

In the experiments described in this thesis, pharmacological ion channel modulators were used to modify ion channel activity. All agents were sourced from Sigma-Aldrich, unless otherwise stated. The K^+ channel blockade was achieved using 4-aminopyridine (4-AP; catalogue no.: 275875), which is an I_{to} blocker, E-4031 (catalogue no.: M5060), a I_{Kr} blocker, or HMR-1556 (Tocris Bioscience), which is a I_{Ks} blocker. In order to achieve the enhancement of K^+ channel activity, I_{KATP} activation was carried out using pinacidil (Enzo Life Sciences). Lidocaine (catalogue no.: L7757) was used for Na^+ -channel blockade, nifedipine (catalogue no.: N7634) for the blockade of L-type Ca^{2+} channels and ORM-10103 (catalogue no.: SML0972) for the inhibition of the Na^+/Ca^{2+} exchanger activity. To reduce gap junction coupling, carbenoxolone (CBX; catalogue no.: C4790) was used.

2.3 Microelectrode array recordings

The electrophysiological properties of cardiac myocytes were assessed using the USB-MEA60-Inv MEA system (MultiChannel Systems, Reutlingen, Germany). An MEA plate consists of 60 gold electrodes arranged on an 8 x 8 matrix (inter-electrode space: 700 μ m, electrode diameter: 100 μ m) with missing electrodes in the corners of the matrix (Figure 2.3). Stimulation was carried out using a STG stimulus generator programmed by MC Stimulus II software (version 3.4.4, MultiChannel Systems). A biphasic stimulus (2 ms duration; 120% of the threshold; voltage: 500-1000 mV) was applied for several minutes from the 6 electrodes located on one of the four external rows of the matrix to reach steady state before obtaining 10sec recordings in incremental rates until loss of 1:1 capture. In the cases of recording fibrillation, 30sec recordings were obtained. The temperature was kept stable at 37°C during stimulation and recording. Unipolar extracellular EGMs were recorded at a sampling frequency of 25kHz.

The responsiveness of cells to ion channel or gap junction modulators was tested 2-4 days after seeding cells on MEA dishes. Control recordings with NRVM administered with 1mL of Hank's buffered saline solution (HBSS) were obtained before changing the control solution for 1mL of incremental concentrations of each pharmacological agent. Signals were displayed and data were analysed offline using MC Rack software (v4.6.2, MultiChannel Systems).

A limitation of this technique is related to the fact that an electrode can only faithfully record up to a certain frequency, because of the nature of the electrode-solution/tissue interface, i.e. when current

passes through an electrode, hydrolysis occurs in the solution resulting to the formation of bubbles at the surface which reduce the frequency range, or when electrodes are in contact with tissue, impurities of the tissue may attach on the electrodes causing the same problems with frequency range. This restricts the ability to record at high frequencies and the highest sampling frequency would be the equivalent to AP upstroke duration – if AP upstroke is 1ms, this is equivalent to approximately 1kHz. The frequency limit imposed by the recording system, and therefore the MEA system, should be ideally at least 10 times that of the signal which is intended to be recorded. The frequency response of an electrode can be tested by injecting an artificial current through the MEA at different frequencies and record the voltage. If the artificial signals were sine waves, the amplitude would start to decrease as the frequency increases. In this case, the attenuation of signal would be assumed to be at greater frequencies than the recorded signal.

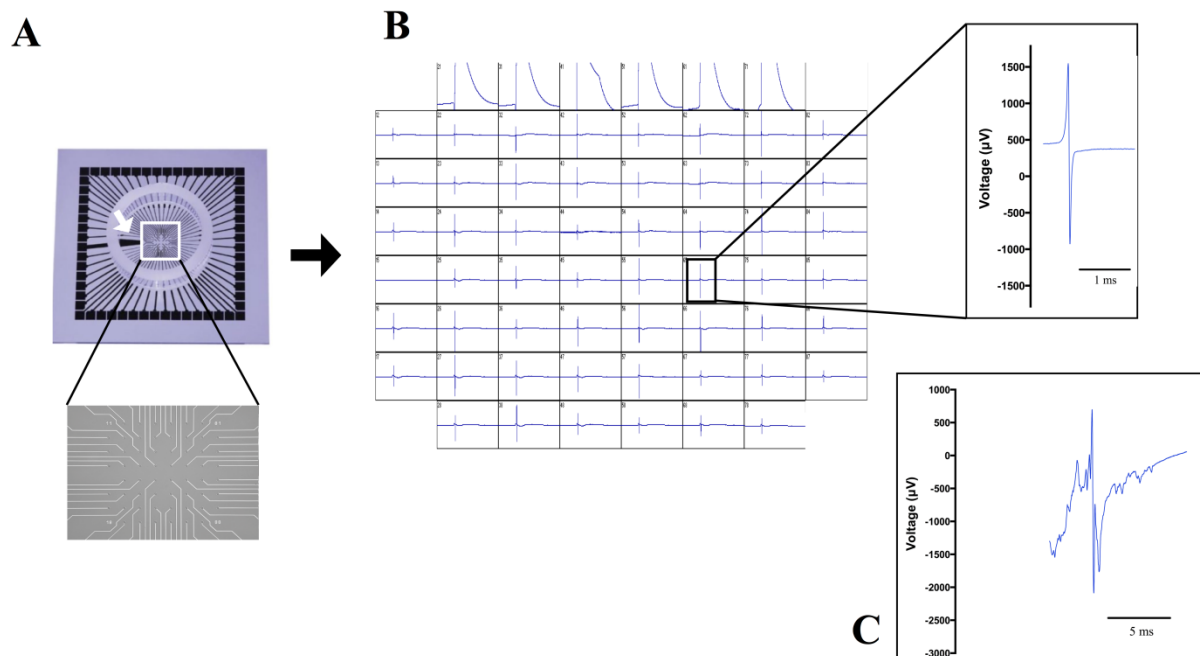


Figure 2.3: (A) 60-electrode MEA dish with the same electrode arrangement as the one used for obtaining EGM data for this thesis. The reference electrode (indicated with a white arrow) can be seen on the left side of the electrode matrix. This electrode arrangement was used mostly for unipolar extracellular recordings. (B) The MEA was used for recording EGMs, as presented in this screenshot. Right, Representative EGM acquired from NRVM monolayer. (C) Representative EGM obtained from a left ventricle tissue slice derived from a donor heart. This EGM was obtained using the same MEA dish as for data collected from cell monolayers.

2.4 Clinical recordings

Eight patients with symptomatic persistent AF, based on the AF classification guidelines published by AHA/ACC/HRS/ESC (Norman, 224), who presented for their first ablation to Imperial College Healthcare NHS Trust were prospectively enrolled for experimental purposes. The study was approved by the Local Research and Ethics Committee for Imperial College Healthcare NHS Trust and written informed consent were obtained from all patients.

Electroanatomic mapping technology records both electrical and spatial positioning data from the electrodes on an intra-cardiac mapping catheter. During this process, and prior to gathering mapping data, the catheter was systematically moved around the full extent of the chamber and spatial electrode position data were collected to generate a point cloud from which a surface representation of the endocardial surface was computed. Sequentially acquired unipolar electrical data were then collected and projected onto the surface.

The electro-anatomic data, which were used for the purpose of this thesis, were collected by Dr Norman Qureshi using a 20-electrode double-loop AFocusII high density mapping catheter (Inquiry™ AFocusII™, Abbott, USA) at a sampling frequency of 2034.5 Hz. Electrodes are 2mm in size with a circumferential spacing of 4mm (Figure 2.4). The catheter was placed in tangential contact with several regions of the left atrial endocardial surface, including the roof and the posterior wall. For each position, unipolar EGM data were collected concurrently from all 20 electrodes under paced rhythm at cycle lengths of 250ms, 300ms, 450ms and 600ms, as well as from a fixed decapole catheter positioned in the coronary sinus. The signals were also converted to units of seconds (t) using $t=n/f$, where n is the number of each recorded sample, the sample number, and f is the sampling frequency at 2034.5Hz. This type of analysis was developed by Dr Rheeda Ali and the outcomes of this analysis were also used for this thesis.

The patient-specific intensity map was derived from manual segmentation of the LGE-MRI and subsequently voxel intensities along a 3mm inward normal were projected onto the segmented atrial shell by Dr Norman Qureshi. Algorithms developed by Ali *et al.* (173) were used by Dr Rheeda Ali for registration between the electro-anatomical surface and the LGE-MRI surface. During this process, a non-rigid surface registration, initialised from the automatic landmark registration, provided the transformation to map the electrode locations onto the LGE-MRI surface. In addition, a method for operator-independent quantification of LGE that correlates with collocated EGM signals was applied using an in-house MATLAB script developed by Dr. Rheeda Ali. This technique was based on the construction of a scar map expressing intensity as multiples of standard deviation (SD) above blood

pool mean (174). The result was the correlation of EGMs with non-scar areas (<0 SD above blood pool mean) or scar regions (>2 SD above blood pool mean) (by Dr Rheeda Ali).



Figure 2.4: (Left) The 7F Inquiry AFocusII catheter with 4F double loop that was used for the acquisition of EGM data. The electrode spacing is 4mm and the loop diameter was 20mm. Adapted from www.cardiovascular.abbott/us/. (Right) The numbering and arrangement of electrodes on the AFocus II catheter.

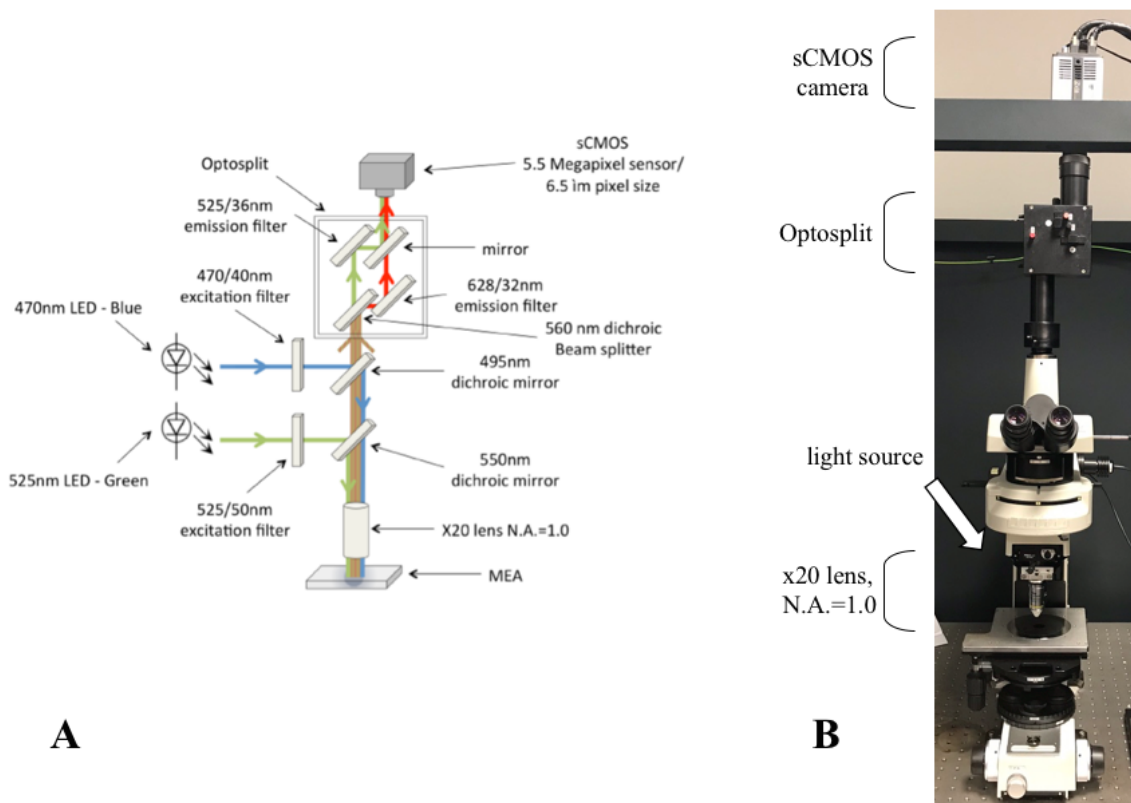


Figure 2.5: Optical mapping system. (A) Schematic diagram of the optical mapping system in which details related to the light source are visible. (B) A LED light source, located on the back side of the system, was directed towards the sample stained with voltage-sensitive dye. The emitted light passes through two dichroic mirrors. Both lights are detected by a sCMOS camera in the end.

2.5 Optical Mapping

Optical mapping is a non-invasive method for recording biological data simultaneously over multiple sites. One application of this technique is the optical mapping of transmembrane voltage (V_m) transients using fluorescent voltage-sensitive dyes, which were also used in this study.

2.5.1 Principles of optical mapping of transmembrane voltage (V_m)

An optical mapping system consists of four components: (1) the biological sample stained with a voltage-sensitive dye, (2) a light source, (3) a system of optics, and (4) a photodetector. During the optical mapping experiment, the sample is stained with a voltage-sensitive dye, which binds to the cell membrane. The dye is then excited by a light source with the appropriate excitation wavelength. When the dye returns to its ground state from its excited state, it emits photons of a wavelength that is longer than that of the excitation light. The system of optics then filters the emitted photons and focuses the light onto a photodetector, which quantifies the amount of emitted light. The change in the quantity of filtered light detected by the photodetector is proportional to changes in transmembrane voltage, as a result of the unique property of voltage-sensitive dyes. This results to the non-invasive assessment of transmembrane voltage changes.

2.5.1.1 Voltage-sensitive dyes

There is a variety of potentiometric dyes, which are designed to allow measurement of transmembrane voltage changes. Some examples include di-4-ANNEPS, di-8-ANNEPS and RH237. These dyes exhibit voltage-dependent shift of their emission spectra. When the dye is excited, at resting membrane potential, it will emit photons with a specific emission spectrum when returning to the ground state. However, when the cell is depolarised, the emission spectrum is blue-shifted towards shorter wavelength. By using appropriate filters to select for light above certain wavelength, a decrease in signal intensity will accompany membrane depolarisation.

2.5.1.2 Motion artefact

One of the main technical challenges of optical mapping studies is dealing with motion artefact. The vigorous contractions of tissue slice, or even the minor contraction of a cell culture, move the sample and prevent the recording of optical action potential signals from the same site over the course of a

cardiac cycle. Multiple strategies have been developed to reduce or eliminate motion artefact. Pharmacological agents are a popular strategy to remove motion artefact and have been used to uncouple cardiac excitation from contraction, thus removing motion without altering electrical activity. Excitation-contraction uncouplers include BDM, which is a non-competitive inhibitor of myosin II, and cytochalasin D, an agent that inhibits actin filament polymerisation (175). However, both of these agents have been shown to alter action potential, they have off-target pharmacological effects on ion channels and they modify restitution and Ca^{2+} handling (176,177). BDM was used during experiments with tissue slices in this thesis. Blebbistatin is a relatively new excitation-contraction uncoupler that acts as inhibitor of adenosine triphosphatases (ATPases) associated with class II myosin isoforms (178). It has no significant effects on pacemaker activity, conduction, repolarisation and intracellular Ca^{2+} cycling (179). However, there are reports showing that blebbistatin can prolong monophasic action potential durations, as well as increasing the maximum slope of restitution and altering fibrillation threshold (180).

There is a number of studies concentrated on applying the ratiometry approach for quantitative measurements of transmembrane potential with dyes. This method is based on simultaneous recordings of fluorescent signal at two different wavelength ranges, one where the dye exhibits a larger voltage-dependent response and another at a wavelength where the potentiometric dye exhibits an inverted or no voltage-dependent response. The ratio of these two signals results in an optical signal free of, or with significantly reduced, motion artefact. A weakness of this method though is that it works well with relatively weak contractions and with biphasic action spectra of fluorescent dyes (181). This method was used when acquiring recordings from cell monolayers in this thesis and it was combined with BDM when obtaining data from tissue slices.

2.5.1.3 Optics

Fluorescent light emitted from the biological sample must be filtered and focused onto a detector. The optics of an optical mapping system typically consists of a collecting lens, one or more dichroic mirrors, focusing lenses and filters. The light from the preparation passes through the collecting lens before being split with dichroic mirrors, which will reflect light below a certain wavelength, whilst allowing light above that wavelength to pass through. The light then passes through the appropriate emission filter and is focused onto the detector or camera.

2.5.1.4 Detectors

Detectors are necessary for the detecting the optical signals, measure their intensity and convert that into an electrical signal. Three main types of photodetectors are used for cardiac optical mapping experiments: (1) photodiode arrays (PDAs), (2) charge-coupled devices (CCDs), (3) complementary metal-oxide semiconductor (CMOS) detectors. Even though PDAs and CCD cameras are the two main forms of detectors, over recent years, CMOS cameras have been developed for optical mapping and these newer cameras have both excellent temporal and spatial resolutions. For the experiments described in this thesis, a sCMOS camera was used.

2.5.2. Optical mapping protocol

Optical mapping was carried out in combination with MEA recordings for the purpose of the work presented in this thesis. Figure 2.5 shows a schematic representation of the custom-made optical system used for simultaneous membrane voltage and MEA recordings (supplied by Cairn Research, UK). The system was built around an upright microscope (Eclipse FN1, Nikon Instruments Europe B.V.) with a modified stage height to hold the amplifier of the MEA system. Excitation light (470nm) was supplied by an OptoLED system (Cairn Research, UK) which provided controlled illumination and modulation. The light was passed through an emission filter of 470/40nm. Light was collimated to the MEA dish by a water dipping objective lens with magnification 20x (NA: 1.0 20x, XLUMPLFLN20XW PL FLUORITE OMJ, Olympus). Samples were stained with 40 μ M di-8-ANNEPS (Molecular Probes[®], Invitrogen) diluted in 1mL HBSS with 2.5 μ L Pluronic[®] F-127 (Life Technologies, USA).

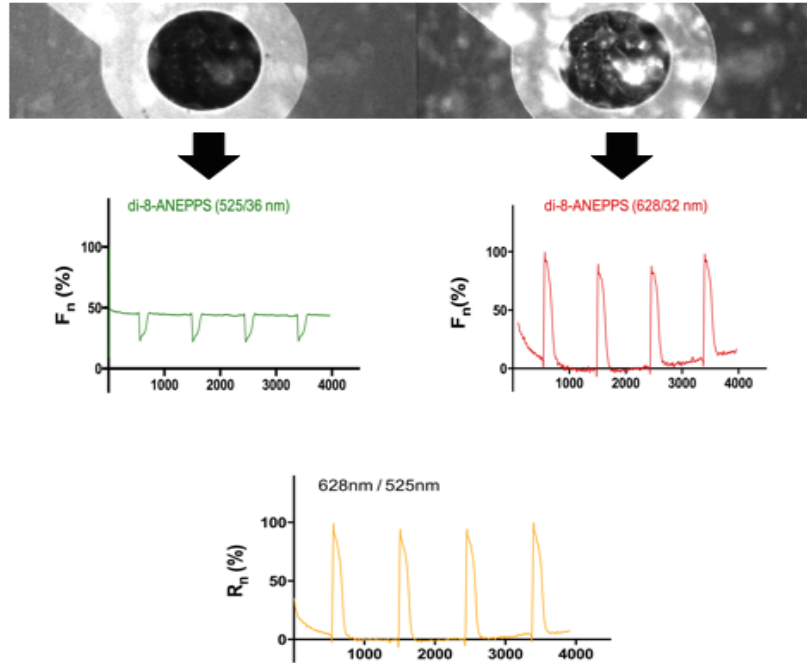


Figure 2.6: (Top) Part of a NRVM monolayer on top of a single electrode and stained with di-8-ANNEPS as captured using the optical mapping system. The same group of cells as seen with simultaneous excitation of green (left) and red light (right). Middle, Fluorescence intensity changes over time used for recording the AP. Simultaneously collected fluorescence at 525/36nm (green) and 628/32nm (red), both expressed as percentage change of normalised signal (F_n). (Bottom) Signal correction by 628/525nm ratiometry expressed as percentage of normalised R_n .

The emitted fluorescence was collected by the same objective lens and passed through dichroic mirrors. Using a 560nm edge BrightLine[®] single-edge dichroic beamsplitter, located in an Optosplit II ‘LS’ emission image splitter (x1.0 magnification), the fluorescent light was divided into two beams that were passed through emission filters (525/36nm and 628/32nm). Subsequently, the light was focused onto a complementary metal-oxide semiconductor camera (Zyla 10-tap sCMOS, Andor Technologies Ltd., Belfast, UK) for the detection of the dual wavelength optical signal with a spatial resolution of 400 x 885 pixels at 525.39 frames/sec. The light intensity measurements were recorded using the Andor Solis software platform (version 4.23.30008, Andor Technologies Ltd.).

However, the I_{Na} activation lasts at most 1-2ms meaning that the temporal resolution of the sCMOS camera may be inadequate in some occasions to properly capture the AP upstroke, as the camera may be able to record maximum 1.05 frames within 2ms, and unable to capture faster activations. Due to this fact and the low voltage stimulation leading to delayed initial depolarisation, as firstly described by Hodgkin and Huxley (475), the AP upstroke duration was measured to be longer than the actual I_{Na} activation. These results are presented in Chapters 4 and 5. However, the modulation of upstroke duration was correlated to conduction velocity and subsequent EGM duration prolongation, as recorded through simultaneous optical imaging and EGM recordings, presented in section 4.2.4. The limiting

factor of temporal resolution did not affect our results, as optical imaging was used in this thesis with the intention to show how changes in AP morphology directly manifest in unipolar EGM morphology, and the focus of this research is the analysis and accurate measurement of EGM features.

2.5.3 Optical mapping data analysis

Optical mapping data were analysed offline using GraphPad Prism (version 6.0f for MacOS X, GraphPad Software, La Jolla California, USA). The data were spatially smoothed using 3 averaging neighbours and 4th order of the smoothing polynomial. Action potential data were undertaken with emission ratiometry (181). The ratiometric signal analysis was achieved by using the fluorescence signals of opposing ΔF (green and red fluorescence). The ratio signal, collected at both wavelengths, was extracted for the attenuation of any motion artefacts, since motion artefacts can appear as a common change on both wavelengths and consequently, they can be cancelled out by the ratiometric calculation. As a result, the introduction of excitation-contraction uncouplers, such as blebbistatin, is not necessary with our system. Analyses were carried out using a custom-written macro in GraphPad Prism for drift removal and normalisation of raw data before calculation of AP morphology. Data analysis included the measurement of action potential duration at 50% and 90% of repolarisation (APD₅₀ and APD₉₀ respectively) and upstroke duration.

2.6 Immunofluorescence

2.6.1 Principles of immunofluorescence

Immunofluorescence is a method used to determine the location and distribution of specific proteins. It is based on the principle of specific interactions between an antibody and its corresponding target, the antigen, which is found on the substrate. The primary antibody is designed to bind specifically to the protein under investigation. This antibody is produced by an animal which was previously injected with a peptide sequence from the protein under investigation. The animal recognises the peptide as a foreign protein and raises an immune response against it. As part of this reaction, B-lymphocytes secrete antibodies that specifically target that peptide sequence. The antibody is then purified from the serum of the animal to be used against the protein of choice in immunofluorescence experiments. The incubation of the sample with the primary antibody is followed by the incubation with a secondary antibody that specifically targets the primary antibody. Secondary antibodies are conjugated to a fluorescent marker, which can be visualised under a microscope using an epifluorescence filter or by

confocal microscopy. The series of specific binding between the protein of interest, the primary antibody and the secondary antibody means that the locations of the fluorescent signal as visualised under the microscope correspond to the locations of the protein of choice. Although an excellent technique for protein localisation, some limitations include potential false-negatives due to poor antibody penetrance into tissue and false-positive due to non-specific antibody binding.

2.6.2 Immunofluorescence protocol

Some of the experiments related to immunolabelling NRVM/fibroblast co-cultures or NRVM cultures following the administration of lidocaine, HMR-1556 and pinacidil were performed by Efthymoulos Sokratous, Kai Wang and Alec Saunders, who were fourth-year Imperial College London medical students doing BSc projects under my supervision. Firstly, cell samples were washed once with Phosphate Buffered Saline (PBS) and then they were fixed with 100% methanol at -20°C for 5 minutes. The remainder of the protocol was carried out at room temperature. Following 5 washes with 1mL PBS, cells were incubated for 10min with 1mL permeabilisation solution containing 0.1% Triton X100 in PBS. Samples were washed again 5 times with 1mL PBS and then they were incubated for 1 hour with a blocking buffer consisting of 5% Bovine Serum Albumin (BSA), 0.05% Tween-20, 20% horse serum and made up in PBS. This was followed by an overnight incubation at 4°C with the primary antibodies against α -actin (diluted 1:1000; Abcam – catalogue number: ab137346), vimentin (diluted 1:5000; ThermoScientific – catalogue number: PA1-16759) and α SMA (diluted 1:1000; Dako – catalogue number: MO85127-2), which were used as markers of cardiac myocytes, fibroblasts and myofibroblasts respectively. All primary antibodies were administered in a blocking buffer solution. Following incubation with the primary antibodies, samples were again washed 5 times with PBS before incubation with the secondary fluorescent antibodies. A mixture of a goat Alexa Fluor[®]568-conjugated anti-chicken antibody (A11041), a donkey Alexa Fluor[®]488-conjugated anti-rabbit antibody (A21206) and a goat Alexa Fluor[®]647-conjugated anti-mouse antibody (A32728) were used in a dilution of 1:400 in PBS for 1hr incubation at room temperature. All secondary antibodies were purchased from ThermoFisher Scientific. Cell samples were then washed 5 times with PBS before being mounted with Vectashield H-1200 non-hardset mountant containing DAPI (Vector Laboratories, UK) and a 19mm glass coverslip was placed onto each cell culture. Immunolabelled samples were stored at 4°C in the dark to preserve fluorescence until use.

2.6.3 Cell Tracking System

HeLa cells were labelled with the short-term cell tracking dye CellTracker™ Red CMTPX (C34552, ThermoFisher Scientific), which is a fluorescent dye with excitation/emission wavelengths at 577/602nm. It is not a technique related to any antibodies, as the whole cell is stained, and it is more appropriate for monitoring the location of specific cells, as it is transferred to the daughter cells, but not the adjacent cells in a population. This method was preferred instead of the traditional immunofluorescence, due to the difficulty finding a protein-marker that could distinguish efficiently HeLa cells from NRVMs.

Firstly, in order to isolate the amount of HeLa cells that was needed for experiments, cells were treated with trypsin without ethylenediaminetetraacetic acid (EDTA) at 37°C/5% CO₂ for 5 min. The equal amount of trypsin inhibitor was then added, in order to end trypsin activity. Cells could be easily removed from the Petri dish where they were growing and they were centrifuged at 1000 rpm for 5min. The cell pellet was resuspended at 4mL 10% MEM medium and the appropriate number of cells was isolated from the resuspension. CellTracker™ Red CMTPX was supplied in lyophilised form containing 50µg powder, that was dissolved into 7.3µL sterile dimethyl sulfoxide (DMSO), in order to prepare 10mM stock solution. This solution was diluted to 5µM using serum-free MEM solution containing 1% non-essential amino acids, 10µL/mL penicillin-streptomycin and 2mM L-glutamine. The cell pellet of needed amount of HeLa cells was resuspended by 200µL of 5µM CellTracker™ Red CMTPX and cells were then incubated for 45min at room temperature in rotation and dark. This was followed by a 1000 rpm centrifugation for 5min and double wash with 10% MEM solution, in order to remove any excess dye. In the end, HeLa cells were mixed with NRVMs. The final cell population consisted of 5% HeLa cells and 95% cardiac myocytes. Cells were seeded onto MEA dishes and they were provided with 10% M199 solution supplemented with 0.12% 5'-bromo-2'-deoxyuridine (BrdU; B5002 – SAFC), in order to inhibit the uncontrolled cell growth of HeLa cells. Cell cultures were incubated at 37°C.1% CO₂ and they were provided with 5% M199 solution supplemented with 0.12% BrdU from the next day after seeding. HeLa cells retained fluorescent signal until 5 days after staining. NRVMs were not affected by CellTracker™ Red CMTPX or BrdU.

2.6.4 Analysis of target-protein distribution

Images were acquired with a Zeiss AxioObserver inverted widefield microscope using a 10x/0.3 objective lens and ZEN software (ZEN 2012, Carl Zeiss Microscopy GmbH). A custom-made macro for FIJI (Image J, v2.0.-rc-43/1.51a), developed by Dr Stephen Rothery (Facility for Imaging by Light

Microscopy), was used for the assessment of individual electrodes covered by target proteins or HeLa cells. The macro scanned the area of the MEA covered by electrodes, it detected each electrode and captured the electrode area, which was manually indicated in the beginning of the process and it was the same for all electrodes. A threshold was added on the image obtained from the whole cell culture or the stack of images of individual electrodes, and it was adjusted, in order to collect the signal coming from the target only. The target was the area covered by vimentin or HeLa cells in most cases. Based on that signal, the area covered by a target protein or cell type was measured as the percentage out of the total electrode area.

2.7 Second-harmonic generation microscopy

2.7.1 Principles of second-harmonic generation microscopy

Second-harmonic generation (SHG) is the optical phenomenon that takes place when the electric field of the exciting light is sufficiently strong to deform a molecule. If the molecule is not symmetrical, the resulting anisotropy creates an oscillating field at twice the frequency, the second harmonic (182). This means that there is greater ability to generate second harmonics when molecules are non-centrosymmetric. Collagen is the most important extracellular structural protein of vertebrates (182). The collagen molecule is non-symmetric and is arranged in a triple helix. As a result, collagen SHG microscopy has recently become a robust tool for imaging tissue structure with cellular resolution, in both *in vivo* and *ex vivo* preparations (183). Since fibrosis, which occurs under myocardial infarction or other reparative or reactive cardiac processes, is quantified through the presence of collagen (32), it is wise to use this molecule for imaging structural modifications of the cardiac tissue. Collagen I molecules, that are predominantly produced in fibrotic regions, are particularly interesting because they are thick, with diameters comparable to the wavelength of visible light (183).

2.7.2 Second-harmonic generation microscopy protocol

Tissue slice was placed on a 35mm-uncoated imaging dish with low walls (catalogue no.: 80131, ibidi). PBS was applied on the slice and a 19mm glass coverslip was placed on top, in order to keep the tissue slice flat. An adapted protocol previously described by Williams *et al.* (183) was used for the

visualisation of collagen fibres. Tissues were visualised using a Leica SP5 two-photon upright confocal microscope. This microscope was equipped with a Spectraphysics Mai Tai 690-1020 DeepSee multiphoton laser and a motorised stage. The multiphoton was tuned to 860nm and simultaneous dual imaging was performed for both collagen I (SHG) detection at 430nm and for cytoplasmic detail (autofluorescence) at 500-550nm. Z-series of images were collected with water dipping objectives (HCX IRAPO L25x/0.95W) throughout the slice. Images were acquired with a 25x magnification and 256x256 or 512x512 resolution and were processed with ImageJ (National Institutes of Health, USA).

2.8 Statistical analysis

Data are expressed as mean \pm SEM. Statistical analysis was carried out using GraphPad Prism software (version 6.0f for MacOS X, GraphPad Software). Statistical significance was evaluated using student's t-test (two-tailed), one-way analysis of variance (ANOVA) for unpaired data or ordinary two-way ANOVA followed by Bonferroni multiple-comparison post hoc analysis where appropriate. Baseline and post-modification data are unpaired, unless otherwise stated. The comparison of MEA recordings with optical mapping was validated with linear regression analysis (confidence interval: 95%) for correlation of corresponding characteristics measured with each technique. The APD₉₀/FPD relationship was compared between the control and modulated data groups by correlation coefficient. All numerical data are presented as mean \pm SEM and *p*-value of <0.05 was considered as significant.

2.9 Machine learning techniques

Machine learning refers to the process of training computational or statistical algorithms to perform pattern recognition, classification and prediction directly from example data. The main classes of machine learning are *supervised*, in which case the algorithms learn to make predictions from training data for which outcomes are provided, and *unsupervised*, which finds patterns in the data with more limited guidance (184). Other categories of machine learning methods are *reinforcement learning* and *semi-supervised learning* (184). This thesis is focused on supervised machine learning.

2.9.1 Supervised machine learning

Every instance of data used by machine learning algorithms is represented by the same set of features. The first step towards the development of a prediction model in feature-based supervised machine learning is constructing the training dataset. The data must be characterised by a set of chosen features

(10). An approach to this is that an electrophysiologist can suggest which features are necessary to best describe the data. Alternatively, a “brute-force” approach could be taken in which the measurement of a wide range of features, that characterise the signal, are computed in the hope that these include all the necessary information. However, the latter strategy may contain noise and significant pre-processing may be required (185).

A critical step in the process is choosing the classification learning algorithm. Once we are satisfied with the training performance and the preliminary testing, the classifier can be available for routine predictions using unlabelled instances to classes. A classifier is usually evaluated according to the *classification accuracy*, which is the percentage of correct predictions divided by the total number of predictions, and how well its general classification performance is. A popular method for the calculation of a classifier’s accuracy is cross-validation. In this technique, the training set is divided into mutually exclusive and equally-sized subsets and multiple rounds of training occur using a combination of randomly selected subsets. Cross-validation does assess accuracy, but its primary role is to test how well the algorithm generalises to unseen data. (10).

A total of 20 supervised machine learning algorithms were considered in this thesis. These belong to 4 families of machine learning algorithms: k-Nearest Neighbours, Support Vector Machines, Linear Discriminant Analysis and Decision Trees. The default MATLAB parameter values were initially used for screening them and searching the most appropriate method for the training dataset. Details of each method are presented below.

A. *k-Nearest Neighbours*: This method is based on the principle that the observations from the training dataset exist in feature-space to other instances with the same or similar features. In the case that instances are labelled with a class, then the label of an unknown instance can be estimated by the class of its nearest neighbours. The kNN method locates the k nearest instances to the target-instance and determines its class by the class with the highest frequency among the k neighbours (Figure 2.6). For this thesis, $k=1$ (KNN fine), $k=5$ (KNN medium) or $k=10$ (KNN coarse) neighbours were considered for the classification of new data with 10-fold cross-validation during the training process. The Euclidean distance between m number of observations x and y was measured using the following equation (2.1):

$$D(x, y) = \left(\sum_{i=1}^m |x_i - y_i|^2 \right)^{\frac{1}{2}} \quad (2.1)$$

The metric applied for the calculation of the relative distance between observations used normalized observations, in order to give them equal weighting. This process minimized the distance between two similarly classified observations and maximized the distance between observations belonging to different classes (10,186).

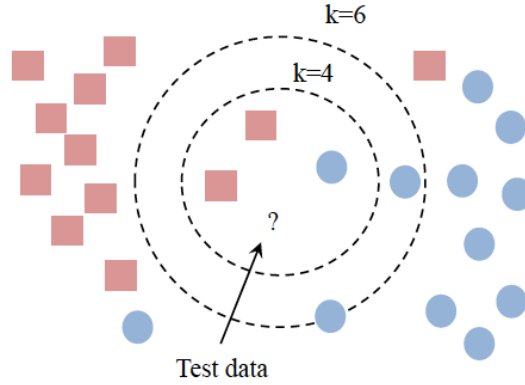


Figure 2.7: k-Nearest Neighbour illustration. The algorithm decides how new data are classified based on the class with the highest frequency in their vicinity. In this case, $k=4$ neighbours or $k=6$ neighbours may be considered for determining the class of the new data.

- B. *Support Vector Machines (SVM)*: SVM is based on the notion that a hyperplane is constructed separating two classes such that the separation, also known as margin, is maximized (10). Creating the largest possible distance between the separating hyperplane and the observations on either side of it has been proved to reduce the generalization error. For linear separation of data (linear SVM), a pair (w,b) exists such that:

$$\begin{aligned} w^T x_i + b &\geq 1, \text{ for all } x_i \in P \\ w^T x_i + b &\leq -1, \text{ for all } x_i \in N \end{aligned} \quad (2.2)$$

where w is the weight vector, b the bias and x_i is the input vector. P and N correspond to the margins on each side of the hyperplane (Figure 2.7). The discriminant function is given by equation (2.3):

$$f_{w,b}(x) = \text{sgn}(w^T x + b) \quad (2.3)$$

The hyperplane is found by minimising the squared norm of the separating hyperplane. Once the optimum hyperplane is computed, the data lying on its margin are known as support vector points and the solution is represented as a linear combination of only these points and the rest are ignored.

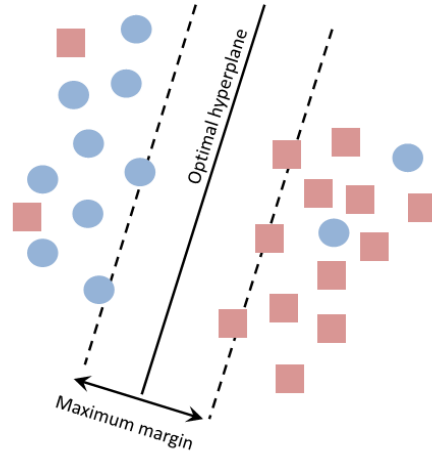


Figure 2.8: An illustration of Support Vector Machine (SVM) for binary classification. An SVM finds the hyperplane that separates the data points of one class from those of the other class. The optimal hyperplane is the one that separates the classes with the largest margin between the two classes. There are no data points within the margin, but they can be on its borders. Adapted from Zhang & Suganthan (187).

C. *Discriminant Analysis:* Both Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) were used in this thesis (Figure 2.8). They are both types of Bayesian classifiers and assume that the data have a Gaussian distribution. Since they are Bayesian classifiers, both methods predict according to the Bayes theorem, described by equation (2.4):

$$P(Y = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{l=1}^K f_l(x)\pi_l} \quad (2.4)$$

where k and l are two different classes. This means that the classifier was given an input X that is the feature vector x . x consists of P different predictors, one for each input dimension. Then, a covariance matrix $\Sigma_k = \Sigma \forall k$ is computed in LDA with respect to each possible output class, i.e. k, l . the classifier computes $P(Y = r \vee X = x)$, which is the probability that the actual output is r with x being the input. The class that returns the maximum value for the above-mentioned probability is given as the output of the classifier. Regarding LDA, $\hat{P}(X \vee Y)$ was estimated, instead of estimating $P(X \vee Y)$, in order to compute how data points were distributed in the p -dimensional space (135).

The difference between LDA and QDA is that while LDA assumes the same covariance matrix Σ for all k classes, there are no assumptions on the covariance matrices Σ_k in QDA (135). Essentially, LDA computed a separate μ_k for each class using data points that belonged to it,

but Σ was computed using the entire training data. And this common covariance matrix was used in the computations corresponding to every class. QDA, on the other hand, computed a separate Σ and μ_k for each possible output class.

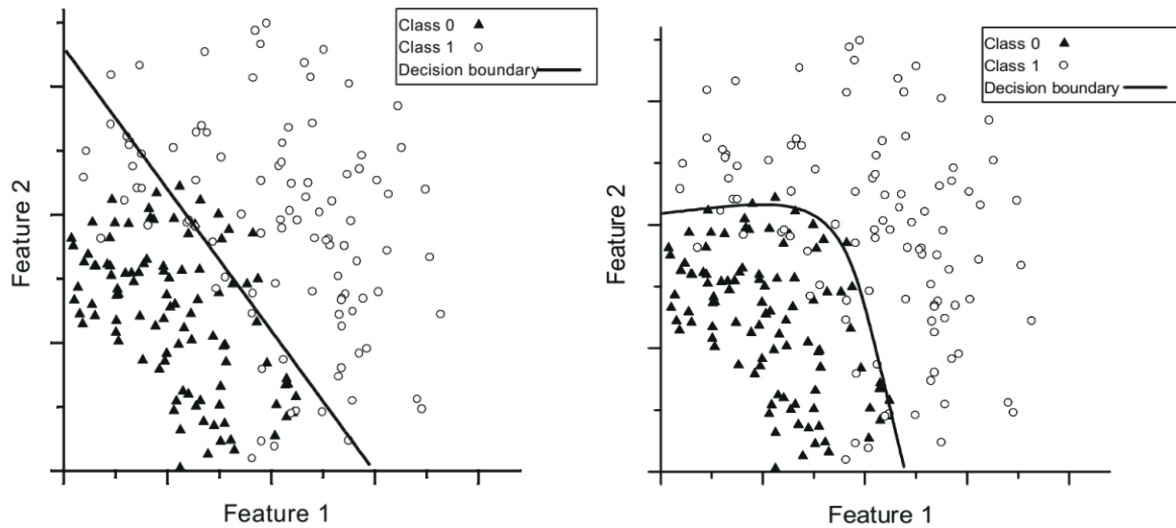


Figure 2.9: Illustration of linear (left) and quadratic (right) discriminant analysis. The linear/quadratic discriminant boundary separates the data into two classes. Linear discriminant analysis correctly separated 81.5% of data in this example, while the quadratic discriminant function improved classification by correctly separating 85.5% of data. Adapted from Kemp et al. (188).

D. *Decision Trees:* A decision tree is a hierarchical partitioning of the observation-space based on a sequence of conditions. This forms a tree structure, whereby the edges are sub-regions of the observation-space and the nodes are conditions which define the partitioning (Figure 2.9). The nodes and the edges are therefore organized in a tree structure. There are two types of nodes in a decision tree, the split nodes (internal) and the leaf nodes (terminal). Each split node is associated with a test function which is used to split the incoming data according to a variety of attributes. Each leaf corresponds to the final decision, i.e. classification label. Decision trees are not sensitive to aberrations during training (189). However, the disadvantage of the decision tree model is that it can easily overfit the training dataset and the final prediction model may be inefficient on the test dataset.

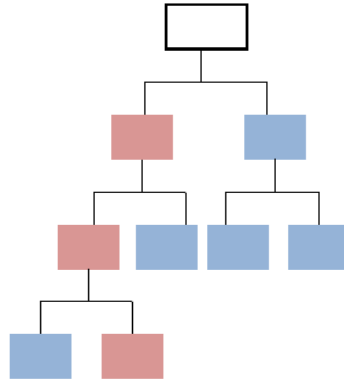


Figure 2.10: Example of a decision tree illustration with its root at the top. Based on a condition (internal node), the tree splits into branches (edges). The end of a branch that does not split anymore is the decision (leaf node).

To address the susceptibility of decision trees to overfitting, a more efficient way is by aggregating the predictions given by several of them for the same problem. This strategy can be also used to improve the accuracy of decision trees.

Bagging – The Bootstrap Aggregating (Bagging) method is a popular Ensemble method. The idea of Bagging is to simultaneously train multiple bootstrap samples from the original training dataset using the unmodified learning algorithm and then average all trained models. Bagging is a randomization method, due to the introduction of some random perturbation in the process that generates an ensemble of M bootstrap samples from a learning sample. The randomness of the model helps to reduce the redundancy of explanatory variables and the individual bootstraps are diversified (190–192). The final prediction is obtained by aggregating the predictions computed from the individual subsets. The final ensemble model has less variance than its components. Random forests can be used in combination with Bagging, and this is the methodology used in this thesis. This approach involves the creation of deep trees fitted on bootstrap samples (Figure 2.11). However, random forests sample not only over the observations, but also over features keeping a random feature subset to build the tree. Bagging considers all features by default when training each bootstrap. The addition of random forests makes bootstrap samples even less correlated and the decision making process becomes more robust to missing data.

Important parameters for the development of a Bagging Ensemble-based prediction model are the leaf size and the number of trees. The *leaf size* is tuned to determine the predictive ability of each model and corresponds to the depth of the tree. A small leaf size results to deep trees

that have adequate generalization ability, but they need more training and prediction time and the process is more computationally expensive. Large leaf size corresponds to shallow trees that need less training time and memory, but the accuracy of the model is not necessarily high in the end. The number of trees is used to control the size and the overall performance of the ensemble model (193).

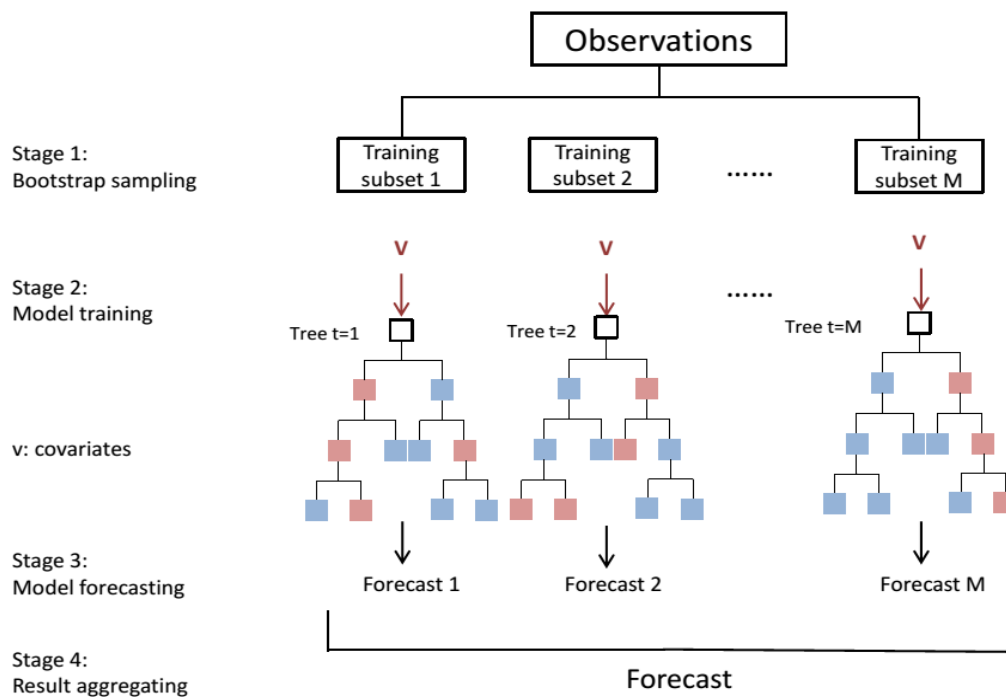


Figure 2.11: Schematic representation of the Random Forests algorithm based on the Bagging method. The process involves 4 stages. Stage 1: M subsets are sampled from the original training dataset using the bootstrap method. Stage 2: M independent decision trees are built for model training. Stage 3: Prediction is obtained from each bootstrap tree over M replications. Stage 4: The final prediction is made based on average or majority voting.

The *out-of-bag* (OOB) observations were used as a way to compute the test error of a model produced while applying the Bagging Ensemble method. OOB corresponds to the subset of observations not selected for training each decision tree. The response for an observation is made using each of the trees for which that observation was an OOB. The majority vote of predictions for a single observation indicates a single OOB prediction for that observation. The same process is repeated for each OOB observation and in the end an overall *out-of-bag classification error* is computed. This metric is an estimate of the test error, since all observations not used during training of each decision tree are used. It can be used as an alternative validation method to cross-validation, which can be computationally expensive for large datasets (194).

Boosting – Boosting is another Ensemble method used in this thesis. The main idea of boosting is also to combine the predictions of several models and obtain a stronger learner with better performance. However, unlike Bagging which aims at reducing variance, Boosting builds the models sequentially. This means that each model in the sequence is built to address the deficiencies of the previous models in the sequence. Each new model focuses on the most difficult observations to fit up to now, ending up with a strong learner at the end of the process with lower bias. Random forests can be used here too, but it is not necessary. These were not used in this thesis. Boosting is mainly preferred for base models with low variance but high bias, such as shallow decision trees with only a few depths, which are also less computationally expensive to fit.

RUSBoost – The Random Under-Sampling Boosting (RUSBoost) is another method applied in this thesis and it is prominently used for classifying imbalanced data, i.e. some class in the training dataset with K classes has many fewer observations than another. The algorithm takes N members of the class with the fewest members in the training dataset as the basic unit for sampling. Classes with more members are sampled by taking only N observations of every class (195). A weighted pseudo-loss for N observations and K classes was also computed, as a measure of the classification accuracy from any learner in an ensemble. This takes into account the confidence of making predictions.

2.9.2 Feature selection

Feature subset selection is the process of identifying and removing irrelevant and redundant features (196). It reduces the data dimensionality and enables machine learning algorithms to operate faster and more efficiently (10).

Feature selection was carried out using the Sequential Forward Selection (SFS) method developed by A.W. Whitney (197). This is a bottom-up search procedure starting from an empty feature matrix D which iteratively adds features. A subset D^* of measured EGM features was incrementally constructed. During each iteration, a proposed feature was selected from the pool of remaining features not already included in D^* . The algorithm terminates when no additional features provide a significant improvement in the ability of the classification algorithm to differentiate the classes. The classification method used during the SFS process was the same with the method next applied for classification training. Four main families of classification methods were used for SFS: kNN, Decision Trees, including Ensemble

Bagging for optimization of the performance, LDA and SVM, as previously presented. Due to this variability, a different feature subset was selected after the application of each algorithm.

Each feature subset was tested during classification training using the same algorithm which was applied during feature selection. The fact that significantly higher classification performance was achieved using the feature subset extracted through the SFS process showed that feature selection was a meaningful step during the machine learning process. Also, the relatively limited size of the training dataset used during this research, which is a reason for overfitting and subsequently false positive classifications (196), made the application of feature selection necessary. Lastly, the feature selection provided feature subsets which could be related to the experimental conditions and were biologically meaningful. These findings, in addition to the general contribution of feature selection towards the production of effective predictive models, are presented in Chapter 7-9. The EGM features used for feature selection, as well as the algorithms used for the detection and measurement of them, are presented in detail at Chapter 3.

2.9.3 Assessment of predictions by cross-validation

Cross-validation is a resampling procedure primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is to use a limited sample in order to estimate how the model, which is under development, is expected to perform in general when used to make predictions on data not used during the training of the model. Cross-validation is a popular method because it results in a less biased or less optimistic estimate of the model skill than other methods, such as a simple training/testing data split (474).

Applying a k -fold cross-validation, the training dataset is randomly divided into k equally sized subsets, as presented in (2.5),

$$l = l_1 \cup l_2 \cup \dots \cup l_k \quad (2.5)$$

The value of k was $k=10$, which is a default value and has been found through experimentation to generally result in a model skill estimate with low bias. The value of k is usually 5 or 10 and, as k gets larger, the difference in size between the training set and the resampling subsets gets smaller. As the difference decreases, the bias of the technique becomes smaller (474). A model is inferred by the learning algorithm from each subset of observations $\frac{l}{l_i} i = 1, \dots, k$ and its performance is determined on the held-out sample, l_i . The $k=10$ subsets are separately used to train 10 models and each model is

tested using a unique test set. The final performance is computed as the average performance over all these models (198).

2.9.4 Measurement of classification performance

The classification accuracy of each supervised machine learning algorithm was assessed by estimating seven statistical indices that are used in multi-class classification (199,200): average classification accuracy, average per-class classification error (error rate), average per-class agreement of the data class labels with those of a predictor (precision), sensitivity (Se), specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV). Sensitivity is the true positive rate of predictions, meaning the ability to correctly classify an observation. Specificity is the true negative rate, which measures the proportion of actual negatives that are correctly identified as such – the percentage of baseline EGMs, which are correctly classified as not being influenced by any cellular modulation. Given an individual class C_i , the assessment is defined by the true positive (tp_i), true negative (tn_i), false positive (fp_i) and false negative (fn_i) data corresponding to each class. The formula used for the calculation of each performance index, following classification training for l classes, is presented in Table 2.3.

In addition, the classification training was evaluated using a receiver operating characteristic (ROC) curve that plots the true positive rate (sensitivity) versus the false positive rate for different values of a threshold, which was used in classification problems to achieve different compromises between sensitivity and specificity (201). ROC graphs are two-dimensional with the tp rate plotted on the y-axis and the fp rate plotted on the x-axis showing different trade-offs between tp and fp . It is a technique to visualize the classification performance and compare the efficiency of training methods. The quality of a ROC curve was assessed by the area under the curve (AUC). The AUC represents the probability of a classifier ranking randomly chosen positive instances higher than randomly selected negative instances. Its values are always between 0 and 1. However, random guessing produces a diagonal curve in the ROC and therefore, AUC can never be less than 0.5, which was also the case in the analysis presented in this thesis.

Table 2.3. Equations for assessing the multi-class classification of a dataset into l classes. tp_i – true positive data of class C_i , fp_i – false positive, fn_i – false negative for class C_i , tn_i – true negative.

Statistical index	Formula
Classification accuracy	$\frac{\sum_{i=1}^m \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{m}$
Error rate	$\frac{\sum_{i=1}^m \frac{fp_i + fn_i}{tp_i + fn_i + fp_i + tn_i}}{m}$
Precision	$\frac{\sum_{i=1}^m \frac{tp_i}{tp_i + fp_i}}{m}$
Recall	$\frac{\sum_{i=1}^m \frac{tp_i}{tp_i + fn_i}}{m}$
Sensitivity (Se)	$\frac{tp_i}{tp_i + fn_i}$
Specificity (Sp)	$\frac{tn_i}{tn_i + fp_i}$
Positive Predictive Value (PPV)	$\frac{tp_i}{tp_i + fp_i}$
Negative Predictive Value (NPV)	$\frac{tn_i}{tn_i + fn_i}$

Chapter 3

Development of automated electrogram morphology analysis

3.1 Introduction

A number of factors may lead to an unsatisfactory classifier during machine learning. If the dimensionality of the data is too high, this may lead to overfitting in which the algorithm learns the training data but does not generalise to unseen data. The choice of features may also insufficiently capture the important characteristics of the data necessary to distinguish classes. Furthermore, if the dataset is imbalanced, in that the number of examples from one class is significantly different to the number from another class, then the training process may be biased over the leading class (202). Therefore, it becomes clear that signal analysis and feature extraction are essential for the development of an efficient classifier in supervised machine learning. The remainder of this chapter describes the EGM features considered for analysis, the development of algorithms for EGM morphology analysis and feature extraction and the validation of accurate detection of some of the features.

3.1.1 Digital signal processing in cardiac electrophysiology

Digital signal processing (DSP) is the group of mathematical techniques used to manipulate signals that originate as sensory data from the real world. These signals are analysed after they have been converted into a digital form (203). This conversion process starts with analogue input signals, which are continuous in time and amplitude, and a transducer is used to convert the input into analogue electrical signal, i.e. voltage. A filtering process follows, in order to limit the frequency range of analogue signals prior to sampling, which also serves to attenuate any signal distortions. The output of the analogue filter is then sampled and converted into a digital signal, which is discrete in both time and amplitude. DSP is already used in a range of medical instruments, such as ECG analysers, X-rays and medical image systems, providing useful information for precise diagnoses (204). The digital signal is then processed according to filtering or other algorithms depending on the application (204).

DSP plays an important role in the electrophysiology laboratory, because successful mapping and ablation is dependent on acquiring multiple, low-amplitude, intracardiac signals in the presence of numerous sources of signal noise and interference. The recordings also need to be displayed in a simple and clinically relevant fashion with minimal or no artefacts (205). The first problem that has to be resolved using DSP techniques is the interference of leakage current, due to the presence of various devices in the electrophysiology (EP) lab, with intracardiac signals. The result may be low signal-to-noise ratio. In order to distinguish signal from noise, good electronic design is necessary, including the placement of pre-amplifiers and amplifiers close to the signal source, the patient, the electric isolation

of patient, the power sources being routed carefully away from the patient table and cable shielding, which helps to shunt capacitively coupled interference to the ground reference of the electronic system.

Both visual review by experienced electrophysiologists and computation of EGM features are highly dependent on the signal quality. However, undesirable components such as high and low frequency noise, the ventricular far field, power line hum, and other parasitic frequencies overlay the EGM of interest (476). Filtering is useful to effectively enhance certain portions of the frequency spectrum while rejecting unwanted portions of the signal, thus increasing the signal to noise ratio (SNR). Filtering is performed at various stages of the electronic system and includes combinations of passive and active filters in conjunction with digital filters. Passive filtering is applied by resistors, capacitors and inductors, while active filtering is carried out by operational amplifiers. Digital filters use a combination of DSP hardware and software algorithms (205). Not only EGMs, but also ECGs are contaminated with baseline wander, high-frequency noise caused by electromyography and motion artefacts during collection and transmission. These components overlap the cardiac component in both spatiotemporal and frequency domains resulting to fairly difficult extraction of weak cardiac components from the corrupted ECGs. Indeed, ECG signal denoising is important for improving the quality of ECG signals (477). However, ECGs require different filtering process from intracardiac EGMs, due to differences in frequency content. Different cut-off frequencies are used depending on the components of the signal derived from various parts of the cardiac tissue, such as His bundle, pulmonary veins and Purkinje-fibre area. In electro-anatomic mapping systems, such as Carto 3, the filter selection is carried out both in the signal acquisition system and the mapping system. A range of 16-500Hz is suggested by the manufacturer of this system, but this can be adjusted in the Carto 3 system depending on a set of EGMs (205). The SNR of EGM and ECG recordings cannot be compared, as it heavily depends on the denoising method and the scholastic investigation of EGM noise quantification methods has been quite recent (476). Despite the current ability to remove far field noise occurring equally in both unipolar leads and overcoming the loss of spatial accuracy and morphological features, as well as the dependence on the direction of excitation propagation, SNR is one of the reasons why unipolar EGMs are rarely used. The lack of efficient noise removal methods, specifically adopted to unipolar EGMs, is a limiting factor. Even the removal of far field noise is demanding because the components do not simply cancel out, as opposed to bipolar EGMs where this process is simpler (476).

3.1.2 Electrogram morphology analysis

The analysis of digitally sampled intracardiac EGMs is commonly used for the automated characterisation and identification of activation patterns of cardiac arrhythmias. There are a number of

algorithms based on time-domain and spectral analyses which have been implemented in clinical electro-anatomical mapping systems for automated signal processing. The results of these analyses are used to guide mapping and ablation therapy.

The interval confidence level (ICL) is a time-domain method suggested by Nademanee *et al.* (106) for quantification of CFAEs recorded during atrial fibrillation. The algorithm used in this method counts the local maxima and local minima based on a fixed noise threshold. The blanking time between two peaks is pre-defined by a duration interval. If the time duration between two adjacent peaks is smaller or larger than the set blanking interval, then both peaks are considered invalid and excluded from the analysis (206). This method is embedded in the CARTO mapping system (Abbott, IL, USA). Another similar algorithm uses the mean of R-peak intervals as an alternative to the previous method. Peaks that are above the pre-defined fixed noise threshold are considered to be valid and the duration between two consecutive peaks has to be greater than the mean fractionation interval (206). This algorithm is implemented in the EnSite NavX mapping system (Abbott). However, while some studies have shown the efficacy of these algorithms for identifying critical sites of ablation (37,108,207), others have shown more moderate success rates (107,208).

Frequency-domain methods have also been considered. The Fourier transform decomposes a signal into a sum of sinusoidal waves of different frequencies. From this, the dominant frequency (DF) can be computed, defined as the frequency component with greatest energy in the signal. Many studies have used the DF as a measure to identify fast-firing locations in the atria which are considered drivers for catheter ablation (118,209,210). Regularity index (RI) is another commonly used frequency-domain feature. The RI is calculated as the ratio of the power of the DF, and its surrounding 0.75Hz frequency band, to the power of the other harmonics in a bandpass frequency range. It is usually used as a reliability parameter for DF (118,209,210). However, the DF and RI algorithms have showed only modest efficacy in the clinic, similar to time-domain methods (211).

Due to weaknesses of the methods previously described and insufficient characterisation of EGM morphology, it is not possible to efficiently correlate EGMs to any abnormalities of the electro-architecture. Thus, there is a need for extended study of EGM morphology and the analysis of new features (153). To date, feature extraction from CFAEs is carried out mainly based on either time-domain or non-linear features. However, a modern approach is the combination of these features, which is considered to achieve a better ablation performance (153,212,213). There are published methods for automated CFAE analysis, based on time-domain, wavelet- and non-linear-based descriptors and the results are used for supervised classification of CFAEs based on different CFAE transition stages (213).

This chapter presents the algorithms used in this thesis for processing the EGM recordings obtained *in vivo*, *in vitro* and *ex vivo*, in order to extract features that characterise EGMs. Some of these algorithms had already been available through MATLAB, but the ones that I developed are stated. The feature vectors extracted in the end of the signal processing analysis are used afterwards for creating datasets. These are necessary for the application of supervised machine learning techniques and the development of predictive models, as described in Chapter 5. Chapter 3 also presents the validation of time-domain analysis algorithms that I developed.

3.2 Methods

3.2.1 Electrogram features: An overview

Automated analysis of the EGM signal was performed using custom-written code (MATLAB R2016b). The code detected and quantified 37 time-domain, frequency-domain and time-frequency features, including features based on information theory and non-linear dynamics. These were features previously published to be extracted from clinical EGM and ECG recordings. When analysing the *in vitro* and *ex vivo* data, only the 36 central electrodes of the 8 x 8 matrix were analysed to exclude the stimulation electrodes along the outer lines of the matrix. These electrodes either recorded no discernible EGM signal or contained high levels of noise due to stimulation. Even though normal EGM signal might be obtained from the boundary electrodes on the opposite side of the array to the stimulus line, the boundary electrodes on the four sides of the MEA were always excluded in order to consistently analyse the data.

3.2.2 Stimulus artefact removal

During electrical stimulation of cardiac cells – under the *in vitro*, *ex vivo* or *in vivo* experimental conditions – the recorded EGM is contaminated by a stimulus artefact. The degree of contamination depends on the stimulation technique, the recording conditions and the experimental model (214). The magnitude of the stimulus artefact dominates the signal emitted from the cells and in the worst case exceeds the dynamic range of the EGM resulting either in a flat line recording or an EGM without clear onset.

Since the morphology of the stimulus artefact was variable between recordings, as shown in Figure 3.1, an adaptable method of removal was needed. The sample of largest magnitude within the 10sec recording – which is the total duration of each electrical recording - was found and all maxima greater than 60% of this value were identified. This threshold was selected, in order to confirm that all R-peaks and stimulus artefacts are included in analysis. By examining the first two of these peaks, four cases were considered. If the first peak was lower and $>0.5\text{ms}$ apart from the second peak, then the stimulus artefact was not filtered or removed (Figure 3.1C-D). Thus, the EGM was kept unmodified. However, in case the first peak was $<0.5\text{ms}$ apart from the second peak – the voltage was not considered in that case – then that was part of the same EGM, any stimulus artefact did not reach the threshold and was not detected, and the locations of both peaks were used for computing the EGM fractionation at a later stage of the feature extraction. The 0.5ms was used, because it was found that this was the longest time interval between two detected peaks in a double-peaked EGM.

When the first peak was higher and $>0.5\text{ms}$ apart from the second peak, which was the most frequent case, that meant that there was a stimulus artefact obstructing the EGM and the mean baseline voltage, V_b , was computed as the average voltage of the first 40ms (1,000 samples). The stimulus artefact region, R_{SA} , was then used, which was empirically determined as 100-200 samples on either side of the stimulus artefact peak depending on the stimulus artefact amplitude. The artefact region R_{SA} was then replaced by the mean baseline voltage, V_b , as illustrated in Figure 3.1B. In case the first peak was higher than the second peak and they were separated by less than 0.5ms , these were considered part of a fractionated EGM and no further action was taken by the algorithm. The processed signal was then used as input for the following steps of the feature extraction process.

3.2.3 Time-domain feature extraction

Features based on time-domain analysis of EGM morphology were derived from the morphology of the QRS complex and T wave. The QRS complex is defined here as the interval between the EGM onset and the EGM offset. An EGM was detected as a single QRS-T part of the signal, as opposed to the whole recording. The algorithms presented in this chapter were developed by me, unless otherwise stated. R-peak was detected as the highest voltage point of each EGM, which was also greater than 60% of the sample of largest magnitude – as presented in section 3.2.2. S-peak was defined as the point of lowest voltage within 4ms (100 samples) following each R-peak. The *EGM amplitude* was defined as the voltage difference between R- and S- peaks. The amplitude was measured for each EGM

throughout a 10sec recording, as pacing resulted in multiple EGMs per recording. Therefore, the *average EGM amplitude* was also calculated. However, the *amplitude of the first EGM* also formed a

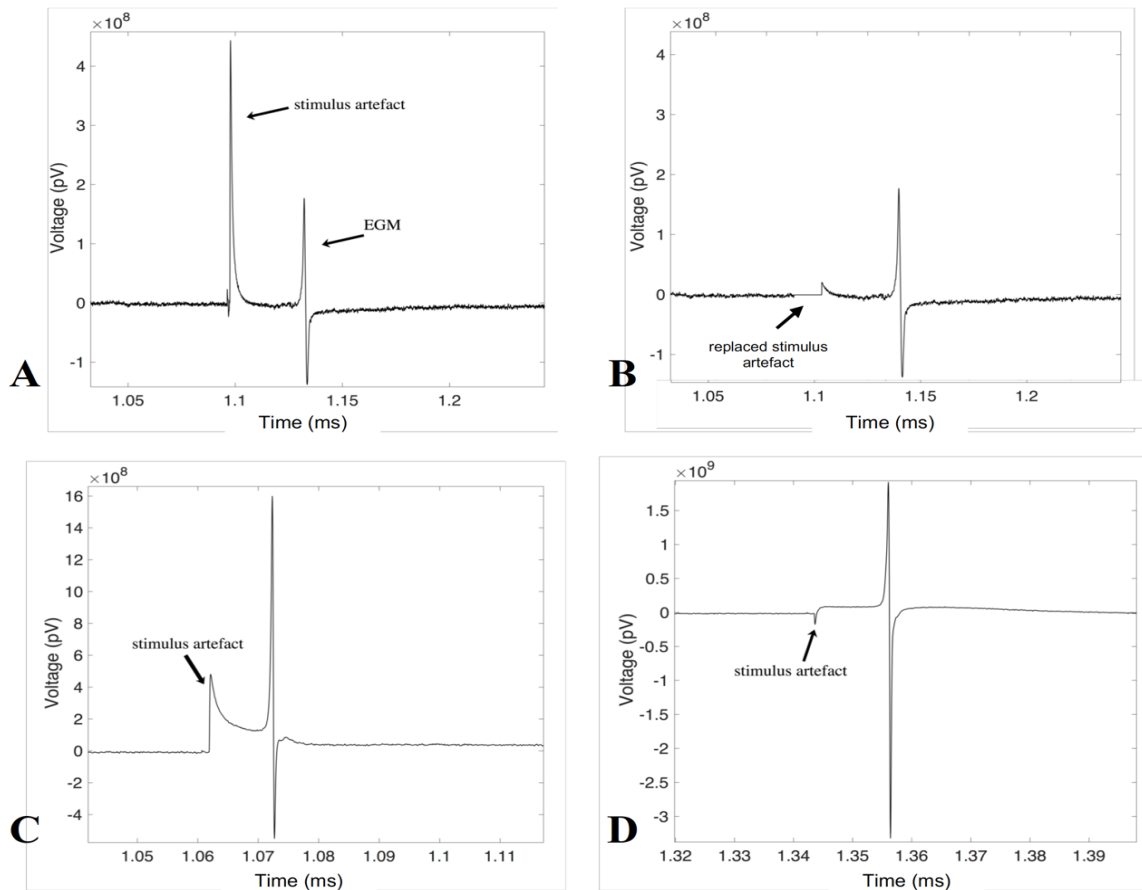


Figure 3.1: Representative stimulus artefacts in EGM recordings obtained from NRVM monolayers. (A) Example of a stimulus artefact with a peak higher than the R-peak of the subsequent EGM. This was the most common case among recordings obtained both from cell cultures and tissue slices. (B) The EGM recording presented in (A) following stimulus artefact removal and its replacement with the mean baseline voltage. (C) Example of a stimulus artefact with a peak of lower magnitude than the R-peak of the EGM. (D) Example of a stimulus artefact with no peak.

feature. Since voltage alternans is known to be related to vulnerability to arrhythmias (215,216), the *variance of EGM amplitude* was also measured. *RS interval* was computed as the time duration between R- and S-peaks. *Fractionation index* was defined as the number of positive deflections above 60% of the sample of largest magnitude of each EGM within a 10sec recording. Some of these features were detected on the EGMs presented in Fig. 3.4.

The RR interval was calculated as the time difference between adjacent R-peaks. If X is a vector of length m , where $X(1)$, $X(2)$ etc. are the R-peaks, then A is a vector of length $(m-1)$. The elements of A are the differences between adjacent elements of X :

$$A = [X(2) - X(1) \quad X(3) - X(2) \quad \dots \quad X(m) - X(m-1)] \quad (3.1)$$

The *RR interval* was defined as the average value of the elements of A :

$$RR \text{ interval} = \frac{\sum(A)}{(m-1)} \quad (3.2)$$

The EGM onset was defined as the first recorded sample at 10% above V_b and before the first R-peak, in order to eliminate noise observed close to the baseline signal. The EGM offset was detected as the first sample following the S-peak and passing above the 10% threshold below baseline signal. The EGM offset detection occurred within a time window that included the S-peak location and 8ms (200 samples) following that. *EGM duration* was measured as the time difference between the EGM onset and offset. *QS interval* was the time difference between the EGM onset and the S-peak.

If y is the signal recording over the first second, the local minimum of the first derivative y' on the interval [R-peak location, S-peak location] was measured, in order to extract the time of the $(-dV/dt)_{max}$ between R- and S-peaks. Based on the location of $(-dV/dt)_{max}$, the *width of R- and S-peaks* were calculated as the time duration between the EGM onset and the $(-dV/dt)_{max}$ and the time duration between the $(-dV/dt)_{max}$ and the EGM offset respectively.

There are features extracted during digital signal analysis which may depend on one another. The construction of new features from the basic feature set is a technique best known as *feature construction* (217). The same strategy was followed in this thesis. Based on the amplitude and RS interval, the average gradient of the signal between R- and S-peaks was calculated as: *average gradient = first EGM amplitude / RS interval*. In a similar way to this, the gradient between the S-peak and the EGM offset and the gradient between the EGM onset and the R-peak were calculated. The amplitude of each peak was also measured, as the voltage change between EGM onset and R-peak or the voltage change between S-peak and EGM offset. Based on these measurements the ratio $(R\text{-peak amplitude})/(S\text{-peak amplitude})$ was calculated. The previously described features were also used as the basis for the calculation of:

$$(R\text{-peak width}) / (S\text{-peak width}) \text{ and } (R\text{-peak width}) / (EGM \text{ duration}) \quad (3.3)$$

The manual detection of T-waves was based on the detection of the positive deflection following the QRS complex in the EGM signal. The automated detection of T-waves was based in the idea presented by M. Nair (473) which involved a number of filtering steps, in order to make T-waves more prominent on the EGM signal. However, this was adopted to the data obtained during this research and a quite different filtering process was followed. Firstly, a 2nd order Chebyshev type I filter (218) designed for the bandwidth of 0.5-10 Hz was implemented. In addition, a 19th order Savitzky-Golay filter with a frame length of 20 was applied (219). These conditions were optimised in order to filter the signal enough without causing signal distortions, i.e. amplitude reduction. The above filters acted to eliminate high frequencies from the signal. The QRS complex was then replaced by a mean baseline signal. Removing the QRS complex made the T-wave the dominant feature in the processed signal. Finally, the peak of T-waves could be detected using the approaches described earlier for the R-peak detection. The location of a T-wave was used for measuring three time-domain features. *QT interval* was calculated as the time duration between the EGM onset and the location of the T-wave peak. *Field Potential Duration (FPD)*, which is equivalent to the APD at 90% of repolarisation (40), was equal to the time duration between the location of S-peak and the T-wave peak. The *T-wave amplitude* was measured as the voltage change between the baseline signal and the T-wave peak.

Table 3.1: List of time-domain features measured in this thesis.

Time-domain features		
1. amplitude of first EGM	7. gradient between EGM onset – R-peak	13. R-peak amplitude/S-peak amplitude
2. average EGM amplitude	8. gradient between S-peak – EGM offset	14. EGM duration
3. variance of EGM amplitude	9. RR interval	15. fractionation index
4. average RS interval	10. R-/S-peak width	16. QT interval
5. QS interval	11. R-width/S-width ratio	17. Field Potential Duration
6. signal gradient between R-/S-peaks	12. R-width/EGM duration ratio	18. T-wave amplitude

3.2.4 Frequency-domain EGM features

Dominant frequency (DF) is the main frequency-domain feature and it is based on the application of FFT (220). DF is defined as the frequency with the maximum power in the Fourier spectrum of a signal.

The DF calculation was based on the method presented by Telgarsky (115). Based on this, the signal was first pre-processed. The Chebyshev type I bandpass filter, followed by rectification, was applied (Figure 3.2B-C). The resulting filtered waveform was filtered once more using a 20Hz low-pass 2nd order Butterworth filter (Figure 3.2D). These pre-processing steps, which were optimised for the EGMs collected at 25 kHz sampling frequency, enhanced the periodicity of the recordings obtained at 1-5Hz pacing for the purposes of this thesis. Moreover, this algorithm is proved to diminish the effects of changing EGM morphology or amplitude due to electro-architectural factors (118). The signal was then de-trended by subtracting the mean signal from the data. The FFT was then applied to obtain the power spectrum of the signal. DF was computed as the frequency of maximum energy in the power spectrum.

Power spectral density (PSD) is a type of frequency-domain analysis in which a signal is subjected to a probabilistic spectrum of harmonic loading to obtain probabilistic distributions of frequency components for dynamic response (221). PSD is deterministic and independent of time. The most common way of generating a power spectrum, in order to compute PSD, is using the discrete Fourier transform (221). If we take a sample of the function $f(t)$ at equal intervals, the FFT can be applied to compute its discrete Fourier transform according to (3.4),

$$f_k = \sum_{j=0}^{N-1} f_j e^{2\pi i j k / N} \quad k = 0, \dots, N-1 \quad (3.4)$$

where N is the total size of samples ranging over integer values.

The periodogram, which is a graphical data analysis technique for examining frequency-domain models, was used as an estimate of the spectral density of a signal to identify the dominant periods (frequencies) of a time series (222). The periodogram is the Fourier transform of the biased estimate of the autocorrelation sequence. For a signal x_n sampled at f samples per unit time, the periodogram is defined as shown by Eq. 3.5:

$$P(f) = \frac{\Delta t}{N} \left| \sum_{n=0}^{N-1} x_n e^{-j2\pi f n} \right|^2, \quad -\frac{1}{2\Delta t} < f < \frac{1}{2\Delta t} \quad (3.5)$$

where Δt is the sampling interval, N is the number of observations in the signal, Δ is $(N+1)/2$ for N odd and $(N+2)/2$ for N even. For a one-sided periodogram, the values at all frequencies except 0 and the Nyquist, $1/2\Delta t$, are multiplied by 2 so that the total power is conserved.

The periodogram then consists of a vertical axis which is the spectrum estimate at the given frequency, and a horizontal axis representing the Fourier frequencies $(1/N, 2/N, \dots, (N/2)/N)$ where N is the number of observations in the signal.

Data windowing is important when calculating the periodogram estimate. The window function for the periodogram estimate is:

$$W(s) = \frac{1}{N^2} \left[\frac{\sin(\pi s)}{\sin(\frac{\pi s}{N})} \right]^2 \quad (3.6)$$

This window function is a function of s , the frequency offset in bins. $W(s)$ has oscillatory lobes, but apart from these, it falls off only about as $W(s) \approx (\pi s)^{-2}$. This is not a rapid fall-off and it results in significant leakage from one frequency to another in the periodogram estimate. The solution to the problem of leakage is data windowing. Therefore, according to P.D. Welch (223), when Welch's data windowing method is implemented, Eq. 3.6 is modified. For each segment of N sampled points for periodogram spectral estimation, an infinite run of sampled data is multiplied by a window function in time, one that is zero except during the total sampling time $N\Delta$ and is unity during that time. That means that the data are windowed by a square window function. According to the convolution theorem, the Fourier transform of the product of the data with this square window function is equal to the convolution of the data's Fourier transform with the window's Fourier transform. Thus, Eq. 3.6 is modified as

$$W(s) = \frac{1}{N^2} \left| \sum_{k=0}^{N-1} e^{2\pi i s k / N} \right|^2 \quad (3.7)$$

The square window function turns on and off rapidly due to large s values. In order to solve this, the input data $c_j, j = 0, \dots, N-1$ are multiplied by a window function w_j , which changes more gradually from zero to a maximum and then back to zero as j ranges from 0 to N .

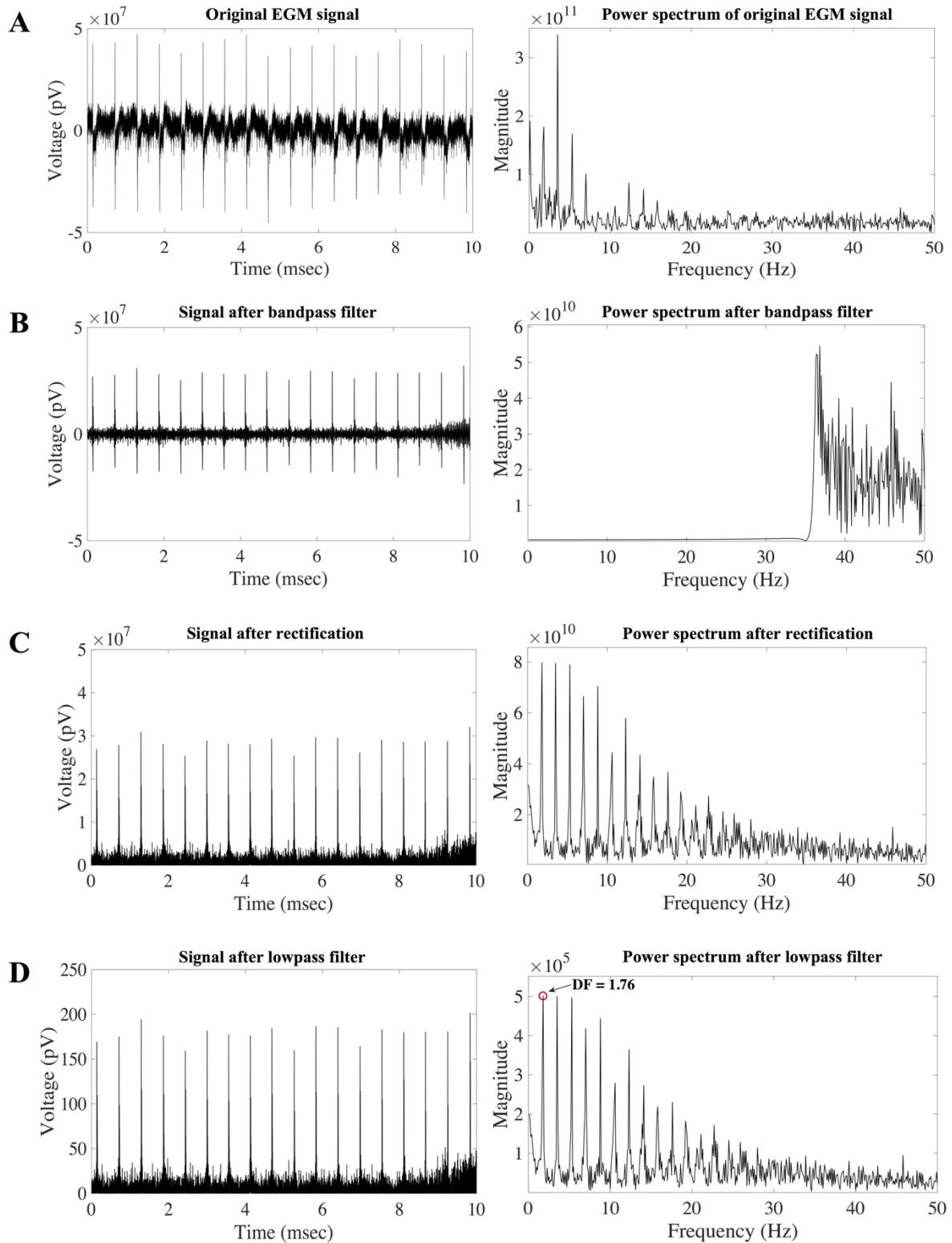


Figure 3.2: Representative resulting signals and corresponding power spectra after each step of analysis during the filtering protocol, in order to compute the dominant frequency. (A) The signal and power spectrum before the application of filters. (B) The signal and power spectrum after bandpass filtering at 40-250Hz. (C) Signal after rectification. (D) The signal and power spectrum after the application of a low-pass 2nd order Butterworth filter with a 20Hz cut-off frequency. The power spectrum shows that the DF of the signal is 1.76. DF = dominant frequency.

For the PSD analysis performed in this thesis, the Hamming window was used,

$$w_j = \frac{1}{2} \left[1 - \cos \left(\frac{2\pi j}{N} \right) \right] \quad (3.8)$$

The Hamming window does not go exactly to zero at the ends.

All the above equations for the calculation of the periodogram during PSD were implemented in MATLAB functions that were used for this type of analysis. The periodogram was subsequently used to extract features, which include the *mean and median frequency of periodogram estimates*, the *maximum estimate* and the *standard deviation of estimates*. Moreover, the *skewness and kurtosis of periodogram estimates* were measured. The skewness, which is a measure of the degree of asymmetry of a given distribution around the mean (224), was defined as

$$s = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^3} \quad (3.9)$$

Kurtosis, which characterizes the relative distortion of a signal compared to the normal distribution (224), was defined as

$$k = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} \quad (3.10)$$

Both in Eq.3.9 and Eq.3.10, x is the PSD estimates, n the number of estimates and i is the value for bias correction.

Table 3.2: List of frequency-domain features measured by using the algorithms presented in section 3.2.4.

Frequency-domain features	
1. Dominant frequency	5. Skewness of PSD estimates
2. Mean of PSD estimates	6. Kurtosis of PSD estimates
3. Median of PSD estimates	7. Standard deviation of PSD estimates
4. Maximum PSD estimate	

3.2.5 Time-frequency EGM features

Time-frequency analysis offers simultaneous interpretation of the signal both in time and frequency allowing local, transient or intermittent components to be elucidated (225). The Continuous Wavelet Transform (CWT) of the EGM signal, $f(t)$, is useful when features in the signal are localised in time allowing for sophisticated filtering to obtain only parts of signal that are of interest. CWT was applied in the research presented in this thesis and it was based on Morlet wavelet (226), which is given by:

$$f(t) = \frac{1}{\sqrt{\pi\beta}} e^{2i\pi f_c t} e^{-\frac{t^2}{\beta}} \quad (3.14)$$

where β is a bandwidth parameter and f_c is the wavelet centre frequency given by the positive solution to,

$$f_c = \sigma \frac{1}{1 - e^{-\sigma f_c}} \quad (3.15)$$

Centre frequency can be solved by a fixed-point iteration starting at $f_c = \sigma$. The parameter σ allows trade between time and frequency resolutions. An example of the result of EGM

The CWT of a signal $f(t)$, for a scale parameter $a > 0$ and parameter $b \in \mathcal{R}$, is defined by Eq. 3.16, (226):

$$C(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\alpha} \psi * \left(\frac{t-b}{b} \right) dt \quad (3.16)$$

While $\psi(t)$ is a continuous function in both time and frequency domains and it is called the mother wavelet, the $\psi^*(t)$, which is present in (3.16), represents the complex conjugation. The *cwt* coefficients are obtained by varying the values of scale parameter a and position parameter b .

Wavelet modulus maxima are used for location characterising singularities in wavelet space (226). Morlet wavelet could provide us with the ability to extract continuous modulus maxima across scales in wavelet space. *Modulus maxima* are defined as any point b_0 in wavelet space $C(a, b)$ (Eq. 3.16) with local extremum $b=b_0$. When b varies, this implies that

$$\partial T(\alpha_0, b_0) \partial b = 0 \quad (3.17)$$

In addition, the conversion from wavelet scale to wavelet frequency was applied to Morlet wavelet and this was defined by the following equation:

$$f = \frac{f_c}{a * \Delta t} \quad (3.18)$$

With a being the wavelet scale, f_c is the centre frequency of the wavelet in Hz as defined by Eq. 3.15, f is the pseudo-frequency corresponding to the scale a in Hz and Δt is the sampling period, $\Delta t = 1/f_s$ (f_s , sampling frequency).

An energy density analysis over frequency was also carried out. The *energy of signal at each scale* was calculated as,

$$E_\alpha = \sum_{\alpha=1} |(Ca, b)|^2 \quad (3.19)$$

and the energy was then expressed as a percentage,

$$\%E = \frac{E_\alpha}{\sum_\alpha E_\alpha} * 100 \quad (3.20)$$

The *variance of signal energy* across all scales was also measured by,

$$\sigma^2 = \frac{\sum (E_\alpha - \mu)^2}{N} \quad (3.21)$$

where E_α on each scale a is defined by Eq. 3.19, μ is the mean energy across all scales and N is the number of scales a in the distribution.

Scale value, a , determines the degree to which the wavelet is compressed or stretched. The high scale CWT coefficients represent the coarse-scale features in the signal. On the contrary, the low scale CWT coefficients represent the fine-scale features in the signal. Therefore, the modulus maximum corresponds to the minimum scale a . The energy correlating to that scale was used for Eq. 3.20. *Scales corresponding to the minimum and maximum energy of the signal*, as these are measured by Eq. 3.19, were also extracted. Moreover, these scales were used for the measurement of subsequent pseudo-frequencies, f , using Eq. 3.18. An example of the visual outcome applying CWT on EGMs is presented in Fig. 3.3. This presents an EGM obtained from a NRVM monolayer which was used for the extraction of a 2D scalogram which is the visual representation of a wavelet transform. The scalogram consists of the absolute value of the CWT coefficients of a signal enabling multi-resolution analysis. Since CWT allows the spatiotemporal analysis of a signal, Fig. 3.3 shows the temporal colocalization of the EGM and the subsequent magnitude of CWT coefficients.

Table 3.3: List of time-frequency features computed in this thesis by using the algorithms presented in section 3.2.5.

Time-frequency features	
1. Maximum modulus	4. Pseudo-frequency corresponding to minimum energy
2. Variance of signal energy	5. Pseudo-frequency corresponding to maximum energy
3. Scale corresponding to minimum/maximum energy	6. Percent of energy

3.2.6 Features based on information theory

Entropy is a statistical tool that quantifies a time series in terms of the information size. In other words, entropy can analyse the amount of randomness in the signal (227). The Shannon entropy applied in this thesis was based on the method suggested by C.E. Shannon (126). A discrete probability space of a dynamic system, which was the EGM signal in this case, was used for the measurement of Shannon entropy and this was described as $A = (A | P)$. The total number of elements in A is N and the characteristic elements was defined as $A = \{a_1, \dots, a_k\}$. The relevant probability set was $P = \{p_1, \dots, p_k\}$ with $1 < k < N$. Each element a_i had probability $p_i = N_i/N$ ($0 < p_i < 1$), where N_i was the total number of elements a_i in A . Therefore, the entropy of space A was defined by Eq. 3.22:

$$\sum_{i=1}^k p_i = 1$$

$$\mathcal{H}(A) = - \sum_{i=1}^k p_i \log_2 p_i \quad (3.22)$$

Shannon entropy (ShanEn) though was the average information content of A . It measured the degree of uncertainty that existed in the dynamic system as:

$$\mathcal{H}_{ShanEn}(A) = - \sum_{i=1}^k p_i^2 (\log_2 p_i)^2 \quad (3.23)$$

Logarithmic energy (LogEn) entropy was also used during EGM morphology analysis. LogEn of A was defined by the following equation (228):

$$\mathcal{H}_{LogEn}(A) = \sum_{i=1}^k (\log_2 p_i)^2 \quad (3.24)$$

The dynamic system A in this study consisted of 250,000 events recorded during the 10sec stimulation.

Autocorrelation function (ACF) has been used for studying the AF and VF signal for many decades (229,230). The formula for computing the normalised ACF in this thesis was given as

$$acf = \frac{\sum_{i=k}^{N-1} x_i x_{i-k}}{\sqrt{\sum_{j=k}^{N-1} x_j^2} \sqrt{\sum_{j=0}^{N-k-1} x_j^2}} \quad (3.25)$$

with x being the signal to be analysed and N the number of signal segments. The direct calculation of the ACF required approximately $N-k$ multiplications at lag k , with $0 < k < x$, in addition to approximately N multiplications at the first lag. The above equation (Eq. 3.25) corresponded to the calculation of correlation coefficient between a signal and a k -lagged version of itself (231). The *standard deviation of the acf(k)* was then extracted, which was an EGM feature used in further analyses.

Mutual information was another feature based on information theory that was applied for EGM feature extraction. Mutual information is defined as the measurement of the general dependence of two variables and it provides a better criterion for the choice of time delay, comparing to the ACF that only measures linear dependence and the time delay is chosen arbitrarily (232). It is suggested that for chaotic oscillations, i.e. like the EGM signal collected during fibrillation events, the mutual information gives a more reliable value of time delay (233). Mutual information was computed between 1sec sub-segments of the EGM signal using the appropriate MATLAB function and based on the method developed by Fraser and Swinney (232). In the end, the mean mutual information was measured and used for characterising an EGM.

Table 3.4: List of features which are based on information theory and which are computed using the algorithms presented in section 3.2.6.

Features based on information theory	
1. Standard deviation of autocorrelation function	3. Shannon Entropy
2. Mutual information	4. Log Energy Entropy

3.2.7 Computational time for morphology analysis

The computational time was variable and dependent on signal complexity. Based on the analysis of 150 randomly selected electrodes, the computational time for feature extraction from a single electrode was 18.84 ± 5.2 sec for cell cultures and 28.18 ± 0.9 sec for tissue slices. However, the EGM analysis was carried out automatically on the whole electrode matrix analysing 36 electrodes in total, as previously explained, and therefore the computational time in that case was 10.38 ± 0.5 min for cell monolayers and 17.15 ± 0.5 min for tissue slices. The duration of feature extraction in clinical EGMs was 17.4 ± 3.5 sec for individual traces (mean \pm S.D. time from 100 EGMs). The duration of analysis was highly variable among datasets obtained from the patients, due to different amounts of available traces correlating to <0 S.D or >2 S.D. of blood pool mean regions.

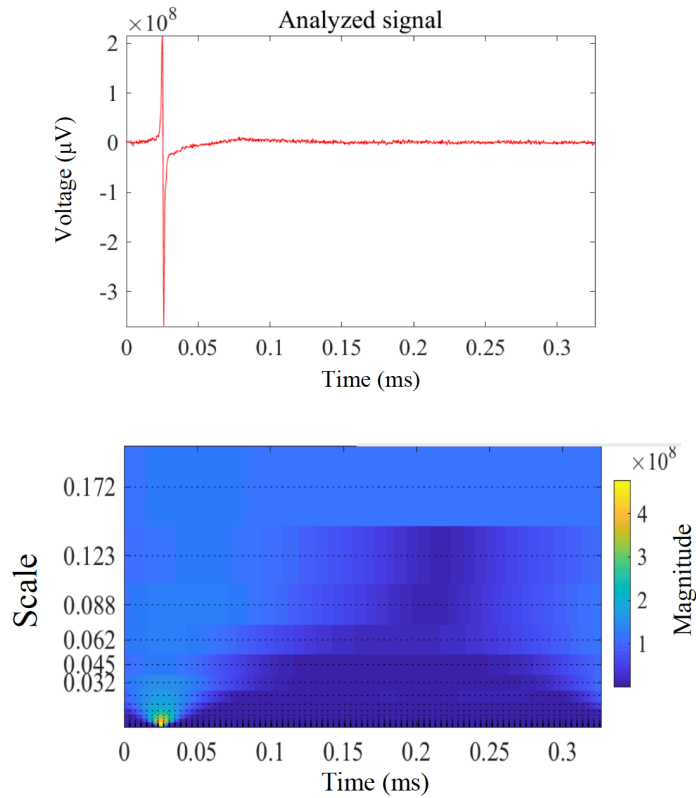


Figure 3.3: Example of an EGM obtained from a NRVM monolayer (*top*) with the subsequent decomposition of the CWT time series using the Morlet wavelet. The processed CWT time series is represented as a scalogram (*bottom*), which shows the percentage of energy for each wavelet coefficient over time. The colour codes the density of the signal component in the corresponding frequency with brighter colours (towards higher magnitude) representing a higher density.

3.3 Results

3.3.1 Quantification of electrogram features

The elimination of the stimulus artefact did not distort the overall signal and it did not affect EGM morphology, as there was no interaction of the stimulus artefact removal algorithm with the signal within the EGM onset/EGM offset interval. As presented in Figure 3.4, the morphology of the signal collected in each case showed a different level of complexity. However, all main morphological characteristics, which are the basis for the successful analysis of all time-domain features, could be detected. These characteristics are the R- and S-peaks, the EGM onset/offset and the location of $(-dV/dt)_{\max}$.

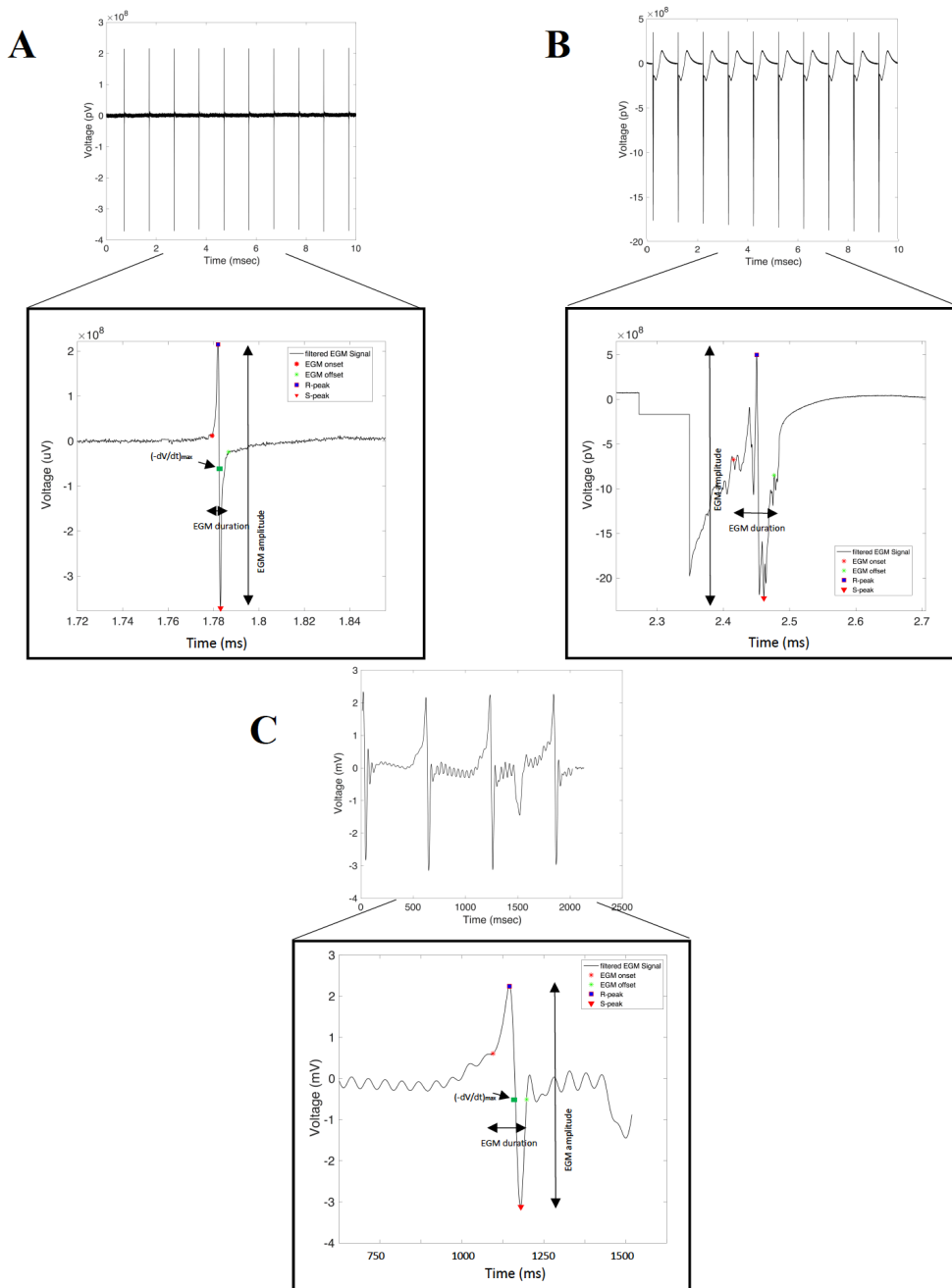


Figure 3.4: Examples of unipolar EGM recordings obtained from different settings. (A) A 10-sec unipolar extracellular signal collected from NRVM monolayer and a single EGM, where the main time-domain features are detected. (B) EGM recording obtained from a tissue slice derived from heart-failure transplant and a single EGM from the same recording, which has a more complex morphology than the EGM presented in (A). (C) EGM recording obtained from the posterior LA of a persistent AF patient and an EGM from the same recording with the basic time-domain features detected on it.

3.3.2 Verification of automated feature analysis

A group of automatically calculated time-domain features were compared against manual calculation of the same features that I carried out, in order to assess their accuracy. These features included the amplitude of the first EGM, EGM duration, fractionation index, the width of R- and S-peaks. This is an important step, since the majority of the algorithms for measuring time-domain features were not implemented into MATLAB functions and I had to develop them. Also, this analysis would show whether morphological characteristics, such as the EGM onset/offset and the local maxima/minima, were correctly detected. For this analysis the number of EGMs obtained before and after each modulation was the same (n=50 EGMs) and this size of data was maintained throughout analysis for all modulations. The same EGMs were used for manual analysis, as well as automated analysis, in order to be consistent.

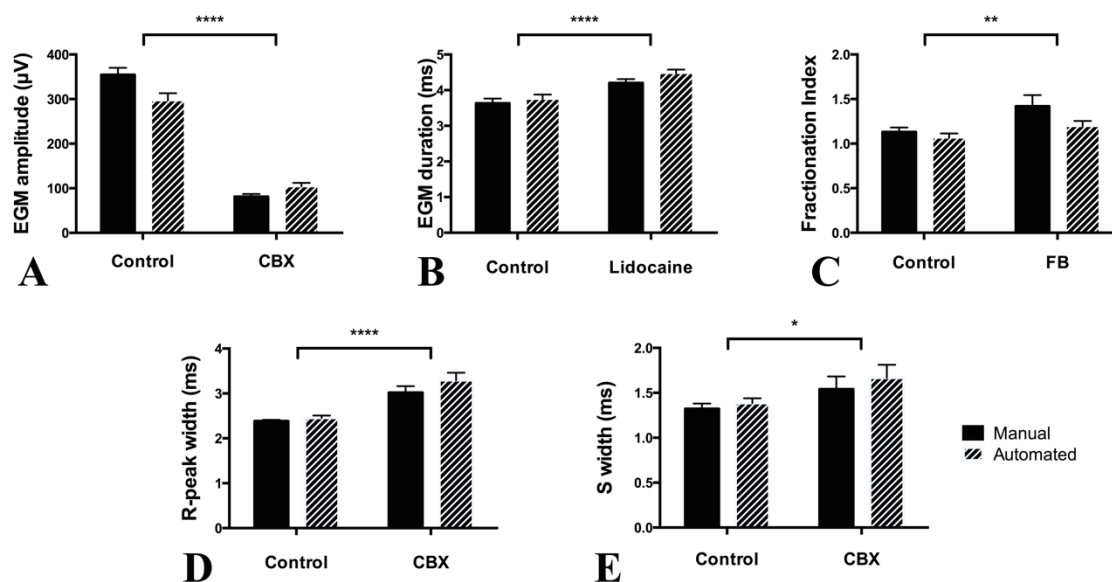


Figure 3.5: Time-domain features, including EGM amplitude, EGM duration, fractionation index, R-/S-peak width, were measured both manually and using the automated feature extraction method. Features changed only due to the administration of pharmacological agents, i.e. 40µM CBX or 20µM lidocaine, or due to the increased amounts of fibroblasts (FB) in the cell culture, but the analysis method gave the same results in each case. The number of EGMs that were analysed before and after each modulation were equal and the same throughout analysis (n = 50 EGMs in each group). All bar charts represent mean \pm S.D.; * $p \leq 0.5$, ** $p \leq 0.01$, **** $p \leq 0.0001$.

As it can be seen on Figure 3.5, each functional and structural modifier was assessed using a specific EGM feature, despite the fact that each intervention may have changed multiple other features. The two-way ANOVA test showed that the EGM amplitude was the same between the manual and the automated analysis both before and after the administration of 40µM CBX at NRVM monolayers. The

same occurred for EGM duration, R-peak width and S-peak width before and after the administration of 20 μ M lidocaine in +20%FB cell cultures and for fractionation index measured in MO and +20%FB cell cultures. The size of each feature as measured either with the manual analysis or the automated analysis before and after the modification is showed at Table 3.5. These results show that there was no change between the manual and automated detection, meaning that the EGM morphological characteristics were correctly detected, and the algorithms functioned as expected.

Table 3.5: Manual and automated measurement of representative time-domain features. The values of each feature are presented as mean \pm S.D (n=50 for both control/modification data). The percentage of difference between each pair of measurements is also presented.

	Control			Modification		
	Manual analysis	Automated analysis	% difference	Manual analysis	Automated analysis	% difference
EGM amplitude (μ V)	354.7 \pm 96	326.9 \pm 106	-10.6 – -16.2	81 \pm 39.8	94.2 \pm 53.4	-0.9 – +12.2
EGM duration (ms)	3.63 \pm 0.8	3.75 \pm 0.8	2.7 – 4.2	4.2 \pm 0.7	4.47 \pm 0.6	3.5 – 10.6
Fractionation index	1.13 \pm 0.3	1.06 \pm 0.3	-4.9 – -7	1.42 \pm 0.8	1.19 \pm 0.4	-8.4 – +17.4
R-peak width (ms)	2.38 \pm 0.2	2.45 \pm 0.4	-6.4 – +10.5	3.02 \pm 1	3.3 \pm 1.2	4 – 11.9
S-peak width (ms)	1.32 \pm 0.4	1.38 \pm 0.4	3.5 – 6.5	1.54 \pm 1	1.67 \pm 1	5 – 14

3.3.3 The T-wave detection

As described in section 3.2.3, the approach for the detection of T-wave peak involved a series of filters, in order to make the T-wave more prominent and its detection easier. Figure 3.6 presents examples of the application of that method on EGMs obtained from NRVM cultures and tissue slices from failing heart. The automated T-wave detection was validated in a group of 390 EGMs obtained from cell monolayers and 145 EGMs from tissue slices. The validation showed that T-waves could be detected effectively using the automated method. In specific, T-waves were correctly identified in 231 EGMs (59.2%) using the morphology analysis algorithm, they were erroneously detected in 125 EGMs (32.1%) and there was no T-wave detection in 34 cases (8.7%). In tissue slices, the T-wave was precisely identified 87 times (60%), while it was incorrectly identified in 58 EGMs (40%). The manual identification of T-waves was proved to be more reliable and this was confirmed by quantifying the APD₉₀ using concurrent optical imaging data, as presented in more detail in Chapter 4.

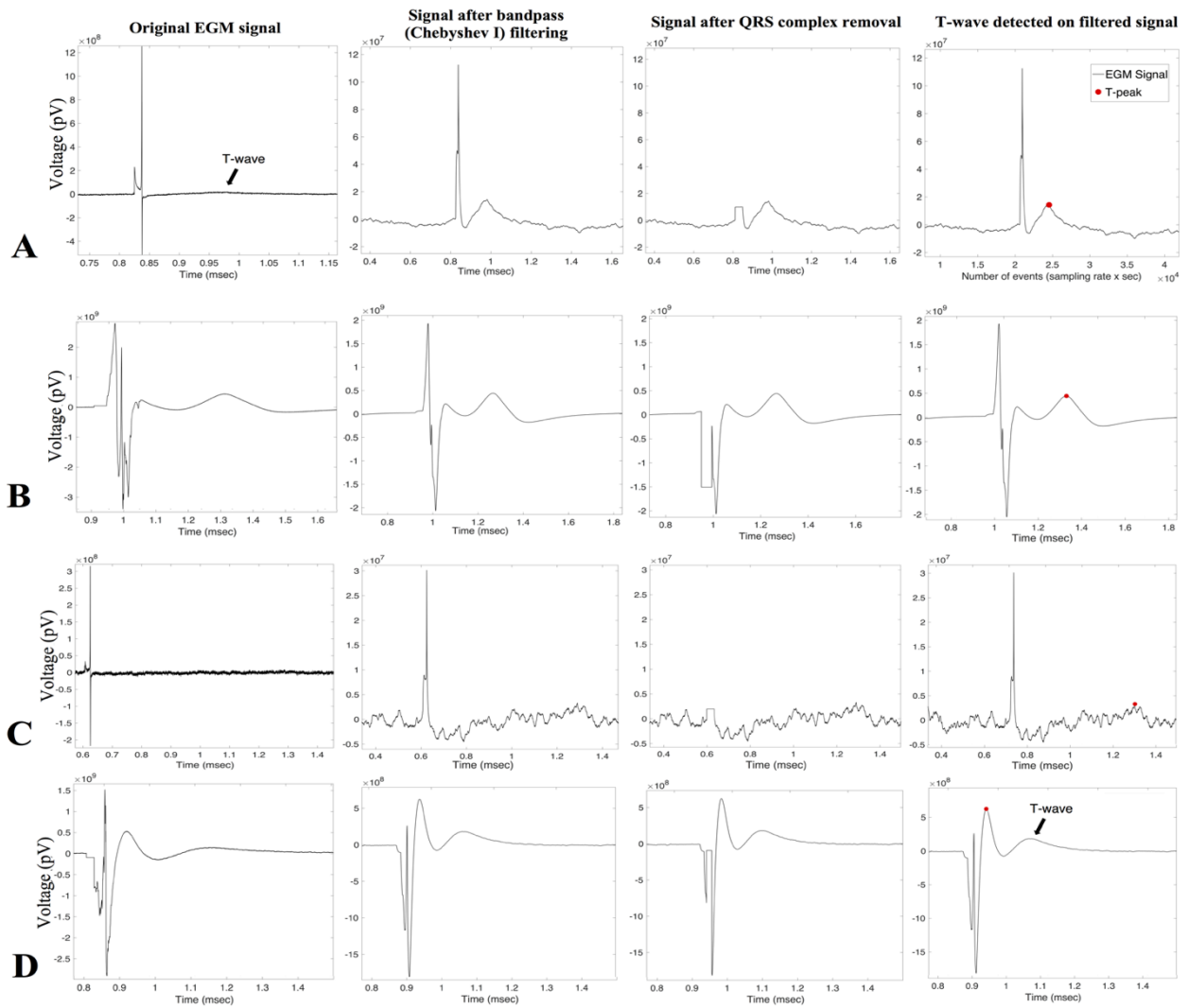


Figure 3.6: The process to identify a T-wave in EGM signal obtained from NRVM monolayers (A, C) and tissue slices from failing heart (B, D). (A) EGM recording where the T-wave is obvious, even though shallow, in the original signal. The T-wave was efficiently detected in the filtering process. (B) EGM signal from a tissue slice with an originally prominent T-wave, which was also successfully identified in the end. (C) An EGM with flattened T-wave. The filtering method did not identify the correct T-wave and the detected location is only part of the baseline signal. (D) An EGM obtained from a tissue slice where the wrong wavelet was identified as a T-wave. The proper T-wave follows the detected one.

One reason for the low accuracy of the algorithm is the absence of T-waves or the presence of shallow T-waves, as is the case in Figure 3.6C. These are issues prohibiting the manual detection of T-waves too. This problem mostly occurred in cell monolayer data though. Another reason for the less effective automated detection of T-waves can be the abnormal ST-segment morphology, especially in EGMs collected from tissue slices. There were a number of observations similar to the one presented in Figure 3.6D, where the ST-segment formation could be interpreted as an elevated J wave-like formation or an inverted T wave. This abnormality led to the identification of the wrong peak as the one belonging to a

T-wave. T-waves could not be detected in any of the clinical recordings, either manually or by the automated method, due to the noise of the signal.

3.4 Discussion

This chapter presents the algorithms that were used for automatically processing unipolar extracellular EGMs collected from cell monolayers, tissue slices and patients, and the algorithms used to analyse EGM morphology and compute signal features. Some of the time-domain analysis algorithms were also validated for their accuracy against manual measurements. The automated analysis was preferred because it is faster than the manual analysis and in order to reduce the errors due to human interaction and increase the consistency of results. Also, the automated analysis enables the extraction of complex non-morphological features.

Supervised machine learning needs a dataset of labelled instances, which are characterised by an array of features. One strategy is that the feature selection is undertaken by an electrophysiologist, because he/she is supposed to have a clear understanding about those which are most informative (10). However, our experience of using unipolar extracellular EGMs for identifying specific cellular and structural abnormalities is currently poor and the contact intracardiac EGM morphology has not been exhaustively investigated. Thus, a range of 37 features was chosen, including those with known clinical relevance, such as the fractionation and peak-to-peak voltage (37), the dominant frequency (118) and the Shannon entropy (127). Other features were chosen based on current trends in the analysis of EGMs for machine learning (153) and the rich experience of training algorithms to classify the electrocardiogram (ECG), which is collected from the body surface. The features that were presented in this chapter were derived from time-domain, frequency-domain and time-frequency analysis, in addition to features based on information theory and non-linear dynamics. This combination of characteristics was used for the analysis of unipolar extracellular EGMs obtained from NRVM monolayers, tissue slices derived from healthy and failing hearts and persistent AF patients.

Early attempts to use the ECG morphology for improving the diagnosis of cardiac arrhythmias occurred in the late 1990s (234). A number of morphological ECG features were used for the development of a dataset which was later used for supervised machine learning (234). Some of those features were also used here, even though they were adapted for EGM morphology, such as the CWT-based features from the time-frequency analysis. There is an increasing number of studies over the last decade which attempt

to develop ECG morphology-based predictive models focused on the improvement of AF treatment (213,235), as these are more accessible and non-invasive electrical recordings. Since there is a necessity for more accurate discrimination of AF types, research groups used not only time-domain ECG features, but also characteristics based on atrial activation times, frequency-domain analysis, signal quantization and the morphological similarity of activation waves, in order to describe the state of organisation of CFAE (213,235), which are intact recordings collected from the cardiac surface. The approach of combining features obtained from different types of signal processing has been recently adopted by researchers who are interested in intracardiac EGMs, in order to optimise the catheter ablation strategies for more efficient treatment of any AF condition (157,212,213). The features used in those EGM studies were mostly based on previous ECG studies and they only extended to the use of innovative time-frequency and non-linear dynamics features (153,157). A similar approach is also followed in this thesis.

The majority of algorithms for time-frequency analysis and information theory features were directly available in the MATLAB environment, which was used for the development of the signal processing script and the automated feature extraction. Other features, such as dominant frequency and T-wave detection, were based on previously published algorithms. However, the algorithms for the identification of QRS complex components were original. For this reason, they were validated, and they were proved to be accurate since there was no difference in measuring some of the time-domain analysis manually or by applying the automated method. The detection of the QRS complex is challenging, due to its time-varying morphology which may be caused by physiological variation factors. Previously published methods for the automated detection of the QRS complex were based on ECG recordings. In most cases, they included a pre-processing stage consisting of both linear and non-linear filtering of the ECG for the exact determination of the temporal location of the assumed QRS candidate (236). This was followed by one of different approaches for signal analysis, including methods based on signal derivatives and digital filters and wavelet-based QRS detection methods (237). The former approaches attenuate signal components, such as P-wave, T-wave, baseline drift and signal noise and the application of filters in order to make the QRS complex more prominent. Wavelet-based methods convert the signal into a time-scale representation using wavelet transform, which uses a set of functions that allows a variable time and frequency resolution for different frequency bands. The peak detection can then be carried out using the wavelet transform output and the wavelet coefficients. Apart from these approaches, which have been applied for detecting T-wave parameters, others are based on Hilbert Transform, mathematical morphology and matched filters (237). The method for QRS complex that was presented in this chapter was preferred compared to previous methods, because it allowed minimal distortions on the original signal, it could be easily applied both for *in vitro* and *ex vivo* data that have

different degree of noise and the results could be easily validated for accuracy and precision of feature detection.

3.4.1 Limitations

In this chapter, I aimed to apply a method for the detection of T-wave peak and the subsequent computation of features related to it, which have more clinical importance and are currently used in the clinic for the identification of arrhythmic conditions, such as QT interval, or features with an experimental importance for characterising EGM morphology changes, such as FPD and T-wave amplitude. However, due to the absence of a detectable T-wave in recordings obtained from cell monolayers and the variable ST-segment morphology in EGMs obtained from tissue slices, the algorithm was susceptible to errors. Moreover, the noise that occurred in the clinical EGM recordings could not be eliminated efficiently, due to the low frequency of the noise and its high voltage. Thus, any low-pass or bandpass filters applied to optimise the EGM signal affected significantly the morphology of the EGMs. Another problem due to the noise was the inability to detect T-waves both manually and using the automated analysis. Time-frequency analysis may be also affected by this fact. In conclusion, the T-wave detection presented in this chapter could be still used for the detection of T-wave features, extending that to mainstream and clinically relevant features, but as already presented in section 3.3.3 and discussed this heavily depends on signal quality. Therefore, and in order to be consistent throughout data analysis, any T-wave related features presented in next chapters, mainly FPD, were manually analysed, as it seems that this method could give safer results.

3.5 Conclusions

Our knowledge about the relationship of the cardiac electro-architecture leading to cardiac arrhythmias and the subsequent unipolar EGM morphology is still limited, and there is no universal set of features to be used for the characterisation of an EGM. The same occurs regarding the mathematical techniques for feature extraction, especially for recordings obtained from cell monolayers and tissue slices, since a standard method does not exist. Combining the signal processing techniques presented in this thesis, this gap in knowledge and signal processing experience is attempted to be reduced. Furthermore, the application of these algorithms on unipolar extracellular EGMs obtained using MEAs has shown that a number of features can be reliably detected, even though there is still space for improvement regarding some of the features.

Chapter 4

Characterisation of EGM morphology modifications
due to ion current modulations

4.1 Introduction

As already mentioned in the introduction chapter of this thesis, it remains unclear how the cellular and tissue level factors specifically influence EGM morphology. Ion channel abnormalities, tortuous conduction paths through fibrotic tissue, conduction slowing, rotational activity, wavefront collision and far-field signals are some of the mechanisms that lead to complex EGMs (242). There has been the assumption for the last decade that complex fractionated EGMs represent areas of disease (106). Based on this, there was a trend ablating these sites as a way to treat atrial fibrillation. The improvement of catheter ablation treatment is restrained by such simplistic assumptions and no recent increase in success rates has been achieved (174,243). A greater understanding of the cellular and tissue-level abnormalities implicated in the cardiac cell function and leading to specific modifications of EGM morphology could be beneficial for ablating strategies targeting the underlying aetiology. This chapter aims to examine this relationship and it is mostly focused on the effects of ion channel blockade on EGMs obtained from cell cultures and how these modulations are evaluated by corresponding modifications in AP morphology. It is believed that once the forward relationship of abnormalities leading to specific EGM changes is established, then the reverse pathway of predicting these abnormalities can be clear as there will be a biological meaning alongside the machine learning predictions.

4.1.1 Simultaneous electrogram and action potential recordings

Cardiac mapping for examination of the spatiotemporal distribution of excitation is carried out either by extracellular potentials (V_e) or optical transmembrane potentials (optical V_m) (244). An accepted approach to determine excitation time at a given location is to identify the time of the maximum downward or upward slope of the deflection that occurs in the V_e or optical V_m (244). Different methods have been developed for the characterisation of EGM correlates of AP changes (40,244,245). Some of these methods combine extracellular recordings of electrical activity using MEAs with simultaneous patch clamp recordings in the vicinity of individual microelectrodes for AP recordings (246,247). Other methods combine MEAs with optical mapping in a variety of settings, as different microelectrode arrangements and optical mapping devices are used. Our team has recently developed a system which allows co-localised and concurrent recordings at a single cell resolution within 2D multicellular preparations (40). Three-dimensional simultaneous recordings have been also presented using elastic membranes, which incorporate multifunctional sensors, electronic and opto-electric components,

shaped to match the epicardium (248). Such systems can be used for direct and comprehensive investigations of how cellular level heterogeneities, which affect AP morphology and its propagation, impact EGM morphology.

4.1.2 AP and EGM morphology affected during electrical remodelling

Some of the features that characterise the waveform of the extracellular EGM can be satisfactorily predicted from the action potential shape using systems for simultaneous EGM and optical mapping recordings, as already described. For example, highly selective hERG channel blockers are expected to prolong both APD and FPD (249). In addition, it has been shown that L-type Ca^{2+} channel blockade using nifedipine shortens both APD₉₀ and FPD in NRVMs (40), rabbit Purkinje fibres (250) and guinea pig papillary muscle (251). However, the effects of multichannel blockers, such as quinidine or flecainide, on AP shape are difficult to predict (249). Since there is variable contribution of individual channels on AP shape depending on species and location within the heart of one species, the effect of a pharmacological agent may depend on the predominating conductance during the plateau or repolarisation phase and any overlapping active concentration ranges. Thus, these phenomena may result in simultaneous APD prolongation and shortening (252). Similar experimental difficulties may also occur when combinations of ion channel blockers are used (249). The result is limited applications of concurrent EGM and AP recordings in cardiac electrophysiology and still a poor understanding of electrical remodelling factors being responsible for a particular EGM morphology.

The variability of EGM morphology has not been fully characterised and, as it has become clear through *in vitro* and *in vivo* studies, we have a limited understanding of the underlying mechanisms leading to EGM modifications. *In vitro* models are preferred for this investigation because of the degree of manipulation which is possible, compared to *in vivo* experimental conditions and the reduction of complexity such that changes can be more directly correlated with a single phenomenon. In addition, NRVM monolayers are a well-established cell model in cardiac electrophysiology and thoroughly characterised (172).

Therefore, I sought to investigate how specific ion channel abnormalities, related to a variety of cardiac diseases, change the EGM morphology and which is the EGM morphology that could characterise each one of these heterogeneities. The hypotheses to be addressed were:

1. The Na^+ -channel blockade can modify both the upstroke duration and the EGM duration.
2. K^+ blockers lead to APD and FPD prolongation.

3. K^+ current enhancement and L-type Ca^{2+} channel blockade can decrease both APD and FPD.
4. EGM morphology is determined by identifiable features of local action potentials and the architectural determinants of their propagation.
5. Characterisation of more complex additional features will identify further EGM modification occurring upon ion channel blockade.

For this purpose, the following studies described in this chapter were carried out:

1. to change the action potential morphology and conduction velocity using a variety of pharmacological agents,
2. to conduct simultaneous optical mapping and EGM recordings, in order to confirm EGM morphology changes due to specific abnormalities, and
3. to fully characterise the EGM morphology that occurs following the reduced ion channel function *in vitro* and identify modifications on specific features.

4.2 Methods

4.2.1 *In vitro* model of cellular component of scar

The isolated cells were seeded on MEA plates as described in section 2.2.1. As a model of cellular scar, four experimental cell groups were used in this thesis: ‘myocytes only’ (MO), which were the after preplating cells received in the end of the cell isolation process (presented in section 2.2.1), the ‘natural composition’ (NC) cell population isolated from neonatal rat ventricles, and myocytes mixed with low (+20%) or high (+40%) amounts of cultured fibroblasts. Regarding the MO and NC groups, 200,000 cells were plated on the collagen substrate prepared on each MEA dish. In NRVM/fibroblast co-cultures, the low-fibroblast cell culture contained 80% after preplating NRVMs (180,000 cells) and 20% 1-week old cultured fibroblasts (20,000 cells), and the high-fibroblast cell culture contained 60% after preplating NRVMs (160,000 cells) and 40% 1-week old cultured fibroblasts (40,000 cells). Cells were mixed as appropriate prior to seeding on MEA dishes, in order to ensure the homogeneous distribution of fibroblasts.

4.2.2 Ion channel blockade

The ion channel activity of cell monolayers was modified using pharmacological agents, as mentioned in section 2.2.3. A range of concentrations was studied for each of the ion channel modulators. The responsiveness of cells to Na⁺ channel blockade was achieved using increasing concentrations of lidocaine (myocytes only/NC: 0 – 100µM; low/high fibroblast – NRVM co-cultures: 0 – 50µM). K⁺ channel blockade was carried out using 4-aminopyridine (4-AP) for I_{to} blockade (myocytes only/NC: 0 – 1.5mM; NRVM/low fibroblast co-culture: 0 – 300µM), E-4031 for I_{Kr} blockade (myocytes only/NC/NRVM-low fibroblast co-cultures: 0 – 10µM; NRVM/high fibroblast co-cultures: 0 – 1µM) and HMR-1556 for I_{Ks} blockade (myocytes only/NRVM-low fibroblast co-cultures: 0 - 30µM; NC: 0 – 100µM). The opening of K_{ATP} channels was achieved using pinacidil (0 – 30µM). Nifedipine was used for L-type Ca²⁺ channel blockade (myocytes only/NC: 0 – 10µM; NRVM/low- or high-fibroblast co-cultures: 0 – 30µM). Electrical remodelling was also achieved by the administration of a mix of lidocaine and nifedipine (Tables 4.1 – 4.2). All agents were dissolved in HBSS supplemented with 1mM MgCl₂ and 1.5mM CaCl₂.

Cell responsiveness was tested 3-4 days after plating cells on MEA dishes. Cell monolayers consisting of NRVMs were decided to be ready for electrical remodelling experiments when conduction velocity was >20cm/sec. However, this rule was not applied on the NC monolayers and co-cultures of NRVM and fibroblasts, as the conduction velocity was ranging between 12-19cm/sec at the baseline state in these cases. EGM data obtained from the range of concentrations for each pharmacological agent, and according to the protocol presented in section 2.3, were firstly used for the creation of dose response curves corresponding to each cell culture group and modification. The IC₅₀ of each drug was extracted using the dose response curves and dual modality performance experiments were carried out using that concentration. The automated EGM morphology analysis was applied on EGMs collected before and after the administration of a drug concentration close to the IC₅₀. Details of the concentrations used are presented in the following sections of this chapter.

4.2.3 Measurement of conduction velocity across each MEA

A custom in-house MATLAB script was used for the calculation of conduction velocity with the acquired MEA data and the generation of isochronal maps for the demonstration of wave propagation. The conduction velocity across each array was measured using electrode activation times, $(-dV/dt)_{max}$, at each electrode and its four neighbouring electrodes, in order to create conduction vectors. These

vectors were used for identifying the direction of wavefront propagation as described by Salama *et al.* (253) and the creation of isochronal activation maps (Figure 4.1).

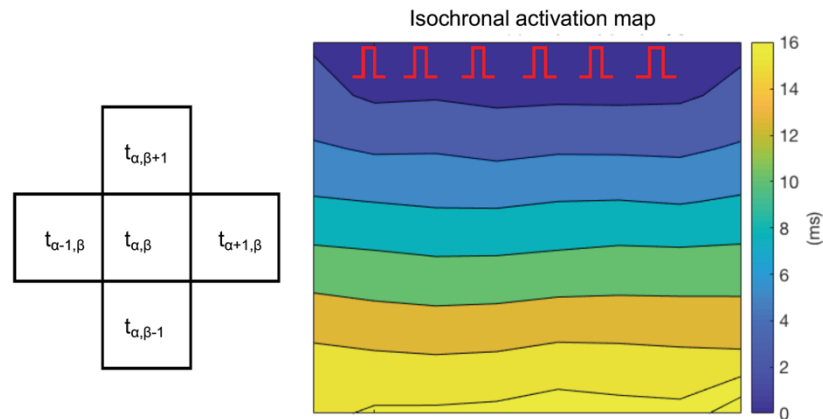


Figure 4.1: Conduction velocity measurement across MEA. (Left) The MEA is arranged in a series of grids with stable inter-electrode distance. Thus, the calculation of the average conduction velocity can be carried out using the activation time of each electrode in position α , β and its neighbours. Adapted by Salama *et al.* with permission of the rights holder, Wolters Kluwer Health, Inc. (Right) Isochronal activation map representing the wavefront propagation in a cell culture.

4.2.4 Synchronisation of dual modality concurrent electrogram and action potential recordings

A custom-made optical system, developed by our team (40), was used for simultaneous membrane voltage and MEA recordings (supplied by Cairn Research, UK) (Figure 2.4). The stage height of the upright microscope in this optical system was modified, in order to hold the amplifier of the MEA system. During dual modality recordings, the MEA plate was connected to the amplifier and the amplifier was loaded on the microscope stage. A BNC connector was built into the MEA stimulator to deliver a signal at the time of stimulus to an external hardware, thus the optical imaging system. for the generation of a blank frame synchronised with the MEA stimulation. A BNC sync cable connected the MEA stimulator to the optical mapping camera, in order to transfer the output signal to the camera. Therefore, the output signal of the MEA stimulator was used for triggering the optical imaging system for recordings for as long as the MEA recordings last ensuring the synchronization between systems. The synchronised trigger resulted to the production of a false stimulation artefact on the optical recording which confirmed the accurate time correlation between the two recording systems in the final results. MEA recordings were made for 10sec and during this time period 5-sec ratiometric voltage optical mapping recordings were simultaneously obtained. The blanking circuit implemented in the

MEA amplifier and the ability of the stimulator to create automated trigger pulses also ensured the synchronisation of recordings. Also, these MEA components made sure that the optical imaging recordings would stop simultaneously with the MEA recordings. More in-depth information about the MEA and optical mapping systems used for concurrent recordings are presented on Chapter 2 (MEA: section 2.3; optical mapping: section 2.5.2).

The IC_{50} of each pharmacological agent was used for dual modality experiments. Control recordings were firstly obtained, and these were followed by the administration of a drug and the incubation of cells/slices with it for 5min. Simultaneous EGM and optical mapping recordings were then obtained, in order to study the effects of electrical remodelling on EGM and AP morphologies. The same procedure was followed for each type of cell culture. For the correlation of MEA and optical imaging recordings, the linear regression of simultaneously occurred feature changes in recordings were analysed.

EGM duration was indirectly correlated to AP upstroke duration and FPD was directly correlated to APD_{90} . EGM duration was analysed using the automated method presented in section 3.2.3. FPD measurements presented in this and following chapters is derived from manual analysis, and therefore this analysis was separate from the automated analysis for the rest of EGM features which was presented in Chapter 3. The manual analysis was conducted as presented in chapter 3 and (40), and this occurred the time duration between S-peak and the peak of the following T-wave.

4.2.5 Statistical analysis

All data are presented as mean \pm SEM. A two-tailed Wilcoxon matched-pairs signed rank test was used for the analysis of dual modality performance data. Linear regression analysis with 95% confidence interval was used for investigating the relationship between AP and EGM characteristics. In order to investigate the quality of linear regression, r^2 and the slope are reported. In addition, and as a way to study how well the model fits the data, the standard deviation of the residuals was used. The residual is the vertical distance of a point from the fit line. $Sy.x$ was used for this purpose and this is measured from the following equation:

$$Sy.x = \sqrt{\frac{\sum(residual)^2}{n - K}} \quad (4.1)$$

where $n - K$ is the denominator, with K being the number of parametres fit by regression. The value $n - K$ is the number of degrees of freedom of the regression. In cases of linear regression using data obtained

from one type of cell culture or tissue slices, then the $Sy.x$ is the same as RMSE. However, when linear regression with data obtained from multiple types of cell culture occurs, then $Sy.x$ is larger than RMSE, but is also a better estimate of goodness-of-fit.

Two-tailed t -test for non-parametric data (Mann-Whitney test) was used for the analysis of EGM features before and after the administration of pharmacological agents unless otherwise stated. A value of $p < 0.05$ was considered significant (*= $p < 0.05$, **= $p < 0.01$, ***= $p < 0.001$, ****= $p < 0.0001$).

4.3 Results

4.3.1 The effects of Na^+ channel blockade on EGM and action potential morphology

Na^+ channel blockade using the IC_{50} of lidocaine ($27\mu M$) increased both the upstroke duration (control: 16.77 ± 0.2 ms; lidocaine: 17.36 ± 1 ms; $n = 6$ arrays; $p < 0.05$) and the EGM duration (control: 3.03 ± 0.5 ms; lidocaine: 4.02 ± 1 ms; $n = 6$; $p < 0.0001$; mean \pm SEM) during concurrent optical mapping and MEA recordings on myocyte only NRVM monolayers (Figure 4.2A-B). A direct correlation between conduction velocity and EGM duration, and therefore upstroke duration, was showed by linear regression analysis (Control: $n=9$ EGMs, lidocaine: $n=27$ EGMs; $r^2=0.61$, $Sy.x = 4.3$, slope = -9.52 ± 1.3 , $p < 0.0001$) (Figure 4.2C). The Bland-Altman plot revealed that EGM duration was systematically 0.77 ms longer on average after I_{Na} blockade (Figure 4.2D). The differences did not seem to be affected by the average EGM durations.

In order to investigate the combined effect of functional and structural modifications on EGM morphology, Na^+ channel blockade was firstly inserted in NC cell cultures and co-cultures of NRVMs with different amounts of cultured fibroblasts. The IC_{50} was calculated for each cell group separately using conduction velocity data obtained at 1 Hz (NC: 370 nM, Low %FB: 701.8 nM; High %FB: $9.7\mu M$) (Figure 4.3A-C). The analysis of simultaneous optical mapping and MEA recordings showed that the upstroke duration did not change significantly within each group following the addition of lidocaine (NC: control – 15.94 ± 0.4 ms, $n=8$ arrays, lidocaine - 16.68 ± 1 ms, $n=5$ arrays; +20% FB: control – 16.54 ± 1.3 ms, $n=4$ arrays, lidocaine – 10.03 ± 4.4 ms, $n=4$; +40% FB: control – 16.05 ± 0.3 ms, $n=5$ arrays, lidocaine – 15.58 ms, $n=1$; mean \pm SEM) and there was no difference among groups ($p = 0.91$) (Figure 4.3D). EGM duration was affected by the amount of cellular scar ($p=0.004$), but not by the Na^+ channel blockade ($p=0.097$) (Figure 4.3E). No differences in EGM duration were observed within each group.

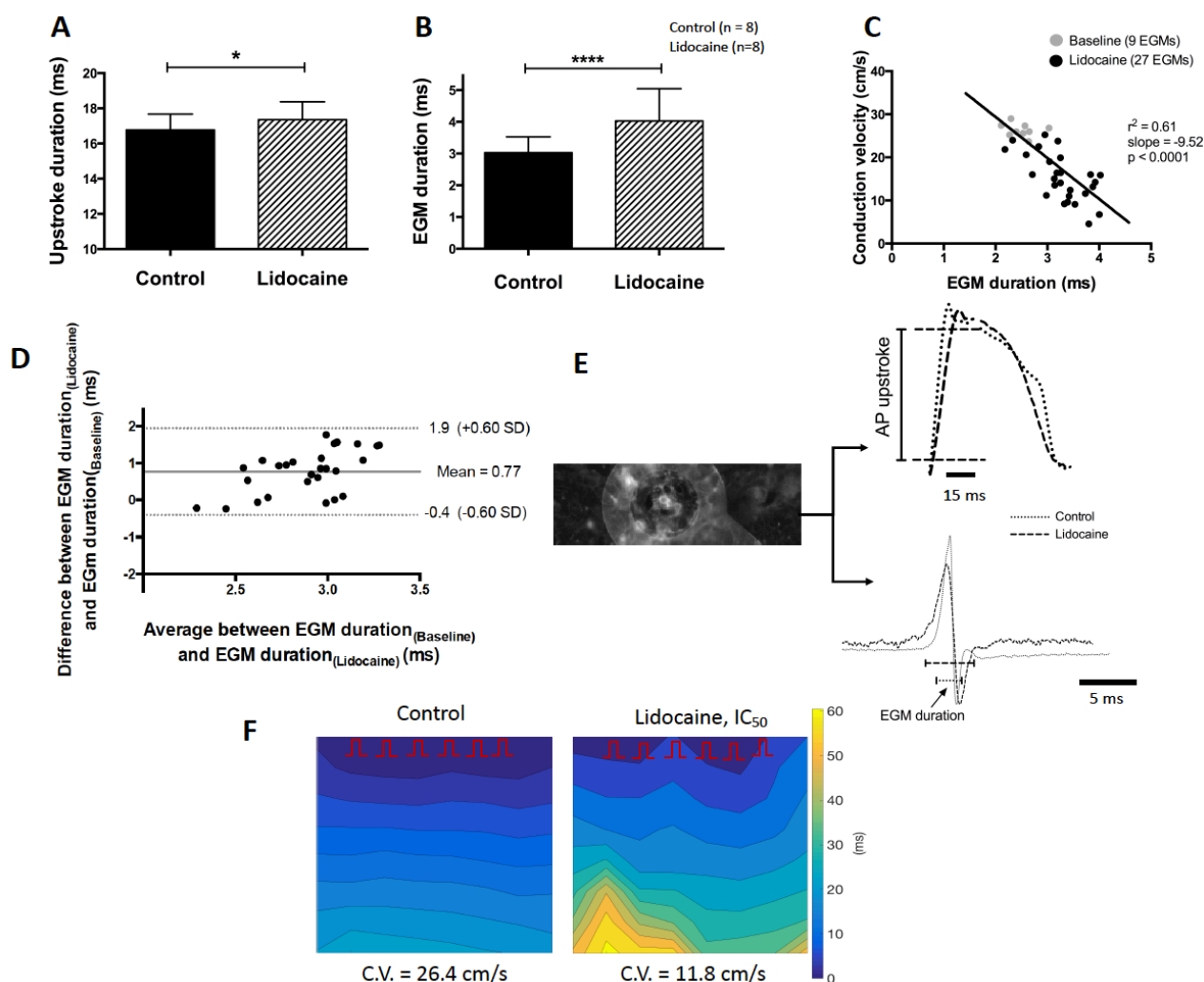


Figure 4.2: Effects of Na^+ channel blockade using lidocaine on action potential upstroke, EGM duration and conduction velocity in NRVM monolayers. (A) Action potential upstroke, recorded by NRVM cultures ($n=8$), was increased after the administration of $27\mu\text{M}$ lidocaine compared. (B) The same lidocaine concentration simultaneously prolonged EGM duration, as shown by the unpaired Student's t-test (two-tailed). (C) Correlation of EGM duration with conduction velocity before and after Na^+ channel blockade ($r^2=0.61$, $\text{Sy.x}=4.3$, $p<0.0001$). (D) Bland-Altman plot for EGM duration at baseline and after I_{Na} blockade. The middle bold line is the average difference between the baseline and I_{Na} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (E) (*Top*) An example of overlapped action potential traces, where the prolonged upstroke duration after the administration of lidocaine is shown. (*Bottom*) Superimposed EGMs obtained before and after Na^+ channel blockade from the same. Action potentials and EGMs were recorded from the same electrode shown on the left. (F) Isochronal activation maps for the wavefront propagation before (left) and after (right) the Na^+ channel blockade using lidocaine. All bar charts represent $\text{mean}\pm\text{SEM}$; * $p<0.05$; **** $p<0.0001$. Adapted from Chowdhury, Tzortzis *et al.* (40).

Conduction velocity changed due to both the amount of cellular scar ($p < 0.0001$) and Na^+ channel blockade ($p = 0.007$). There was a significant reduction in conduction velocity even within each cell group following the addition of lidocaine (NC: control - $21.62 \pm 6.9 \text{ cm/s}$, lidocaine - $14.6 \pm 7.6 \text{ cm/s}$, $n = 7$ arrays; +20% FB: control - $8.65 \pm 3.5 \text{ cm/s}$, lidocaine - $6.4 \pm 2.5 \text{ cm/s}$, $n = 5$; +40% FB: control - $24.62 \pm 4.5 \text{ cm/s}$, lidocaine - $16.24 \pm 4.7 \text{ cm/s}$, $n = 5$; $p < 0.01$; mean \pm SEM) (Figure 4.3F). A direct correlation between conduction velocity and EGM duration was revealed during linear regression analysis ($r^2 = 0.62$, slope = -2.9 ± 0.4 , $p < 0.0001$), but even though that relationship was maintained within the NC ($r^2 = 0.64$, $S_{y.x} = 4.9$, slope = -5.17 ± 1.2 , $p = 0.001$, $n = 13$ EGMs) and +20% FB groups ($r^2 = 0.51$, $S_{y.x} = 2.3$, slope = -1.05 ± 0.4 , $p = 0.03$, $n = 9$ EGMs), it was not present in the +40% FB group ($r^2 = 0.27$, $S_{y.x} = 5.6$, slope = -3.7 ± 2.2 , $p = 0.12$, $n = 10$ EGMs) (Figure 4.3G). APD_{90} and FPD were also affected by the amount of cellular scar ($p = 0.01$ and $p = 0.049$ respectively), and not by Na^+ channel blockade (Figure 4.3H-I). Linear regression analysis showed a direct correlation between APD_{90} and FPD (NC: $n = 122$ EGMs, +20% FB: $n = 71$ EGMs, +40% FB: $n = 106$ EGMs; $r^2 = 0.76$, slope = 0.86 ± 0.03 , $p < 0.0001$) (Figure 4.3J).

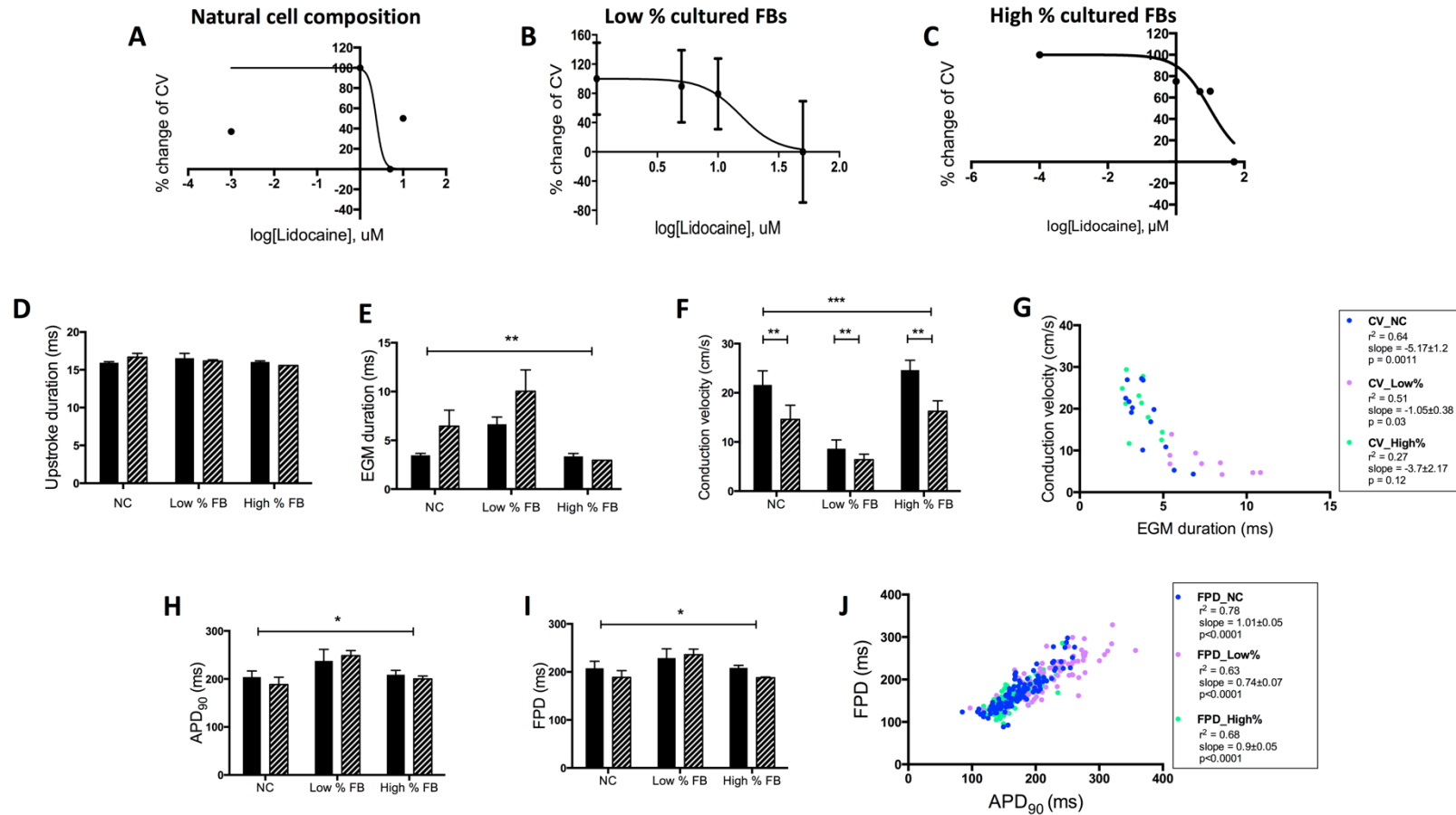


Figure 4.3: (A-C) Dose response curves generated after treating NC cultures (A) and co-cultures of NRVMs with low (B) or high amounts of fibroblasts (C) with a range of lidocaine concentrations. The IC₅₀ of each type of cell culture was extracted using the dose-response curves. (D-F) Effects of Na⁺ channel blockade on the upstroke duration (D), the EGM duration (E) and the conduction velocity (F) obtained from each cell culture type using lidocaine (IC₅₀). (G) Linear regression analysis for the relationship between conduction velocity and EGM duration. (H-I) Effects of lidocaine activity on the APD₉₀ (H) and FPD (I) obtained from each cell culture. (J) Linear regression analysis for the relationship between APD₉₀ and FPD extracted from the optical mapping and MEA data of each cell culture. Mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

4.3.2 EGM and action potential morphology manifestations of transient outward K^+ current blockade

During dual modality experiments, the blockade of the transient outward K^+ (I_{to}) current, due to use of 4-AP, resulted in significant effects both on AP duration and FPD (Figure 4.4). Cells were stimulated at progressively reducing pacing intervals from 1000 to 250 ms both before and after the administration of 4-AP ($IC_{50} = 739.9\mu M$) resulting to electrical restitution (Figure 4.4B-C). FPD was measured manually, as presented in section 4.2.4, due to the automated detection inefficiency as presented in 3.3.3. T-waves were manually detected as explained in section 3.2.3. APD at 90% repolarisation (APD₉₀) was affected by pacing cycle length (control: 1Hz 139.7±19.2 ms, 4Hz 118.7±9.4ms, n=4 arrays; 4-AP: 1Hz 154.6±29.7 ms, 4Hz 116.3±14.7 ms; n=4; p = 0.04; mean ± SEM) (Figure 4.4B). Both pacing rate (p = 0.04) and I_{to} blockade (p = 0.02) affected FPD (control: 1Hz 107.09±12.9 ms, 4Hz 87.3±20.3 ms, n=4; 4-AP: 1Hz 140±18 ms, 4Hz 105±16.6 ms, n=4; mean ± SEM) (Figure 4.4C), which was manually measured. Linear regression analysis showed strong correlation between APD₉₀ and FPD (Control: n = 50 EGMs, 4-AP: n = 41 EGMs; $r^2=0.76$, $Sy.x = 13.3$, slope=0.9±0.07, p<0.0001) (Figure 4.4D). The APD₉₀/FPD linear relationship did not show significant differences before and after the I_{to} blockade (control: $r^2=0.5$, slope=0.63, p=0.06; 4-AP: $r^2=0.62$, slope=0.83; p=0.18). The Bland-Altman plot revealed that FPD was systematically 14ms longer on average after I_{to} blockade (Figure 4.4E). The differences did not seem to be affected by the magnitude of average FPD.

The IC_{50} of 4-AP (803.3 μM) for NC cell cultures was measured based on the change of FPD at 1Hz EGM data. The I_{to} blockade at NC cell cultures had as a result the prolongation of APD₉₀ at 1Hz (control: 104.9±13.4ms, 4-AP: 125.9±10.7ms, n=1 array, p=0.016) (Figure 4.5B). There was also a trend towards prolongation of FPD at 1Hz (control: 106.2±19.7ms, 4-AP: 128.5±18.6ms, n=1, p=0.16; mean ± SEM) (Figure 4.5C). Electrical restitution occurred when stimulating cells at progressively reduced pacing intervals from 1000 to 250ms both in control state and after the I_{to} blockade (Figure 4.5D-E). Both APD₉₀ (control: 1Hz 132.3, n=1 array, 4Hz 93.6±1.77, n=3 arrays; 4-AP: 1Hz 133.2±6.9, n=4, 4Hz 105.1, n=1, p=0.001; mean ± SEM) and FPD (control: 1Hz 144, n=1, 4Hz 102±8.6, n=3; 4-AP: 1Hz 144±19.9, n=4, 4Hz 126, n=1, p=0.04; mean ± SEM) were affected not only by I_{to} blockade, but also by the pacing rate (p<0.01 and p<0.05 respectively). There was also a direct relationship between APD₉₀ and FPD ($r^2=0.64$, $Sy.x=12.6$, slope = 1.04±0.2, p<0.0001; n=19 EGMs) (Figure 4.5F).

The FPD results presented in this section involved the correct manual detection of T-waves. A T-wave was detected as the positive deflection after each QRS complex, a curve that can be seen after that complex. Due to the shallowness of the T-wave though in some occasions, this might be more difficult to be seen in certain examples, as the one presented in Fig. 4.5G.

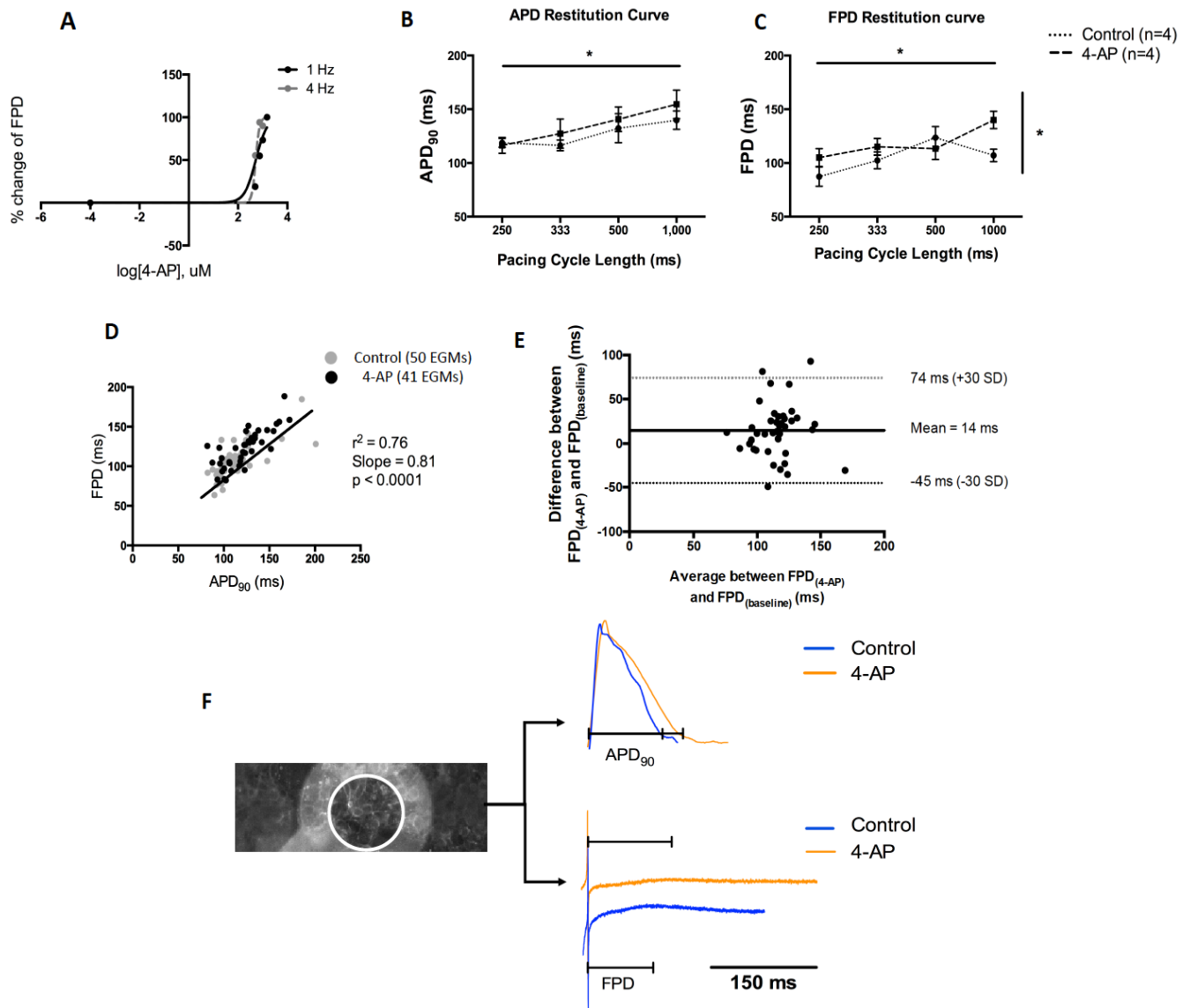


Figure 4.4: (A) Dose response curve for the extraction of IC_{50} of 4-AP following the administration of myocyte only NRVM cultures with a range of 4-AP doses. (B-E) Correlation between APD₉₀ and FPD in NRVM cultures following the I_{to} current blockade using 739.9 μM 4-AP. APD₉₀ restitution curve (B) and FPD restitution curve (C) (control: n=4 arrays; 4-AP: n=4 arrays). Both APD₉₀ and FPD changed significantly across the range of pacing rates ($p < 0.05$), but only FPD was affected by I_{to} blockade ($p < 0.05$, ordinary two-way ANOVA). (D) Linear regression correlating APD₉₀ with FPD, using data obtained before and after the administration of 4-AP ($r^2 = 0.76$, $Sy.x = 13.3$, $p < 0.0001$). (E) Bland-Altman plot for FPD at baseline and following the administration of 4-AP. The middle bold line is the average difference between the baseline and I_{to} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Superimposed raw traces derived before and after I_{to} current blockade (top) and the concurrent EGMs (bottom). All traces were obtained from the same electrode indicated on the left. Mean \pm SEM; * $p < 0.05$.

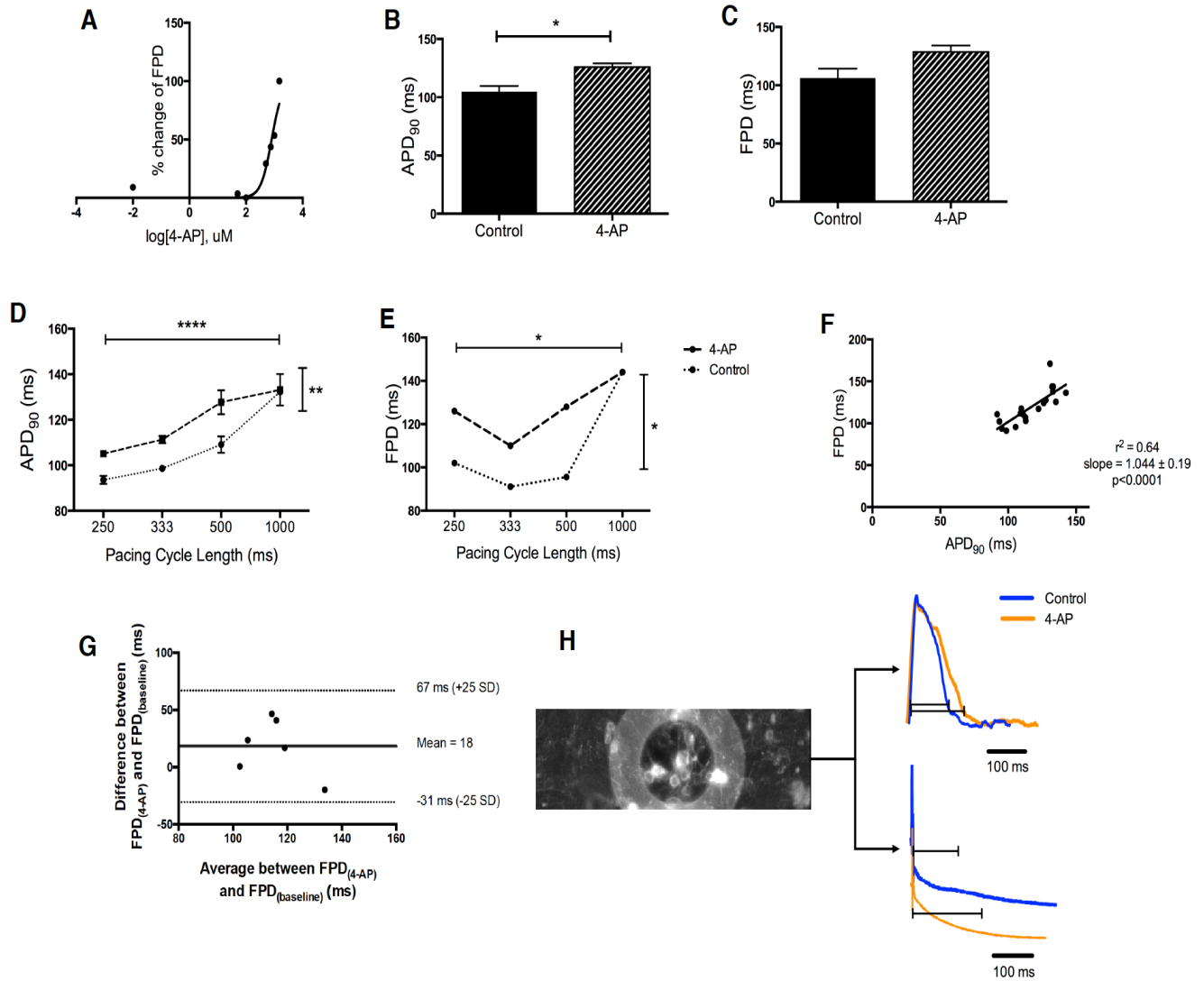


Figure 4.5: (A) Dose response curve for the extraction of IC_{50} . (B-C) APD_{90} (B) increased ($p = 0.016$ following the administration of 4-AP on NC cell culture, and there was a trend for simultaneous FPD increase (C), which was not significant ($p = 0.16$) (control: $n=1$ array; 4-AP: $n=1$ array). (D-E) APD_{90} and FPD restitution curves showing that both pacing rate and 4-AP had a significant effect on both. (F) Linear regression analysis for the relationship between APD_{90} and FPD ($r^2 = 0.64$; $Sy.x=12.6$, slope = 1.04 ± 0.2 ; $p<0.0001$). (G) Bland-Altman plot for FPD at baseline and following the administration of 4-AP on NC cell culture. The middle bold line is the average difference between the baseline and I_{to} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (H) Superimposed raw traces (top) and EGMs (bottom) obtained simultaneously before and after the administration of 4-AP (IC_{50}) on NC cell population. Mean \pm SEM. * $p<0.05$; ** $p<0.01$; *** $p<0.0001$.

4.3.3 The effects of L-type Ca^{2+} current blockade on EGM and action potential morphology

Rate adaptation of APD_{90} and FPD was observed, through manual analysis of both features, before I_{CaL} blockade using the IC_{50} of nifedipine (772nM). The I_{CaL} blockade caused both APD_{90} (control: 1Hz 143.44±29 ms, 4Hz 116.13±18 ms; nifedipine: 1Hz 111.3±23 ms, 4Hz 107±19 ms, n=8 arrays, p<0.001) and FPD (control: 1Hz 138.3±29 ms, 4Hz 108.8±19 ms; nifedipine: 1Hz 98.5±31 ms, 4Hz 92.4±17 ms, n=8, p<0.0001) abbreviation (Figure 4.6B-C). However, following the addition of nifedipine both indices were not affected significantly by pacing rate. There was a strong correlation between APD_{90} and FPD as shown by linear regression analysis ($r^2=0.86$, $\text{Sy.x}=12$, slope = 0.93 ± 0.04 , p<0.0001; n=49) (Figure 4.6D). The APD_{90} /FPD correlation was similar before and after the administration of nifedipine (control: $r^2=0.84$, slope=0.74; nifedipine: $r^2=0.81$, slope=0.75; p=0.65; n=49). The Bland-Altman plot revealed that FPD was 29ms shorter on average after I_{CaL} blockade in a stable basis (Figure 4.6E). The differences did not seem to be affected by the magnitude of average FPD and approximately 4.6% of the differences were out of the limits of agreement for FPD.

The IC_{50} of nifedipine for NC cell cultures was 65nM. No restitution curve could be created, using the dual modality performance experiments, as cells could be stimulated only at 1Hz and 2Hz. The data obtained from concurrent optical mapping and MEA showed that APD_{90} was reduced after the I_{CaL} blockade (control: 250.2±14.2ms, n=7 APs; nifedipine: 230.2±21.4ms, n=12 APs; p=0.04; mean ± SEM) (Figure 4.7B). FPD was simultaneously shortened after the addition of nifedipine (control: 247.7±18.3ms, n=7 EGMs; nifedipine: 225.7±25.3ms, n=12 EGMs; p=0.04; mean ± SEM) (Figure 4.7C). Linear regression analysis showed a strong correlation between APD_{90} and FPD ($r^2=0.59$, $\text{Sy.x}=14.4$, slope = 0.91 ± 0.15 , p<0.0001, n=19). Conduction velocity did not change significantly due to nifedipine (control: 16.83±5.9cm/s, nifedipine: 12.74±3.2cm/s, n=3, p=0.25; mean ± SEM) (Figure 4.7E). The Bland-Altman plot revealed that FPD was systematically 22ms shorter on average after I_{CaL} blockade in NC cell monolayers (Figure 4.7F). The differences seem to be affected by the magnitude of average FPD, as there is a trend towards decrease, but this may be biased by the limited of data.

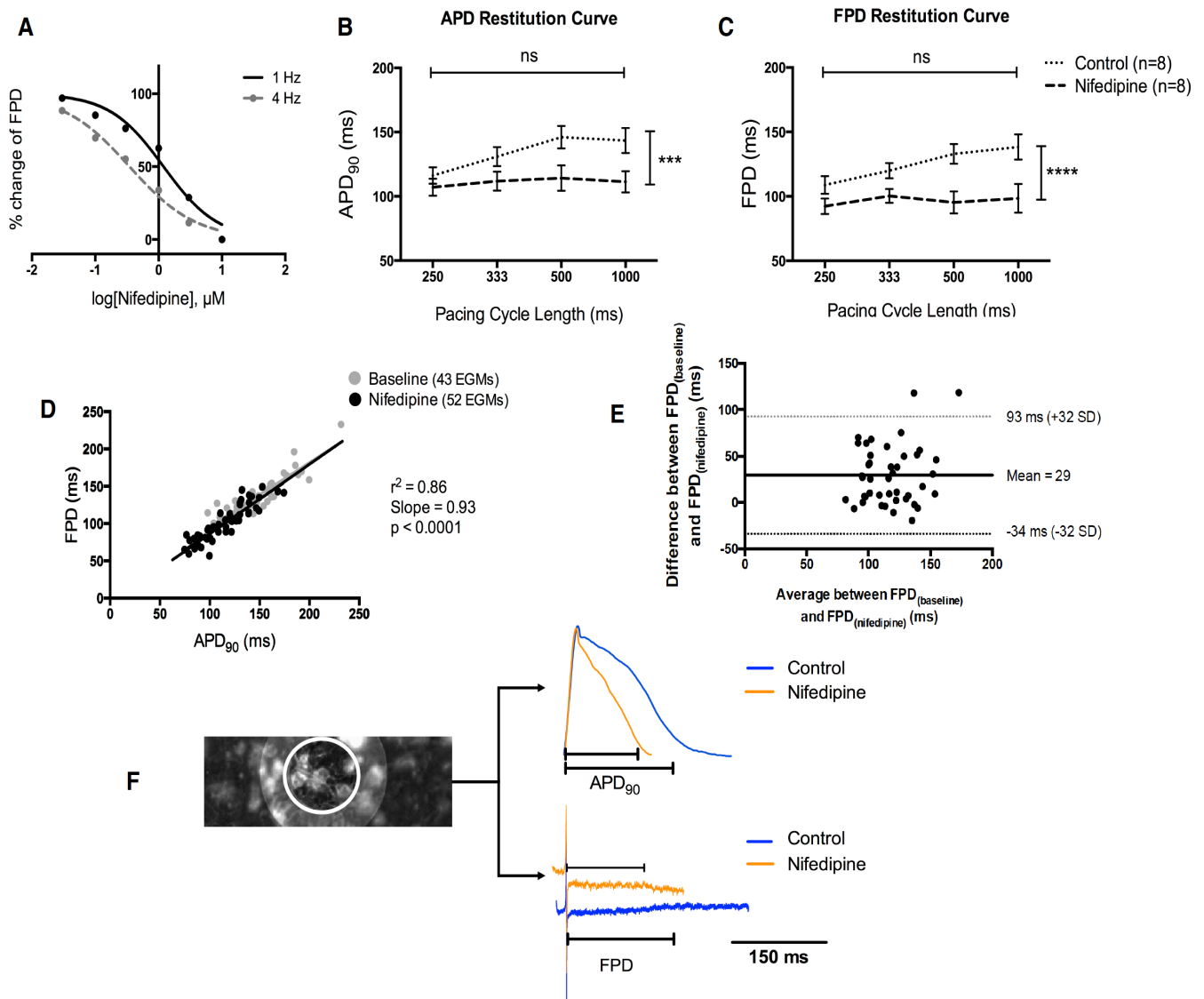


Figure 4.6: (A) Dose response curve for the measurement of the IC_{50} of nifedipine based on the effect of I_{CaL} blockade on FPD. (B-C) Correlation between APD_{90} and FPD in cardiac myocyte monolayers after the L-type Ca^{2+} channel blockade using 772nM nifedipine. APD_{90} (B) and FPD (C) restitution curves (control: n=8 arrays; nifedipine: n=8 arrays). APD_{90} and FPD were significantly affected by nifedipine ($p < 0.001$ and $p < 0.0001$ respectively, ordinary two-way ANOVA), as shown following the manual analysis of both features. (D) Linear relationship between APD_{90} and FPD ($r^2 = 0.86$, $p < 0.0001$). (E) Bland-Altman plot for FPD at baseline and following I_{CaL} blockade in NRVM cells. The middle bold line is the average difference between the baseline and I_{CaL} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Superimposed action potential traces (top) and EGMs (bottom) before and after the administration of nifedipine. All traces were derived from the same electrode, shown to the left. Mean \pm SEM; *** $p < 0.001$; **** $p < 0.0001$.

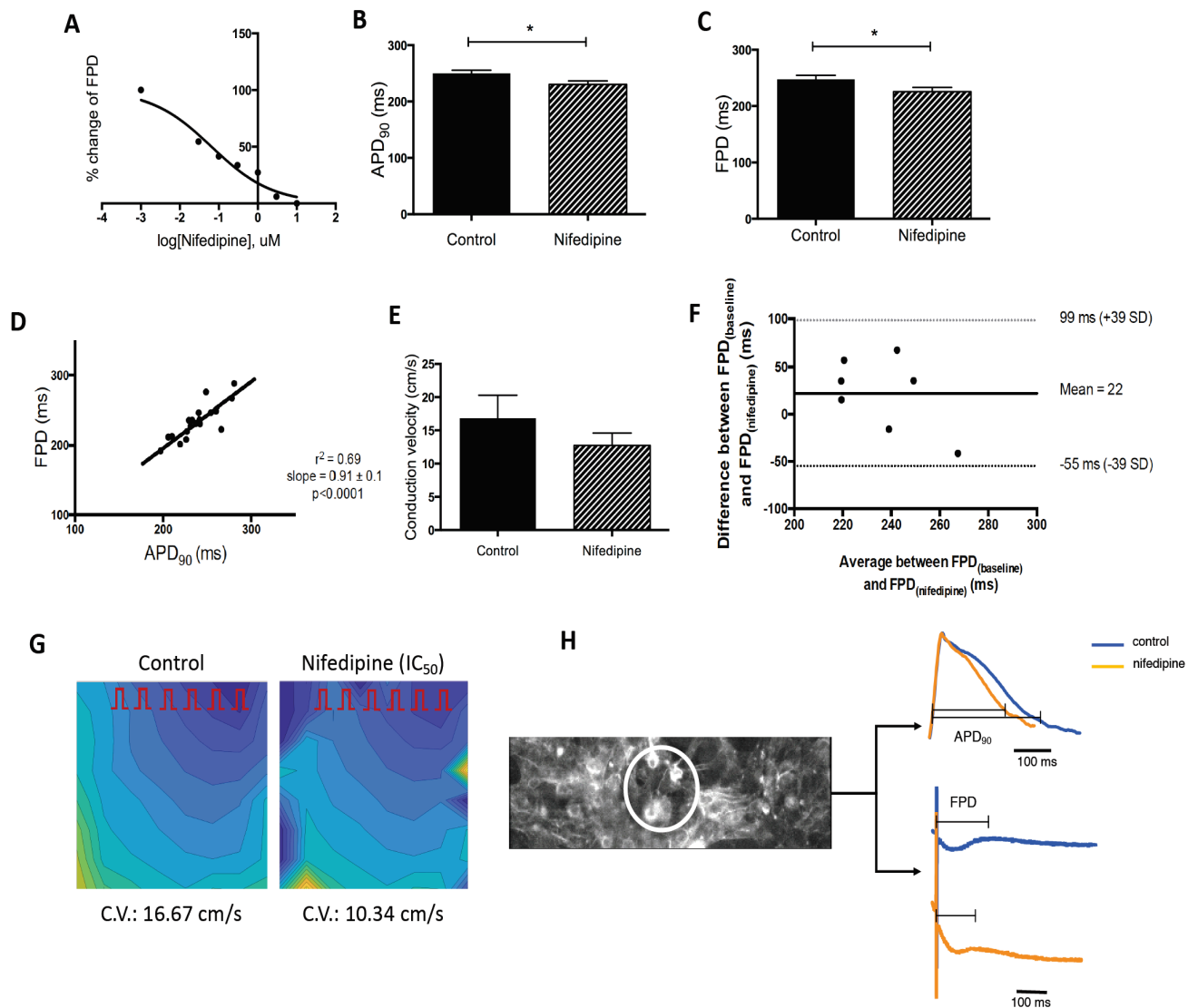


Figure 4.7: (A) Dose response curve for the extraction of IC₅₀ of nifedipine after its administration on NC cultures. (B-C) Mann-Whitney test results presenting the effect of I_{CaL} blockade on action potential (B) and EGM morphology (C) using nifedipine. APD₉₀ and FPD manually annotated data were used for this analysis. (D) Linear regression analysis for the relationship of FPD and APD₉₀ ($r^2 = 0.69$; $Sy.x=14.4$, slope = 0.91 ± 0.15 ; $p < 0.0001$; $n=19$). (E) No significant change of conduction velocity was observed following nifedipine administration ($p=0.25$). (F) Bland-Altman plot for FPD at baseline and following I_{CaL} blockade in NC cell monolayers. The middle bold line is the average difference between the baseline and I_{CaL} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (G) Isochronal maps for the wavefront propagation before (left) and after (right) the I_{CaL} blockade. (H) Superimposed action potential traces (top) and EGMs (bottom) before and after the administration of nifedipine in NC cells. All traces were derived from the same electrode, shown to the left. All bar charts are mean \pm SEM. * $p < 0.05$.

4.3.4 I_{Kr} -related EGM and action potential morphology modifications

I_{Kr} blockade was carried out using E-4031. The IC_{50} of E-4031 (303.3nM) was calculated as the average of IC_{50} concentrations for 1Hz (332.4nM) and 4Hz (274.1nM) data, which were the lowest and highest stimulation rates respectively at preliminary experiments. Following manual analysis of APD_{90} and FPD, it was concluded that both I_{Kr} blockade and the pacing cycle length increased APD_{90} (control: 1Hz 160.69±32.5 ms, 5Hz 125.74±2.8 ms, n=8 arrays; E-4031: 1Hz 216.34±47.8 ms, 4Hz 136.28±0.7 ms, n=4 arrays; $p<0.0001$ for both factors; mean ± SEM) and FPD (control: 1Hz 141.34±28.5 ms, 4Hz 113.37±17.9 ms, n=8; E-4031: 1Hz 206.4±28.8 ms, 4Hz 127.38±12.5 ms, n=4; $p<0.0001$; mean ± SEM) (Figure 4.8B-C). Linear regression analysis showed a correlation between APD_{90} and FPD ($r^2=0.71$, $Sy.x=21.6$, slope = 0.76, $p<0.0001$, n=83) and this relationship was stronger after the administration of E-4031 (Control: n=16, $r^2=0.38$, $Sy.x=20$, slope = 0.49±0.17, $p=0.01$; E-4031: n=67, $r^2=0.79$, $Sy.x=18.9$, slope=0.77±0.05, $p<0.0001$) (Figure 4.8D). The Bland-Altman plot revealed that FPD was systematically 37ms longer on average after I_{Kr} blockade (Figure 4.8E). The differences did not seem to be affected by the magnitude of average FPD and approximately 6% of the differences were out of the limits of agreement for FPD.

The dose response curve corresponding to NC culture EGM data (Figure 4.9A) was used for the extraction of the IC_{50} of E-4031 (49.6nM). The increased amount of fibroblasts in this type of cell culture compared to the myocyte only culture meant that there should theoretically be a smaller I_{Kr} current in those cell cultures and thus, the effect of E-4031 would be different. That dose response curve was based on the manual measurement of FPD in 1Hz data. NC cell cultures could be stimulated up to 2Hz, therefore the limited stimulation range prohibited the creation of restitution curves. There was a trend towards APD_{90} prolongation following the addition of E-4031 (control: 1Hz 226.6±3.5ms, E-4031: 1Hz 247.9±7.4ms, n=2 arrays, $p=0.16$; mean ± SEM) (Figure 4.9B). A significant increase in FPD was also seen at 1Hz (control: 1Hz 215.5±5ms, E-4031: 1Hz 269.8±6.8ms, n=2, $p=0.016$; mean ± SEM) (Figure 4.9C). Linear regression analysis showed a direct relationship between APD_{90} and FPD ($r^2=0.52$, $Sy.x=22.2$, slope=1.25±0.3, $p=0.0007$, n=18) (Figure 4.9D). FPD was systematically 55ms longer on average after I_{Kr} blockade, as the Bland-Altman plot shows in Figure 4.9E. Unclear conclusions are made about the effect of the magnitude of average FPD on differences between E-4031 and baseline data, due to the limited amount of data.

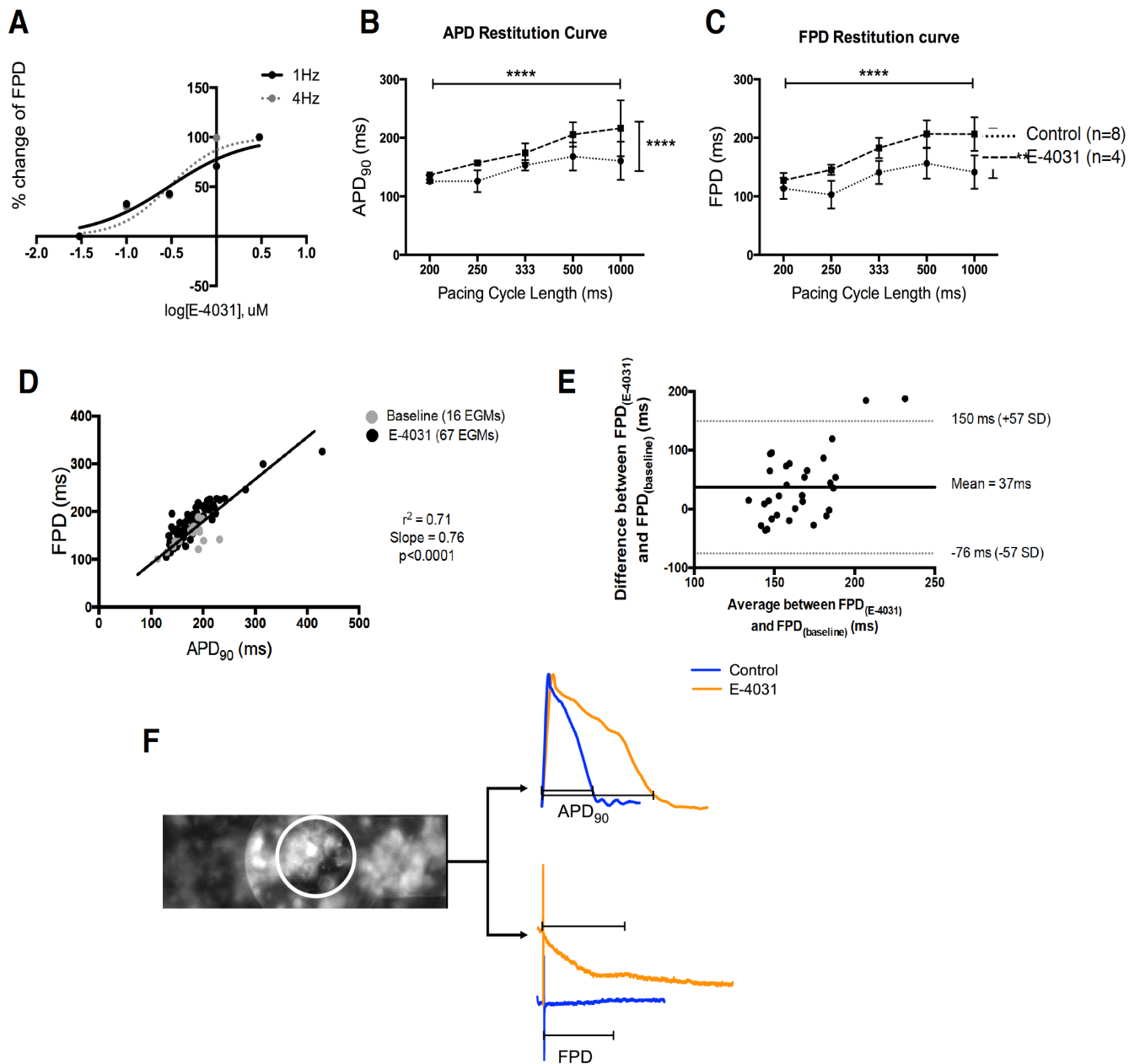


Figure 4.8: Effects of I_{K_r} current blockade using 303.3nM E-4031 on action potential duration and FPD. APD_{90} (B) and FPD (C) restitution curves (control: $n=8$ MEA plates; E-4031: $n=4$ MEAs). APD_{90} and FPD manually annotated data showed that both are affected by pacing rate and the E-4031 activity (all $p<0.0001$, ordinary two-way ANOVA). (D) Linear regression analysis showing the linear relationship between APD_{90} and FPD ($r^2=0.71$, $p<0.0001$). (E) Bland-Altman plot for FPD at baseline and following I_{K_r} blockade in NRVM cell monolayers. The middle bold line is the average difference between the baseline and I_{K_r} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Superimposed raw traces (top) and EGMs (bottom) obtained before and after I_{K_r} blockade from the same electrode to the left. Mean \pm SEM; **** $p<0.0001$.

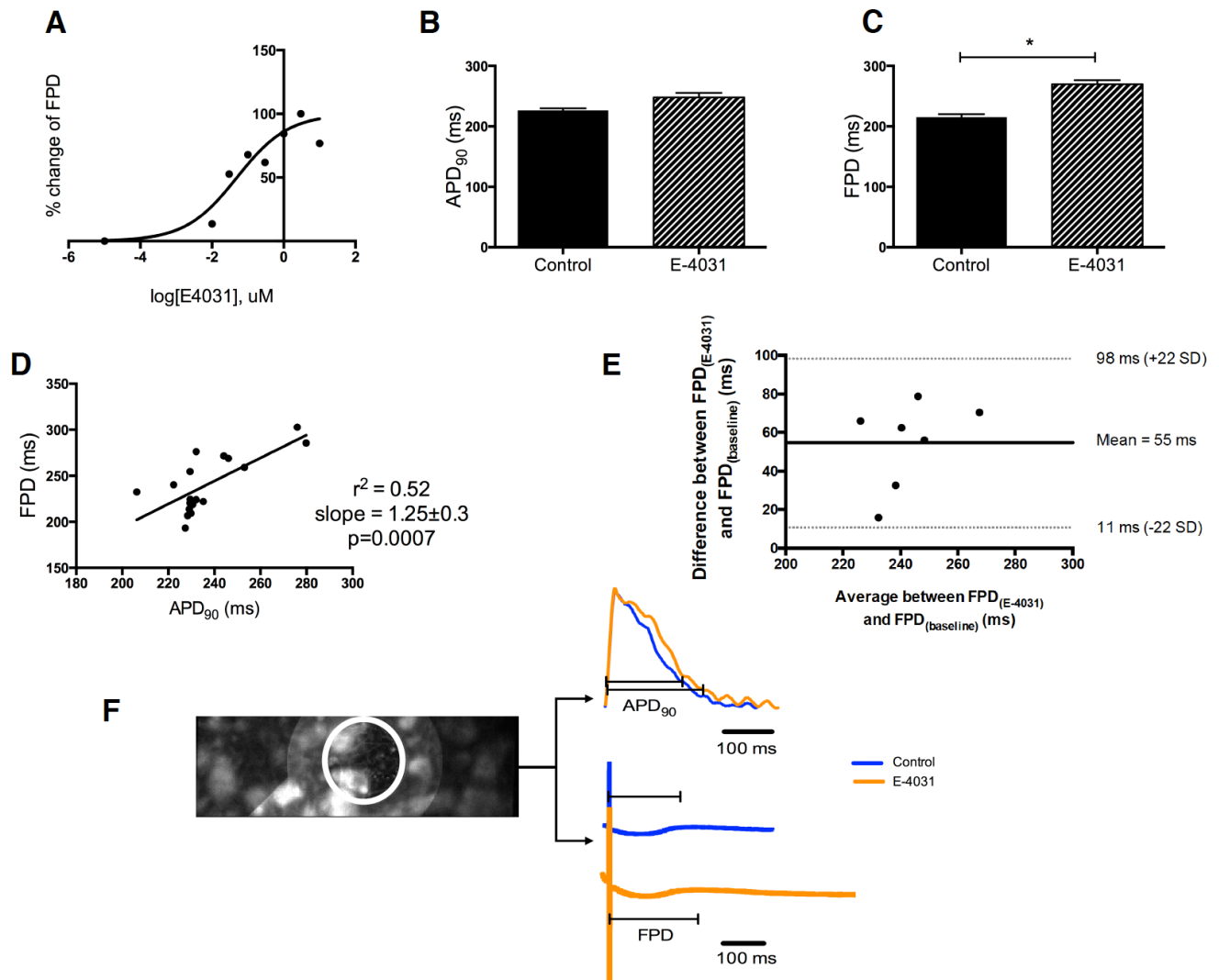


Figure 4.9: (A) Dose response curve for the effect of a range of E-4031 concentrations on the FPD obtained from NC cell cultures. (B) No significant change of APD₉₀ was caused by 4-AP ($p=0.16$). (C) I_{Kr} blockade increased significantly the FPD ($p=0.02$). (D) Linear regression analysis showed FPD and APD₉₀ trended towards a linear relationship between them ($r^2=0.52$, $Sy.x=22.2$, slope= 1.25 ± 0.3 , $p=0.0007$, $n=18$). (E) Bland-Altman plot for FPD at baseline and following I_{Kr} blockade in NC cells. The middle bold line is the average difference between the baseline and I_{Kr} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Superimposed action potentials (top) and EGMs (bottom) obtained simultaneously before and after the I_{Kr} blockade. Mean \pm SEM. APD₉₀ and FPD plots are from manually annotated data. * $p<0.05$.

4.3.5 The effects of I_{Ks} blockade on EGM and action potential morphology

HMR-1556, which is a I_{Ks} blocker, was used for its effect on EGM morphology. The calculation of its IC₅₀ (450.5nM) was based on the effect on FPD, even though there was no clear FPD prolongation (Figure 4.10A). However, a trend towards electrical restitution was observed, as a result of the effect of

pacing frequency both on APD₉₀ (control: 1Hz 175±11.7ms, 3Hz 144±0.5ms; HMR-1556: 1Hz 194±20.7ms, 3Hz 148.3; n = 2 arrays; p=0.06; mean ± SEM) and FPD (control: 1Hz 178.5±17.6ms, 3Hz 143.5±14.4ms; HMR-1556: 1Hz 196.5±24.7, 3Hz 150ms; n = 2 arrays; p = 0.16; mean ± SEM) (Figure 4.10B-C). There was no significant difference due to the HMR-1556 activity (APD₉₀: p = 0.2; FPD: p = 0.45). But the linear regression analysis showed a correlation between APD₉₀ and FPD ($r^2=0.89$, $Sy.x=8.1$, slope = $1.04±0.07$, $p<0.0001$, $n=30$) (Figure 4.10E). This correlation was also present before and after the I_{Ks} blockade (p = 0.14). The Bland-Altman plot revealed that FPD was systematically only 19ms longer on average after I_{Ks} blockade with HMR-1556 (Figure 4.10F). The differences did not seem to be affected by the magnitude of average FPD and none of the differences were out of the limits of agreement for FPD.

The IC₅₀ of HMR-1556, which was used in dual modality experiments on NC cell cultures, was 740nM. This was based on changes in the FPD over a range of HMR-1556 concentrations (10nM – 100µM) (Figure 4.11A). However, no significant difference in APD₉₀ (1Hz control: 225.7±11.5, HMR-1556: 223.4±8.2ms, n = 2 MEAs, p = 0.96; mean ± SEM) and the FPD (1Hz control: 212.2±10ms, HMR-1556: 212.5±8ms, n = 2, p = 0.87; mean ± SEM) was observed after the I_{Ks} blockade using the IC₅₀ of HMR-1556 (Figure 4.11C-D). The linear regression analysis showed direct correlation between APD₉₀ and FPD ($r^2 = 0.76$, $Sy.x=15.4$, slope = $0.81±0.09$, $p<0.0001$, $n=26$), which was the same before and after the administration of HMR-1556 (p=0.73). Conduction velocity was maintained stable over the range of drug doses (n = 3 MEAs, p = 0.53). FPD was systematically only 13ms longer on average after I_{Ks} blockade with HMR-1556 in NC cell monolayers (Figure 4.11F).

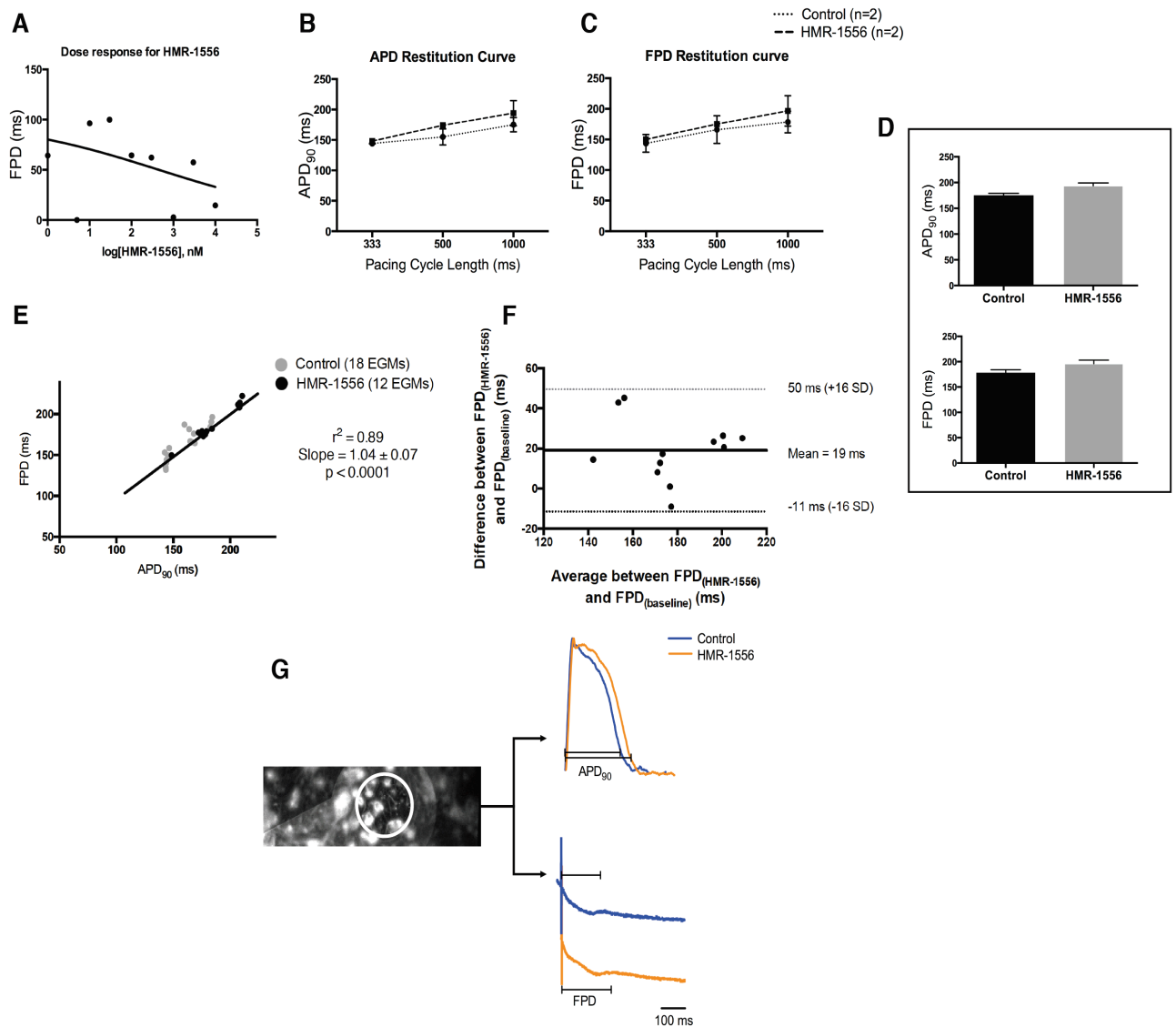


Figure 4.10: Correlation between APD₉₀ and FPD in NRVM cultures after I_{Ks} blockade using HMR-1556. APD₉₀ (B) and FPD (C) restitution curves show that neither pacing rate nor I_{Ks} blockade have a significant effect on APD₉₀ and FPD. (D) No significant change of APD₉₀ and FPD was observed when NRVMs were treated with IC₅₀ concentration of HMR-1556 (control: 6 EGMs; HMR-1556: 8 EGMs) followed by pacing at 1Hz. (E) Linear relationship of APD₉₀ and FPD ($r^2=0.89$, $Sy.x=8.1$, $p<0.0001$). (F) Bland-Altman plot for FPD at baseline and following I_{Ks} blockade in NRVM cell monolayers. The middle bold line is the average difference between the baseline and I_{Ks} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (G) Superimposed raw traces (top) and EGMs (bottom) obtained from the same electrode (left), showing the changes at APD₉₀ and FPD. Mean ± SEM. APD₉₀ and FPD plots were made using manually annotated data.

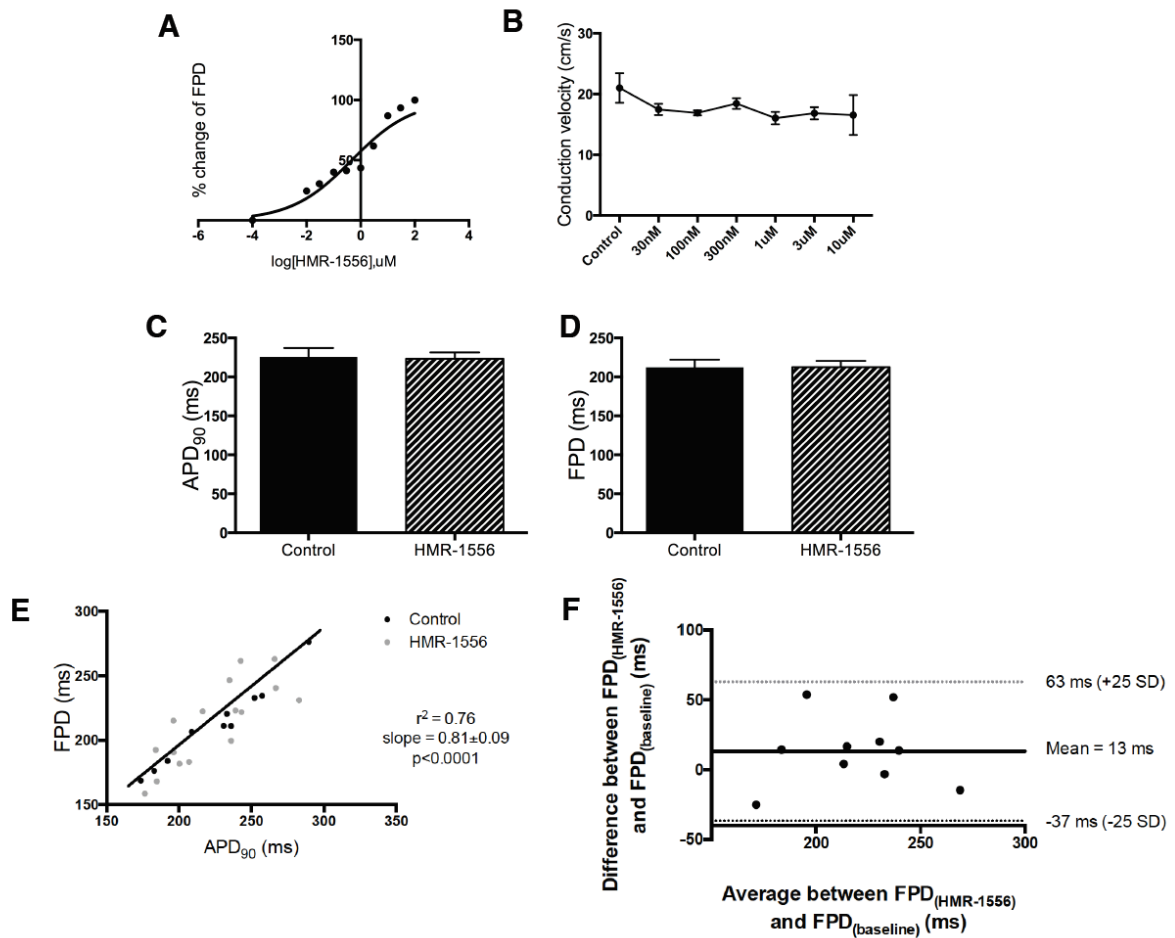


Figure 4.11: (A) Dose response curve for the calculation of the IC_{50} of HMR-1556 in NC monolayers based on the effect of I_{Ks} blockade on FPD. (B) Conduction velocity of NC cell cultures remained stable throughout the administration of a range of HMR-1556 doses ($n = 3$ arrays). (C-D) APD_{90} (C) and FPD (D) did not change following the I_{Ks} blockade using HMR-1556 (IC_{50}) on NC cell cultures ($n = 2$ MEAs). (E) Linear regression analysis for the relationship between APD_{90} and FPD before and after the administration of HMR-1556. (F) Bland-Altman plot for FPD at baseline and following I_{Ks} blockade in NC cell monolayers. The middle bold line is the average difference between the baseline and I_{Ks} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. Mean \pm SEM. APD_{90} and FPD plots were made using manually annotated data.

4.3.6 EGM and action potential morphology modifications following K_{ATP} channel opening

The calculation of IC_{50} of pinacidil ($1.98 \mu\text{M}$) was based on the drug effect on FPD (Figure 4.12A). The data obtained from concurrent optical mapping and MEA recordings showed that APD_{90} was significantly reduced due to pinacidil (control: $220.5 \pm 16 \text{ms}$, pinacidil: $164.2 \pm 8 \text{ms}$; $n = 2$ MEAs; $p =$

0.04; mean \pm SEM) (Figure 4.12B). FPD was also shortened (control: 223.6 ± 13 ms; pinacidil: 166.1 ± 9.5 ms; $n = 2$ MEAs; $p = 0.005$; mean \pm SEM) (Figure 4.12C). Electrical restitution could not be studied, due to the limited ability to capture EGMs at higher frequencies than 2Hz. A direct correlation between APD_{90} and FPD was observed ($r^2 = 0.84$, $Sy.x=11.2$, slope = 0.72, $p<0.0001$, $n=44$) (Figure 4.12D). The Bland-Altman plot revealed that FPD was systematically 61ms shorter on average after K_{ATP} channel blockade with pinacidil in NRVM monolayers (Figure 4.12E).

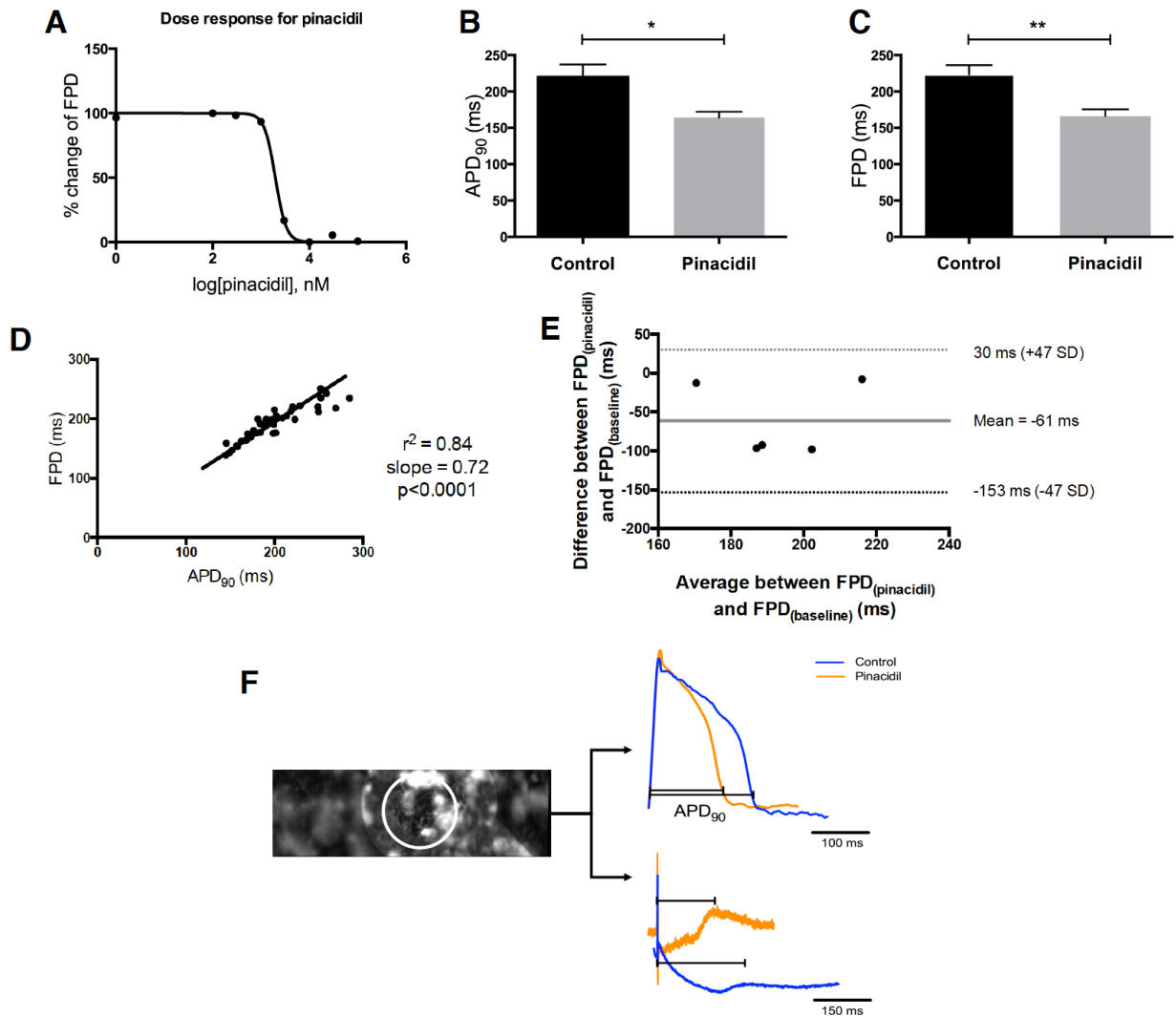


Figure 4.12: Effect of K^+ current enhancement using pinacidil on AP duration and FPD. Both APD_{90} (B) and FPD (C) decreased significantly after the administration of $30\mu M$ pinacidil on NRVM monolayers ($p=0.04$ and 0.005 respectively). (D) Linear regression analysis for the relationship between APD_{90} and FPD ($r^2=0.84$; $Sy.x=11.2$; $p<0.0001$). (E) Bland-Altman plot for FPD at baseline and following K_{ATP} channel opening in NRVM cell monolayers. The middle bold line is the average difference between the baseline and I_{Ks} blockade. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Superimposed raw traces (top) and EGMs (bottom) showing that both APD_{90} and FPD are shortened after the administration of pinacidil. Mean \pm SEM; * $p<0.05$; ** $p<0.01$. APD_{90} and FPD plots were made using manually annotated data.

4.3.7 EGM and action potential morphology manifestations of double ion channel blockade

The simultaneous Na⁺ and L-type Ca²⁺ channel blockade was carried out by concurrently treating NRVMs with lidocaine and nifedipine. A range of dose combinations were used based on the IC₅₀ of each ion channel blocker (Table 4.1). Following the administration of “combination 4”, the double ion channel blockade had a disrupting effect on cells by initially inducing fibrillation. Even higher doses managed to immediately interrupt the electrophysiological activity, and therefore no high enough concentrations could be used, comparing to the individual use of lidocaine or nifedipine. Dose response curve for each ion channel blocker indicated concentrations close to the ones from “combination 3”, and therefore that one was used for dual modality experiments. Conduction velocity was not affected by any drug combination (p=0.81), but EGM duration was affected by increasing doses (control: 2.39±0.01ms, drug comb.4: 2.5±0.2ms, n=3 arrays, p<0.001; mean ± SEM) (Figure 4.13A-B). While APD₉₀ showed a trend towards reduction following the ion channel blockade (control: 1Hz 195.2±30.5ms, ion channel blockade: 1Hz 178.2±23ms, n=5 APs, p=0.15; mean ± SEM) (Figure 4.13C), FPD was significantly abbreviated (control: 1Hz 229±17.5ms, ion channel blockade: 1Hz 188.5±24ms, n=5 EGMs, p<0.05; mean ± SEM) (Figure 4.13D). Linear regression analysis showed a strong correlation between APD₉₀ and FPD (r²=0.84, Sy.x=12.7, slope = 1.01±0.16, p=0.0002, n=10) (Figure 4.13F).

Table 4.1: Combination of lidocaine/nifedipine concentrations used for double ion channel blockade treatment of myocyte only NRVM cultures. The concentration of each drug was based on its IC₅₀.

		Lidocaine (µM)	Nifedipine (µM)
Combination 1	(0.75x IC ₅₀)	20.25	0.58
Combination 2	(IC ₅₀)	27	0.77
Combination 3	(1.5x IC ₅₀)	40.5	1.16
Combination 4	(2x IC ₅₀)	54	1.54

A range of combinations of lidocaine and nifedipine were used for double ion channel blockade in NC cell monolayers (Table 4.2). These concentrations were based on the IC₅₀ of each drug. Separate dose response curves for lidocaine and nifedipine were made for the extraction of IC₅₀ (lidocaine: 904.5nM, nifedipine: 96nM). The dose response curve for lidocaine was based on EGM duration data (Figure 4.14A), since conduction velocity remained stable throughout the range of ion channel blockers that was tested on NC cultures (p = 0.52) (Figure 4.14C). The dose response curve for nifedipine was based on the FPD (Figure 4.36B).

The final combination of drug doses, which was also used for the dual modality experiments was combination 7, according to Table 4.2 (lidocaine: 740nM, nifedipine: 130nM), since that was between the drug combinations which were closer to the new IC₅₀. Doses higher than 5µM lidocaine and 300nM nifedipine prohibited optical mapping and EGM recordings due to the absence of electrophysiological activity. Following the double ion channel blockade by lidocaine and nifedipine, no significant changes were observed on APD₉₀ (1Hz control: 281.4±1.9ms, ion channel blockade: 247.9±16.4ms, n = 1 array, p = 0.06), FPD (1Hz control: 249.4±7ms, ion channel blockade: 241.5±9.9ms, n=1, p = 0.63), upstroke (1Hz control: 16.58±0.1ms, ion channel blockade: 15.72±0.2ms, n=1, p = 0.06) and EGM duration (1Hz control: 3.15±0.1ms, ion channel blockade: 3.4±0.2ms, n=1, p = 0.4) (Figure 4.14D-G). There was also no linear relationship between APD₉₀ and FPD ($r^2 = 0.39$, slope = 0.33±0.18, p = 0.13).

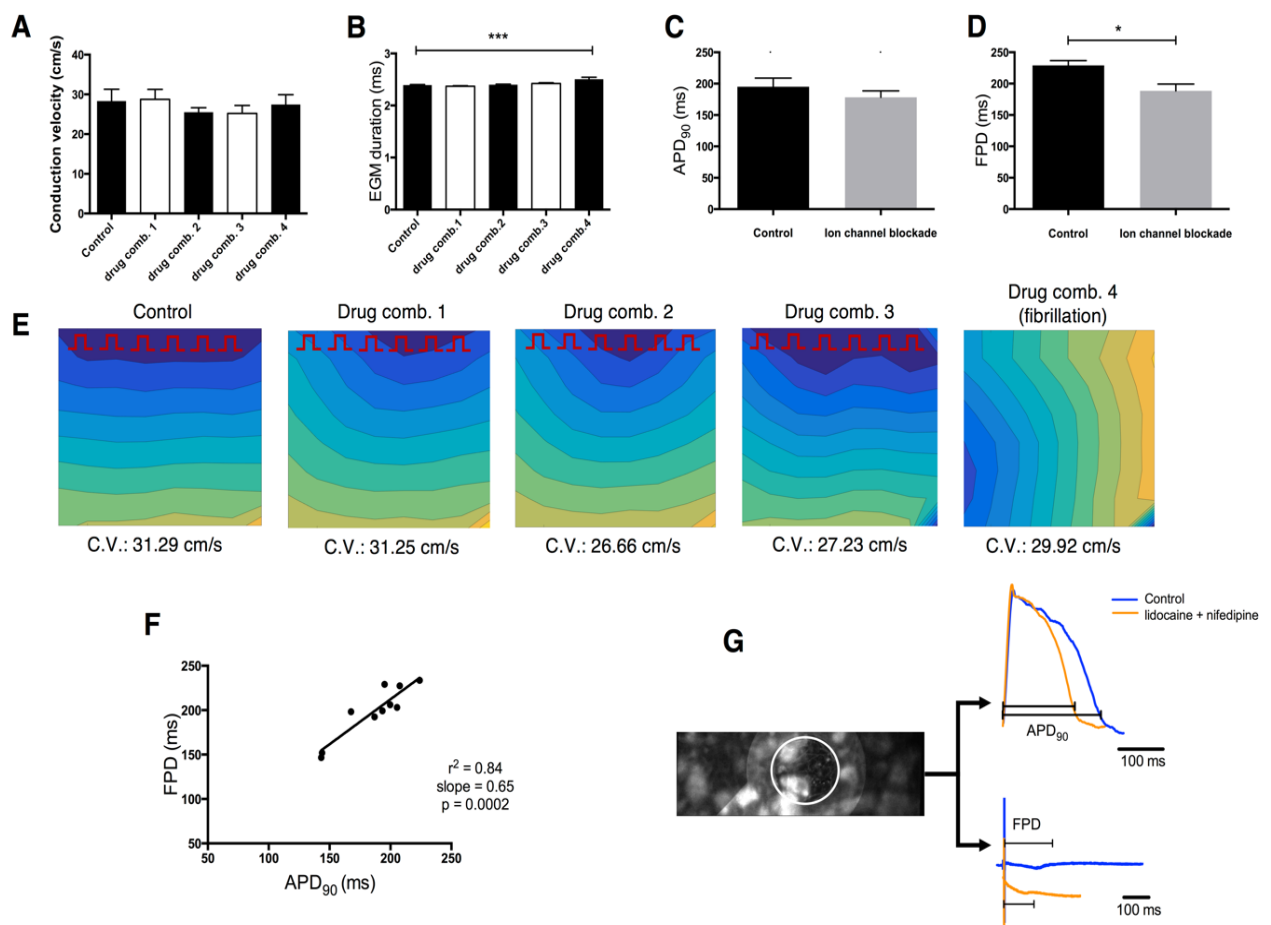


Figure 4.13: (A) Conduction velocity remains stable and is independent of any combination of lidocaine/nifedipine concentrations (p=0.81). (B) EGM duration increases due to the combined ion channel blockade (p=0.0007). (C) APD₉₀ does not change significantly (p=0.15), when NRVM cultures are treated with 40.5µM lidocaine and 1.16µM nifedipine, but FPD (D) is reduced after the ion channel blockade (p=0.16). APD₉₀ and FPD plots were made using manually annotated data. (E) Isochronal maps presenting the activation wavefront propagation and the conduction velocity in each occasion. The last isochronal map (right) presents a different direction of the activation wavefront, due to the fibrillation that occurs when cells are treated simultaneously with high doses of lidocaine (54µM) and nifedipine (1.54µM). (F) Linear regression analysis for the relationship between FPD and APD₉₀ before and after the double ion channel blockade. (G) Superimposed raw traces (top) and EGMs (bottom) obtained simultaneously from the same electrode before and after the combined administration of lidocaine and nifedipine. Mean ± SEM; *p<0.05; ***p<0.001.

Table 4.2: Combination of lidocaine/nifedipine concentrations used for double ion channel blockade treatment of myocyte only NRVM cultures. The concentration of each drug was based on its IC₅₀ from experiments on NC cell cultures using individual ion channel blockers.

		Lidocaine (nM)	Nifedipine (nM)
Combination 1	0.25x IC ₅₀	92.5	16.25
Combination 2	0.5x IC ₅₀	185	32.5
Combination 3	0.75x IC ₅₀	277.5	50.25
Combination 4	IC ₅₀	370	65
Combination 5	1.25x IC ₅₀	462.5	81.25
Combination 6	1.5x IC ₅₀	555	97.5
Combination 7	2x IC ₅₀	740	130
Combination 8	3x IC ₅₀	1,110	195
Combination 9	-	5,000	300

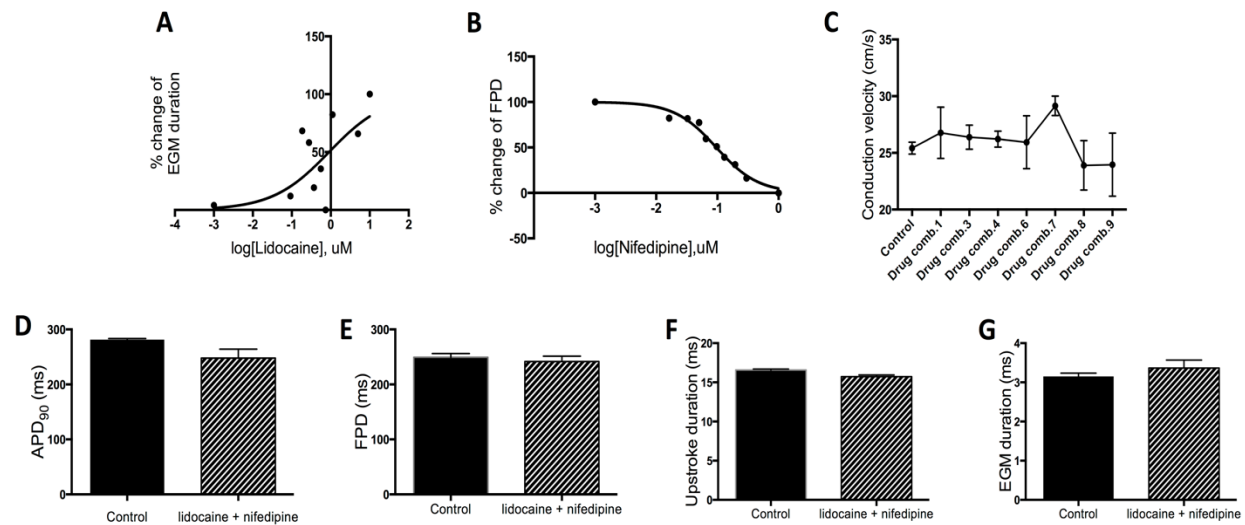


Figure 4.14: (A) Dose response for the calculation of IC₅₀ of lidocaine, based on the effect of Na⁺ channel blockade on EGM duration, when this is administered on NC cell cultures. (B) Dose response curve for the extraction of IC₅₀ of nifedipine based on the effect of I_{CaL} blockde on FPD. (C) Conduction velocity remained stable throughout the administration of any combination of lidocaine and nifedipine. (D-G) No significant change was seen in the APD₉₀ and upstroke duration in action potential morphology and the FPD and EGM duration in EGMs following the double ion channel blockade. Mean ± SEM.

4.3.8 Comparative analysis of EGM features across different *in vitro* models

The experiments presented so far in this chapter prove the efficiency of the pharmacological agents presented in section 4.2.2, in terms of modifying EGM morphology. Therefore, I proceeded into further investigating the specific changes. Comparing the effects of the various forms of ion channel modulations either in isolation or in combination with differential amounts of fibroblasts in NRVM cell cultures, it was observed that they caused significant changes to the majority of features in overall. The comparisons were made between groups of EGM collected from a number of MEA dishes used across experiments. Table 4.3 presents the overall number of EGMs used per experimental condition and the number of cell cultures used to obtain them. Control data for each type of cell culture were obtained from all experiments carried out on the same type of cell monolayer, even though control data were always recorded prior to the addition of a pharmacological agent. The data obtained before and after ion channel blockade or opening on the same cell group were compared, as presented in Figures 4.15-4.25.

Features were measured from EGMs obtained from the same type of cell culture. When comparing the modifications caused by the same modulation, it was observed that each feature was significantly different ($p < 0.0001$ in majority of cases). Exemptions to this were the S.D EGM amplitude measured in +20% FB and +40% FB cultures ($p = 0.03$ and $p = 0.002$ respectively; Fig.4.15), R-/S-width ratio in +20% FB and +40% FB cultures (+20%FB: $p = 0.12$ – non-significant; +40%FB: $p = 0.0002$; Fig.4.18), R-width/EGM duration in +20%FB ($p = 0.17$ – non-significant; Fig.4.19), peak ratio in +20%FB ($p = 0.002$; Fig.4.19) and the S.D. autocorrelation measured from +20%FB data ($p = 0.003$; Fig.4.25). In addition, comparisons of the functional and structural modulations in overall for the same feature showed that all feature modifications were significantly different among modulations ($p < 0.0001$). A detailed presentation of all changes happening under each experimental condition individually, and not comparing to other conditions as in this chapter, is presented in Appendix A (sections I-VII).

Table 4.3. Number of EGMs collected per experimental condition from a number of arrays (cell cultures) under control conditions and following the administration of each pharmacological agent using the range of concentrations presented in section 4.2.2.

	# EGMs	# arrays		# EGMs	# arrays
Control - MO	347	28	KATP opening	125	6
I _{to} blockade	118	4	Control – NC	493	36
I _{Kr} blockade	112	5	NC+I _{to} blockade	138	6
I _{Na} blockade	149	8	NC+I _{Kr} blockade	126	6
I _{Ks} blockade	109	5	NC+I _{Ks} blockade	111	6
I _{Na} +I _{CaL} blockade	117	6	NC+I _{Na} blockade	139	7
I _{CaL} blockade	113	8	NC+I _{CaL} blockade	104	5
Control - +20%FB	268	28	NC+I _{Na} +I _{CaL} blockade	107	6
+20%FB + I _{to} blockade	150	8	+20%FB + I _{Kr} blockade	68	4
+20%FB + I _{Ks} blockade	51	3	+20%FB + I _{Na} blockade	131	6
+20%FB + I _{CaL} blockade	126	7	Control - +40%FB	221	12
+40%FB + I _{Kr} blockade	56	3	+40%FB + I _{Na} blockade	53	6
+40%FB + I _{CaL} blockade	75	3			

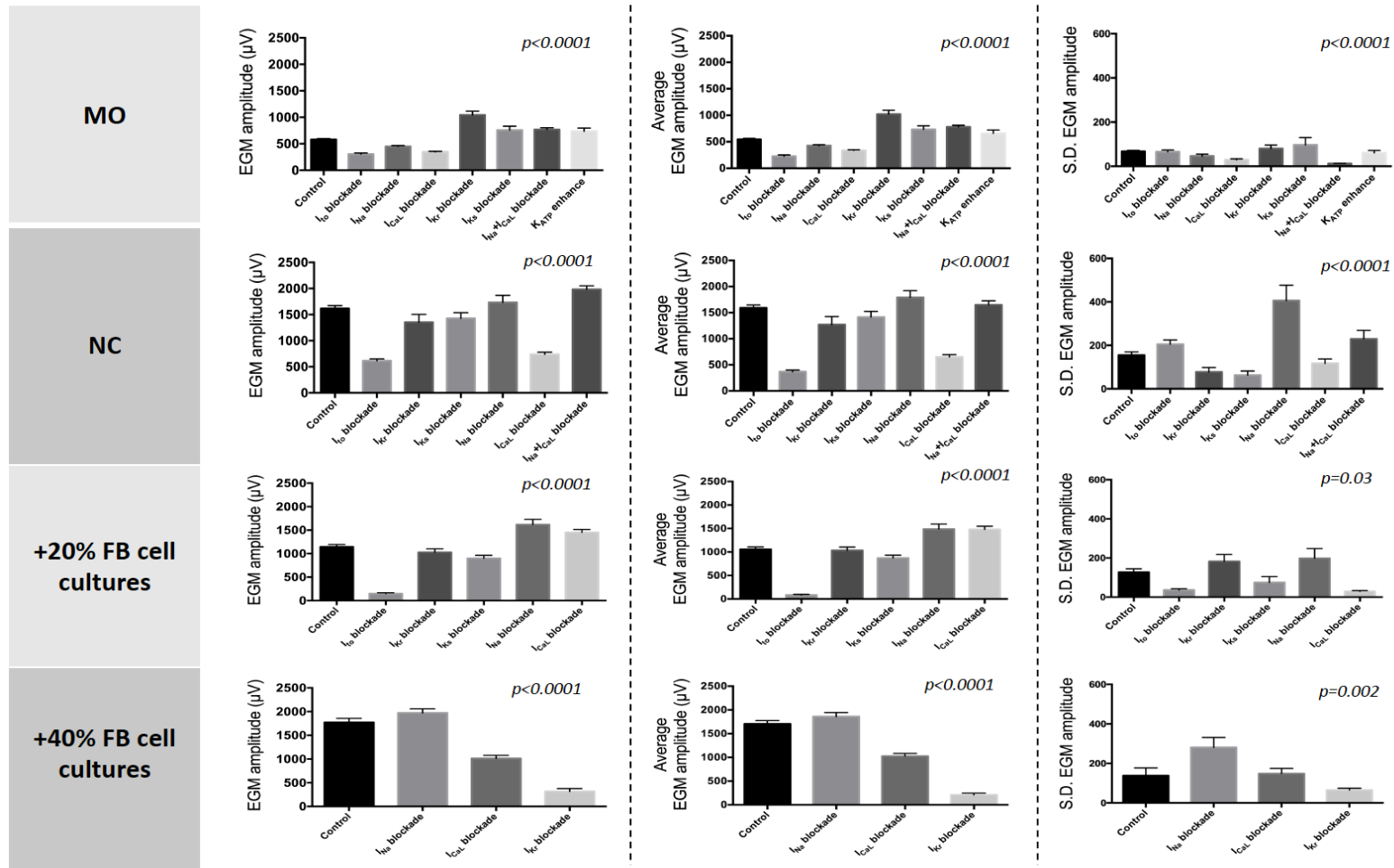


Figure 4.15. Bar charts presenting the modification of the EGM amplitude of the first EGM in a recording, the average EGM amplitude and the S.D. EGM amplitude due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

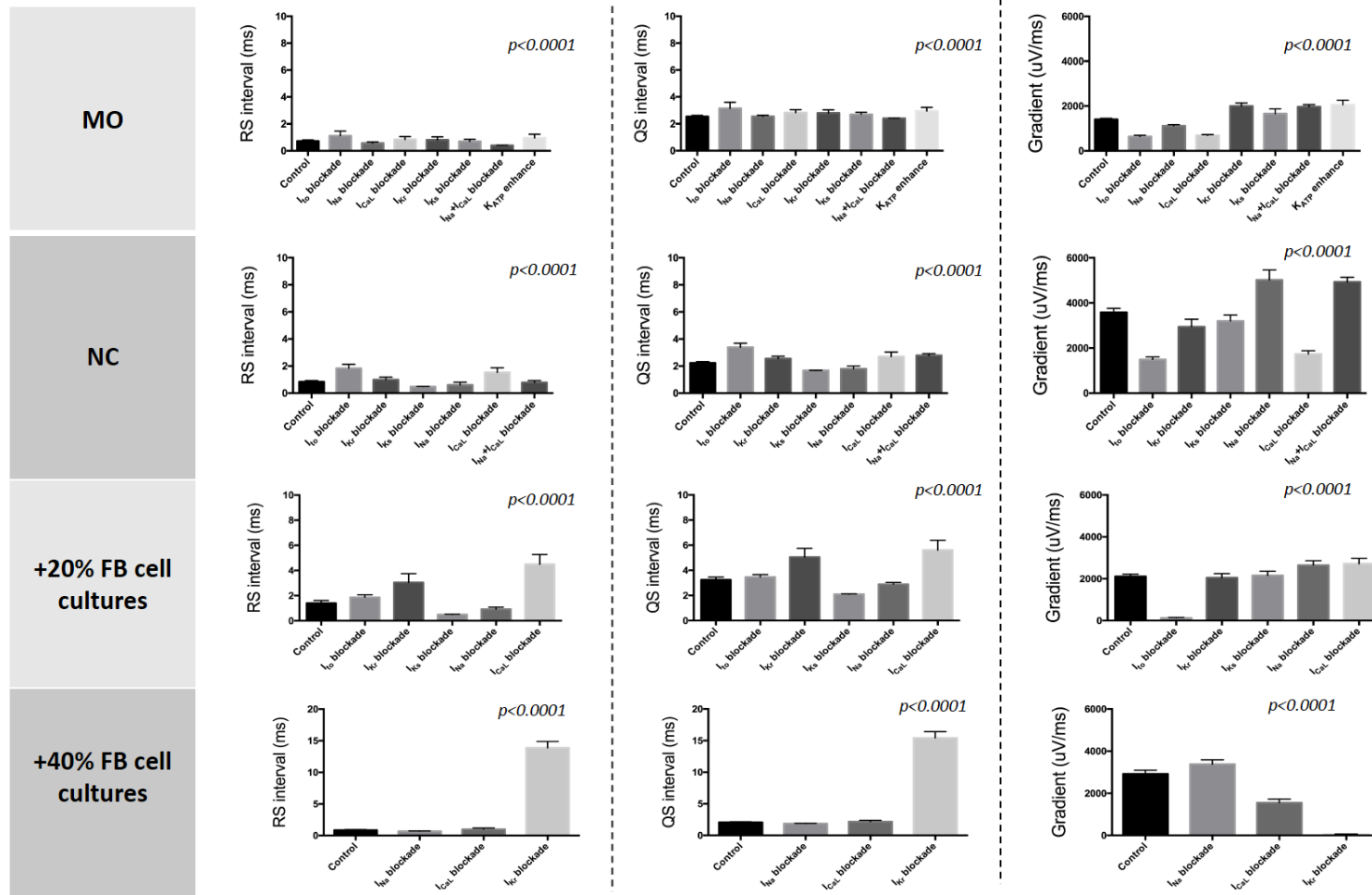


Figure 4.16. Bar charts presenting changes in RS interval, QS interval and the RS interval gradient, due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

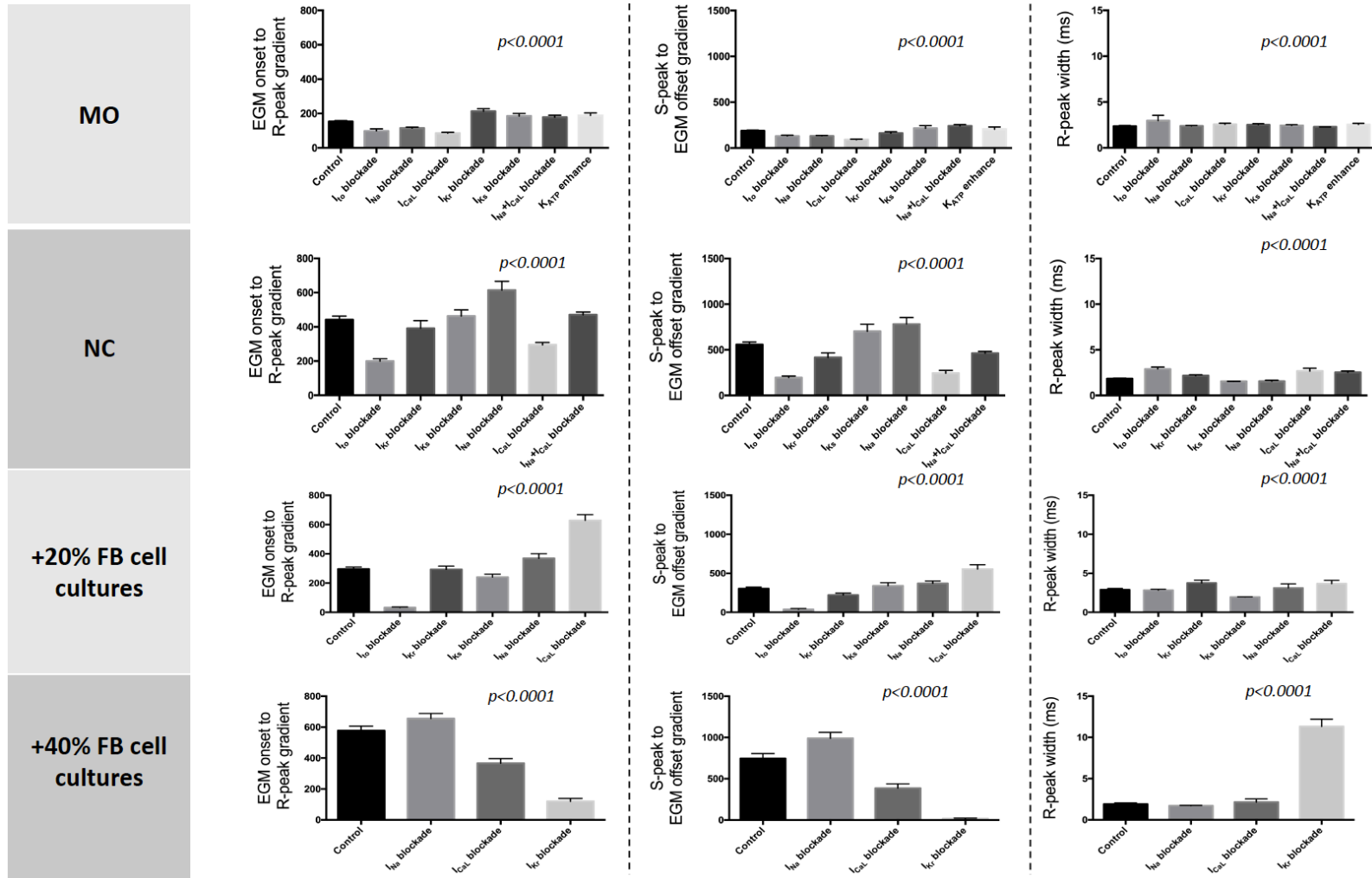


Figure 4.17. Modification of the EGM-onset-to-R peak gradient, S-peak-to-EGM offset gradient and R-peak width, due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

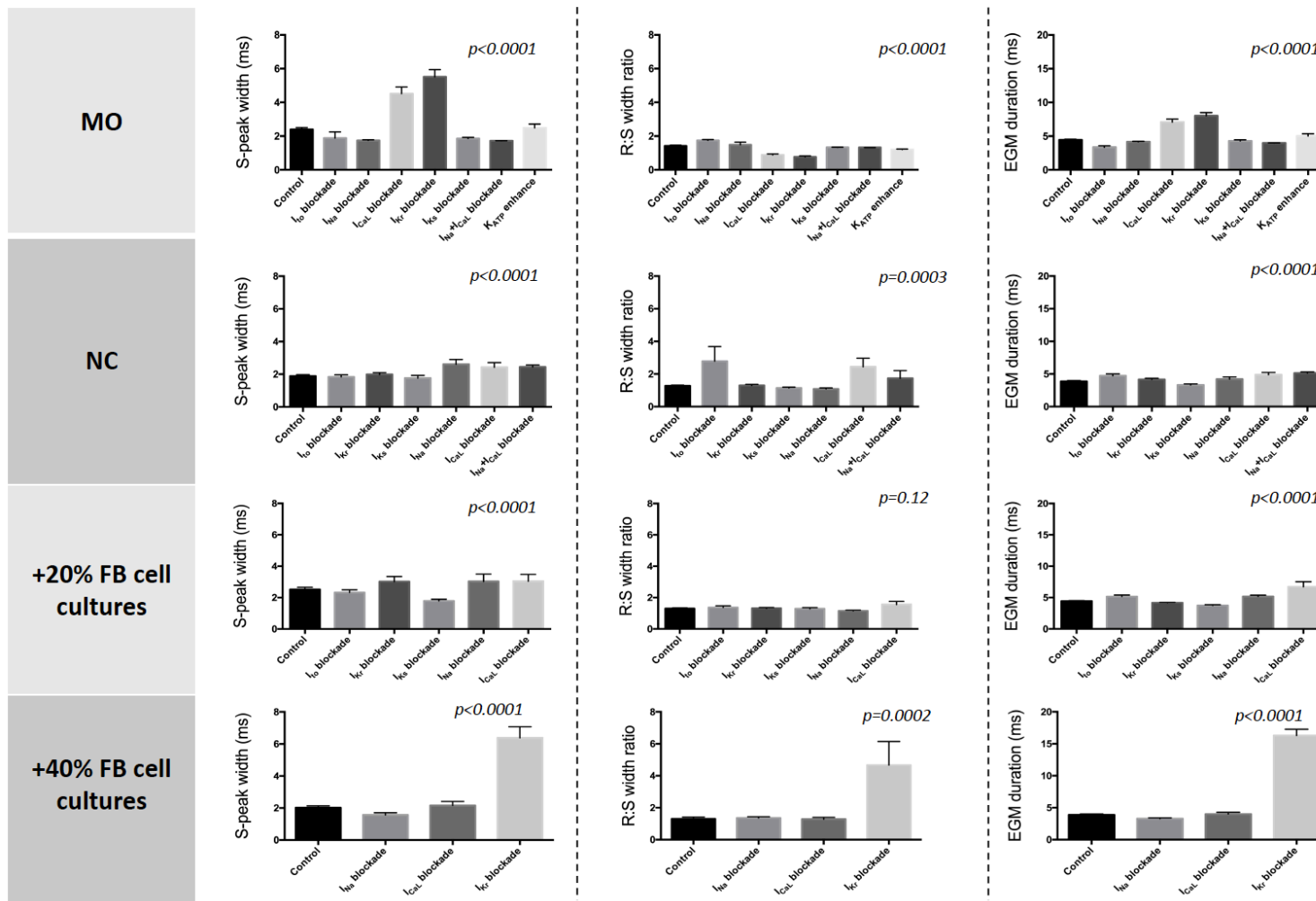


Figure 4.18. Bar charts presenting the modification of the S-peak width, R-/S- width ratio and EGM duration, due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

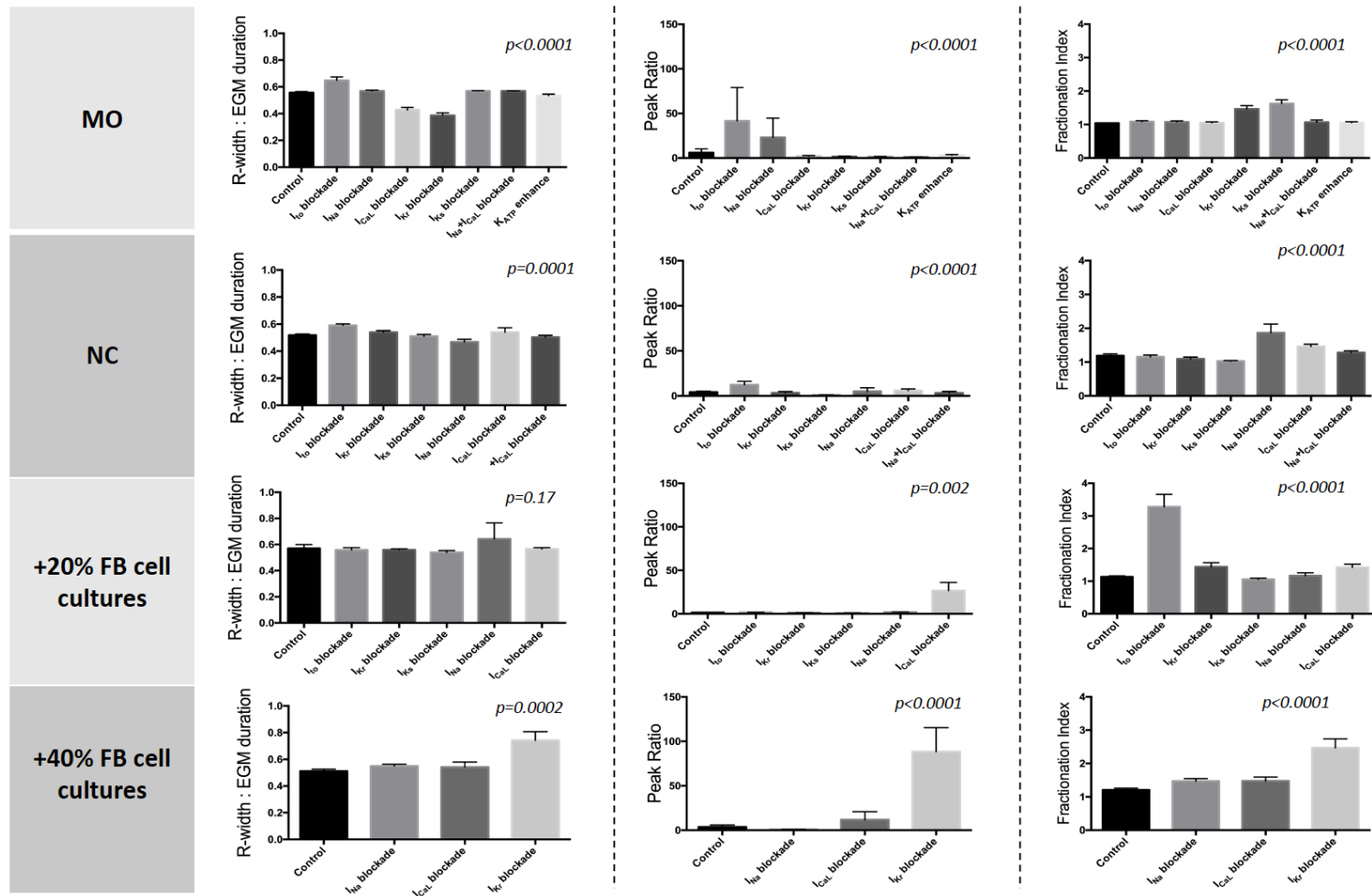


Figure 4.19. Bar charts presenting the modification of the R-width:EGM duration ratio, R/S-peak ratio and fractionation index due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

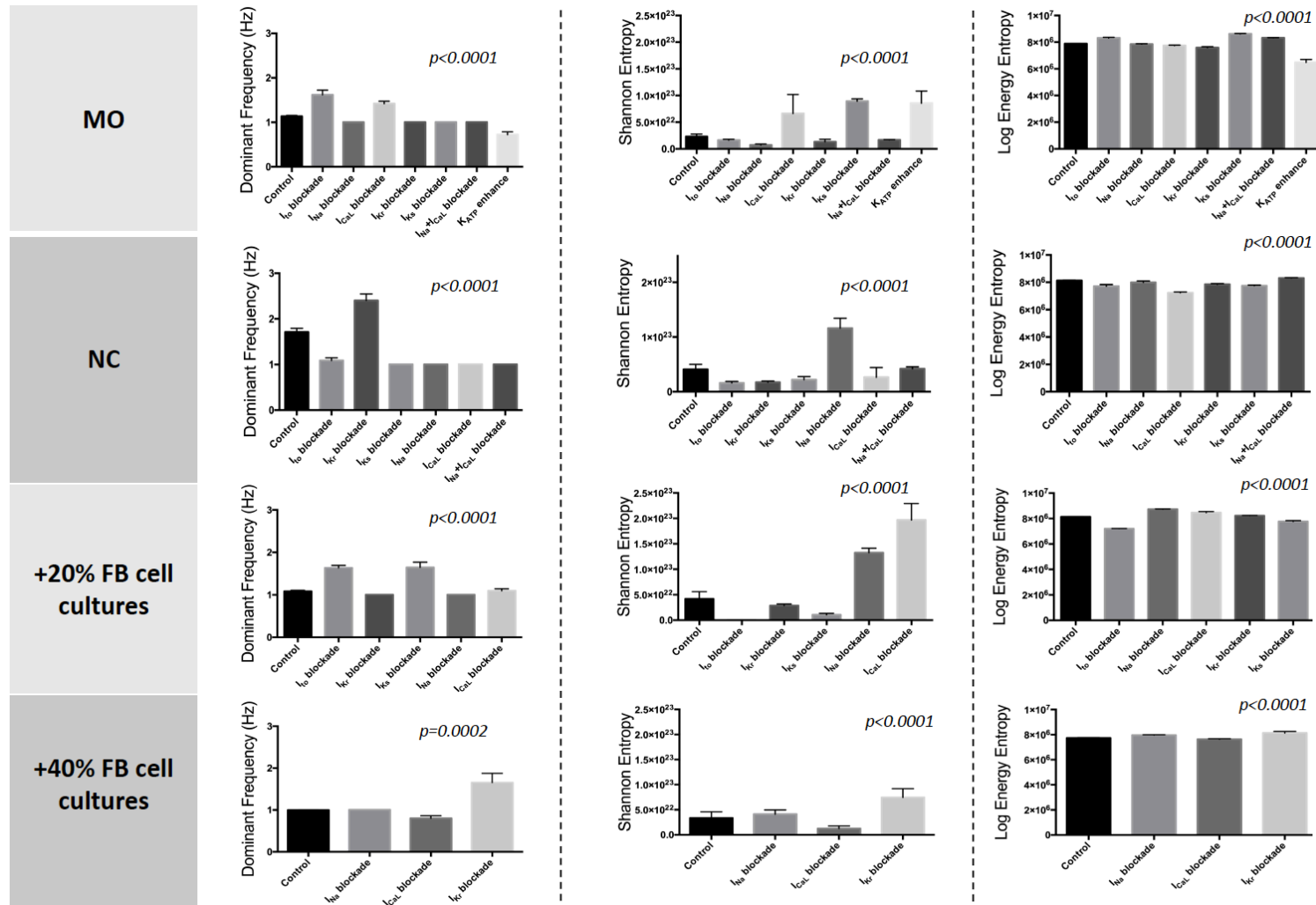


Figure 4.20. Bar charts presenting the modification of the dominant frequency, Shannon entropy and logarithmic energy entropy due to ion channel blockade in different types of NRVN cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

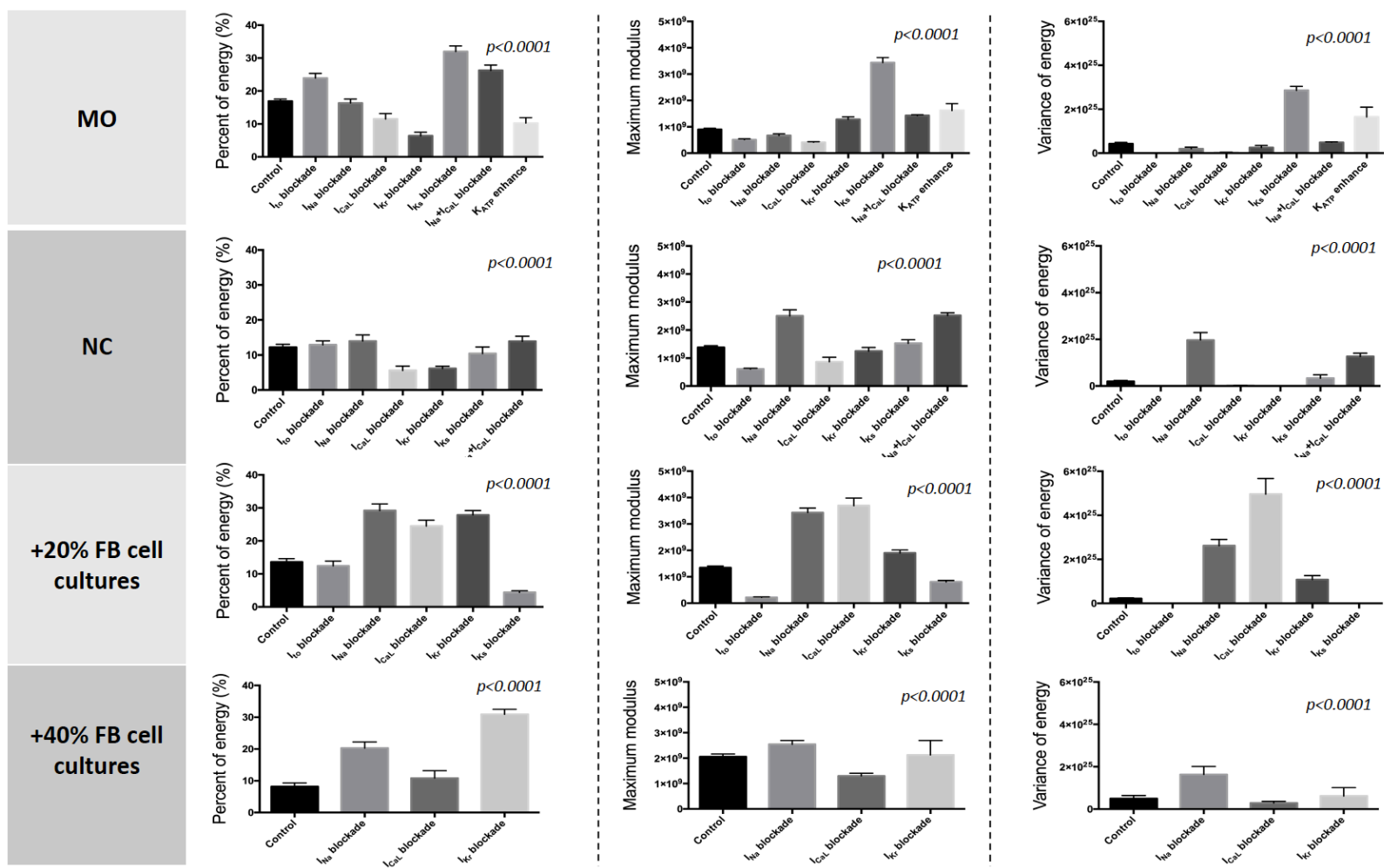


Figure 4.21. Bar charts presenting the modification of the percent of energy, maximum modulus and variance of energy due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

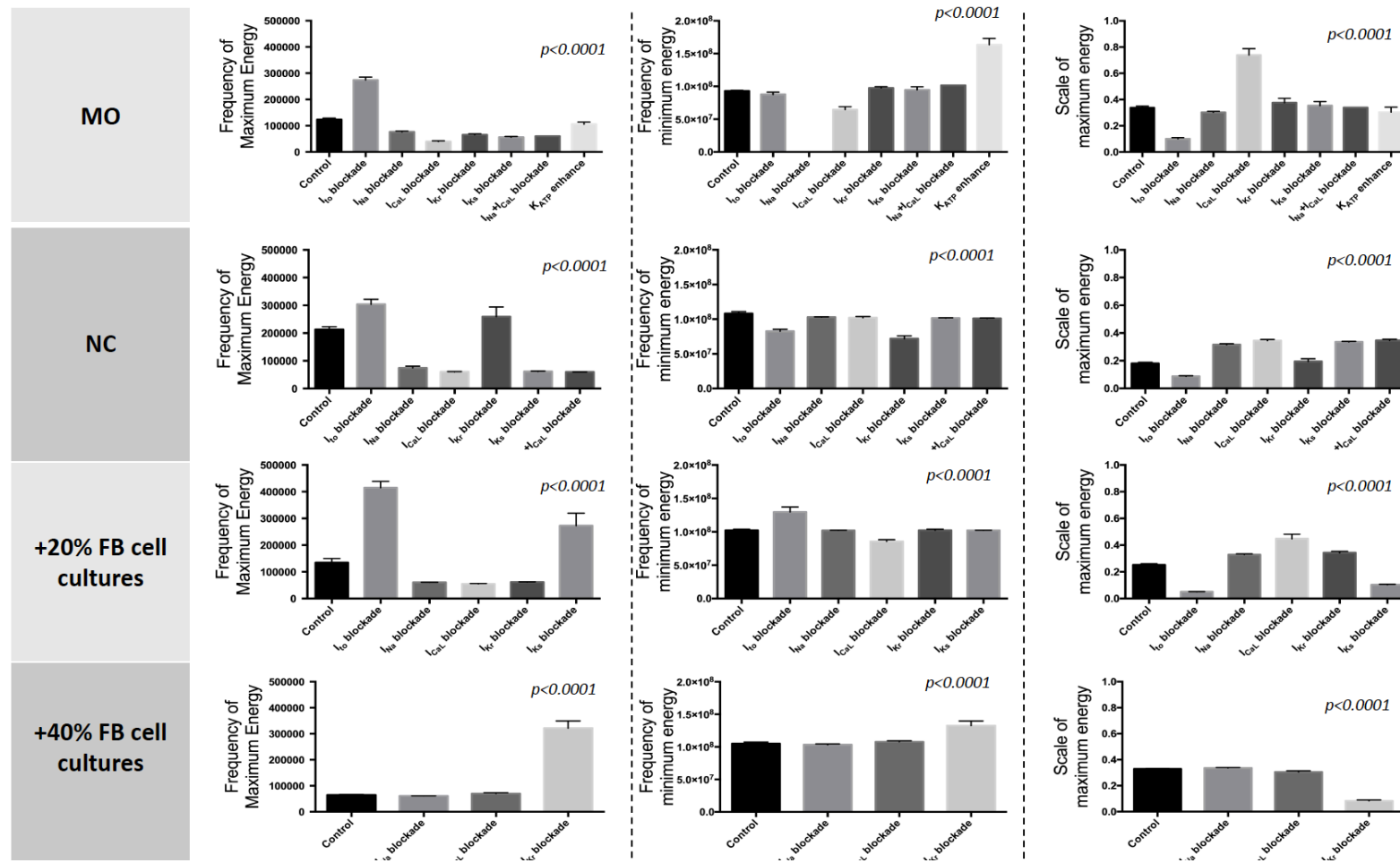


Figure 4.22. Bar charts presenting the modification of the frequency of maximum wavelet energy, frequency of minimum energy and the scale of maximum energy due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

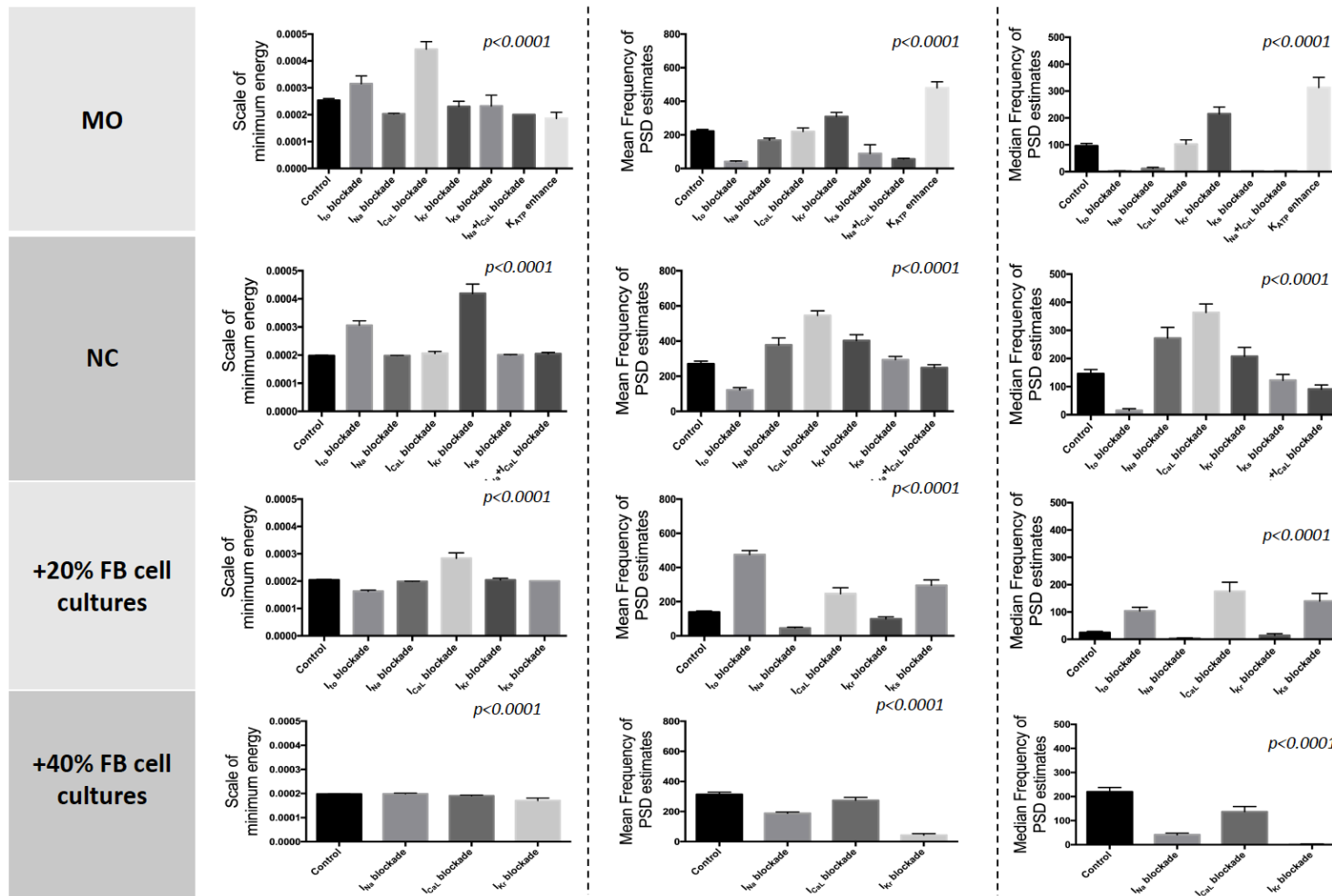


Figure 4.23. Bar charts presenting the modification of the scale of minimum energy, mean frequency of PSD estimates and the median frequency of PSD estimates due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

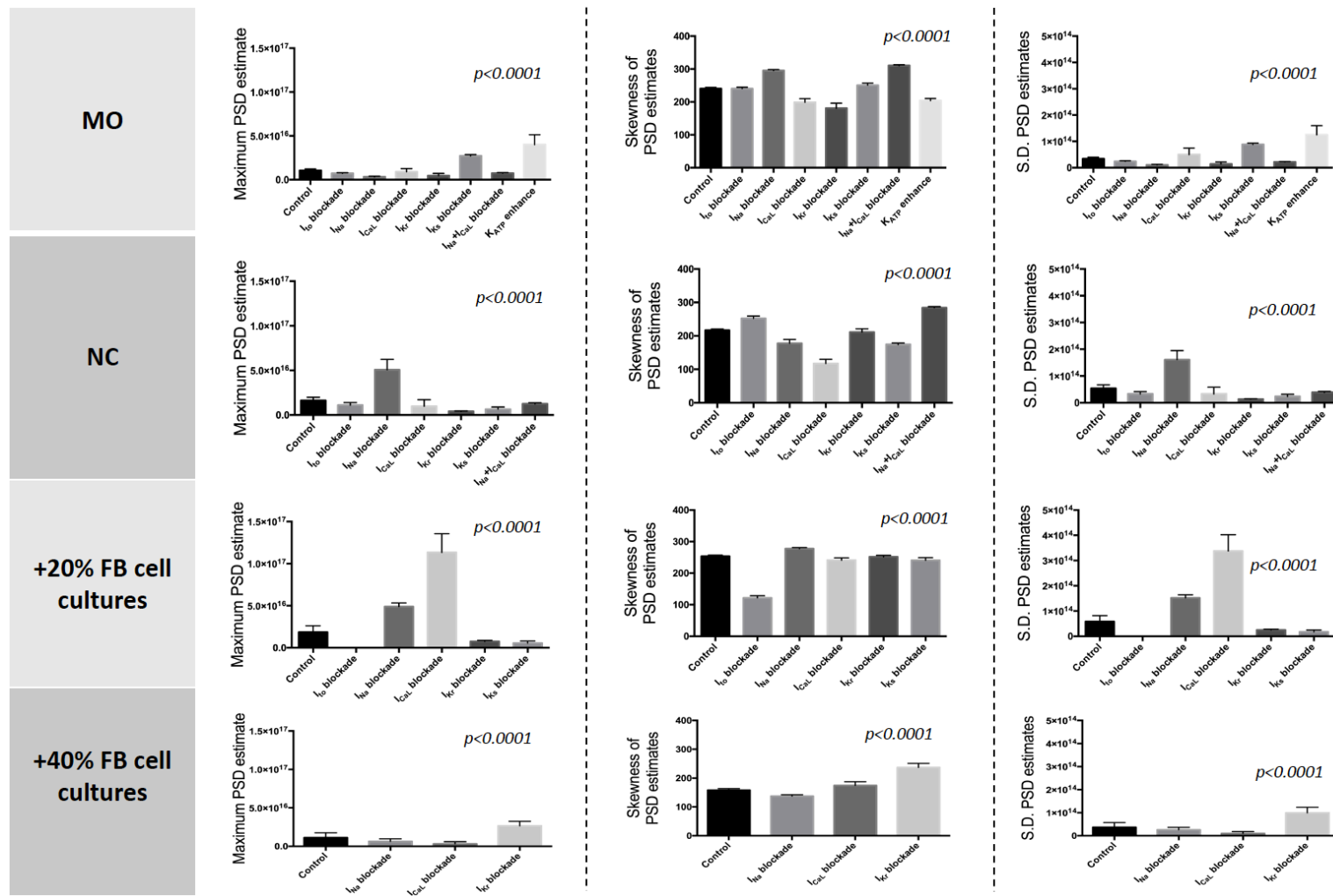


Figure 4.24. Bar charts presenting the modification of the maximum PSD estimate, the skewness of PSD estimates and the S.D. of PSD estimates due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent mean \pm SEM (n values presented in Table 4.3).

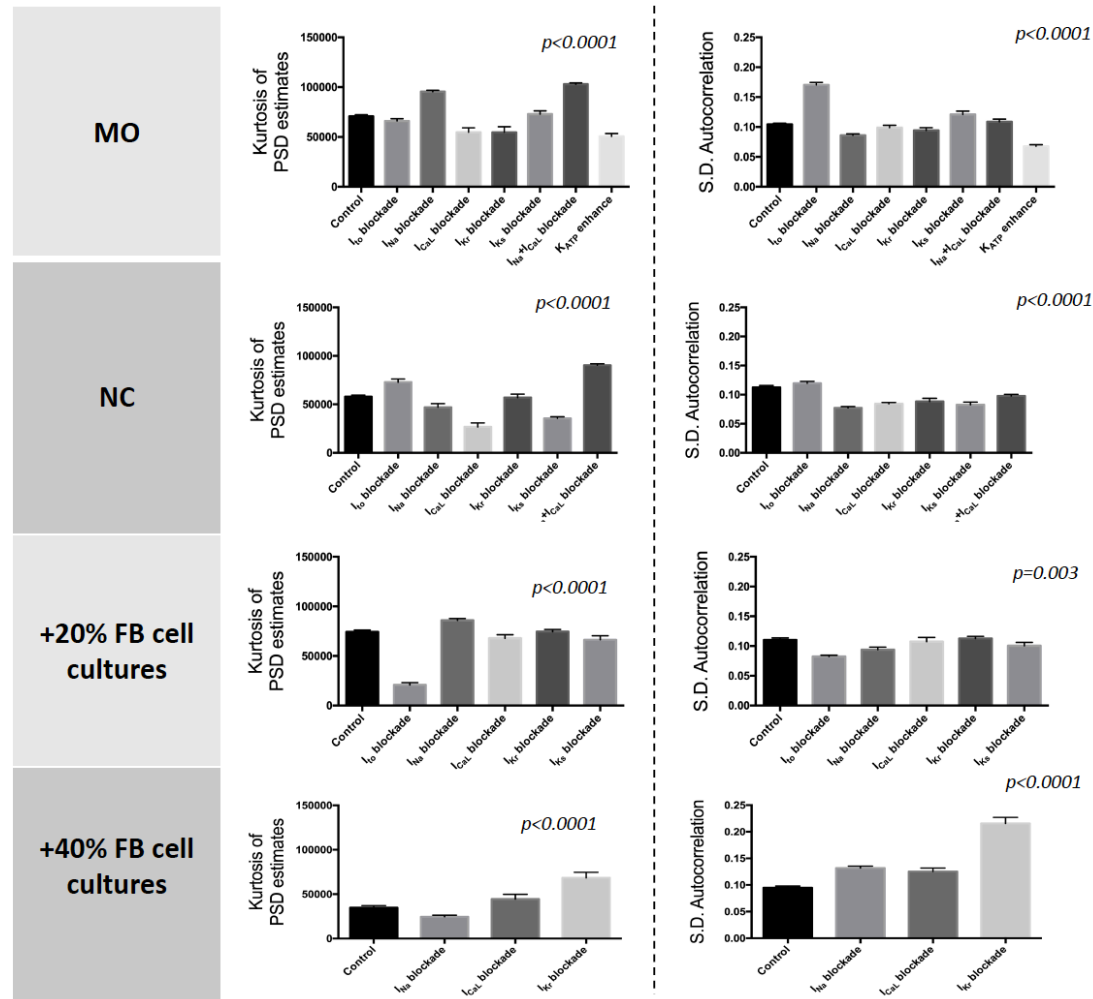


Figure 4.25. Bar charts presenting the modification of the kurtosis of PSD estimates and the S.D. of autocorrelation due to ion channel blockade in different types of NRVM cell culture with the presence of various amounts of fibroblasts. Majority of comparisons are significant with $p < 0.0001$, apart from a few exceptions. Kruskal-Wallis test analysis for non-parametric data. All bar charts represent $\text{mean} \pm \text{SEM}$ (n values presented in Table 4.3).

4.3.9 EGM morphology across different types of K⁺ current modulation

The results of automated feature extraction taken under baseline and following exposure to a K⁺ current related drug (n=72 EGMs for 750μM 4-AP, n=73 EGMs for 300nM E-4031, n=27 EGMs for 1μM HMR-1556, n=73 EGMs for 3μM pinacidil) were then compared. K⁺ current enhancement with pinacidil had different effects to K⁺ current blockade with 4-AP, E-4031 and HMR-1556 in many cases (Figures 4.32-34). There were only three features modified due to pinacidil, but not any due to an ion channel blocker (RR interval: p<0.0001; Frequency of minimum energy: p<0.0001; Log energy entropy: p<0.0001) (Figure 4.26).

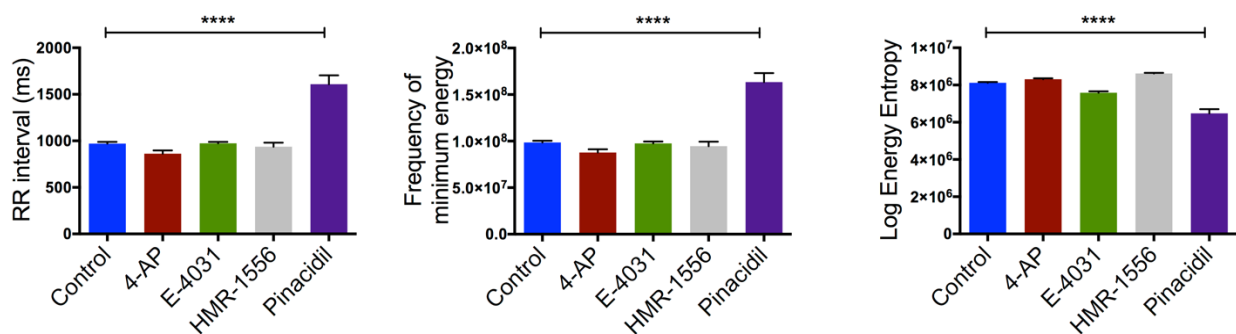


Figure 4.26: Features changed by K⁺ current enhancement with pinacidil, but not by K⁺ current blockade by 4-AP, E-4031 or HMR-1556. There is no significant difference among control, 4-AP, E-4031 and HMR-1556 data. Kruskal-Wallis test analysis. All bar charts represent mean±SEM; ****p<0.0001. (Control: n=190 EGMs; 4-AP: n=72 EGMs; E-4031: n=73 EGMs; HMR-1556: n=27 EGMs; pinacidil: n=73 EGMs).

FPD was the only feature measured manually compared to the rest and it was found that it changed in opposing directions in response to blockade (control: 156.2 ± 8.7ms vs. 4-AP: 164.7 ± 6.6ms – p=0.47; E-4031: 207.7 ± 15ms – p=0.003; HMR-1556: 196.2 ± 8ms – p=0.047; mean ± SEM) and enhancement (pinacidil: 146.6 ± 2.9ms – p = 0.01) (Figure 4.27). However, pinacidil influenced EGM features in the same direction as some of the K⁺ current blockers in some occasions. Such an example is EGM amplitude, where 4-AP caused a voltage reduction (control: 574.5 ± 31μV vs. 4-AP: 305.3 ± 21μV; p<0.0001), while E-4031, HMR-1556 and pinacidil caused an increase (E-4031: 1044 ± 73μV – p<0.0001; HMR-1556: 758.2 ± 73μV – p=0.003; pinacidil: 737.3 ± 60μV – p=0.007). There were also instances, such as RS interval, S-peak width, scale of minimum energy and mutual information (Figure 4.27-4.), in which pinacidil did not have any effect on EGM features compared to the ion channel blockers.

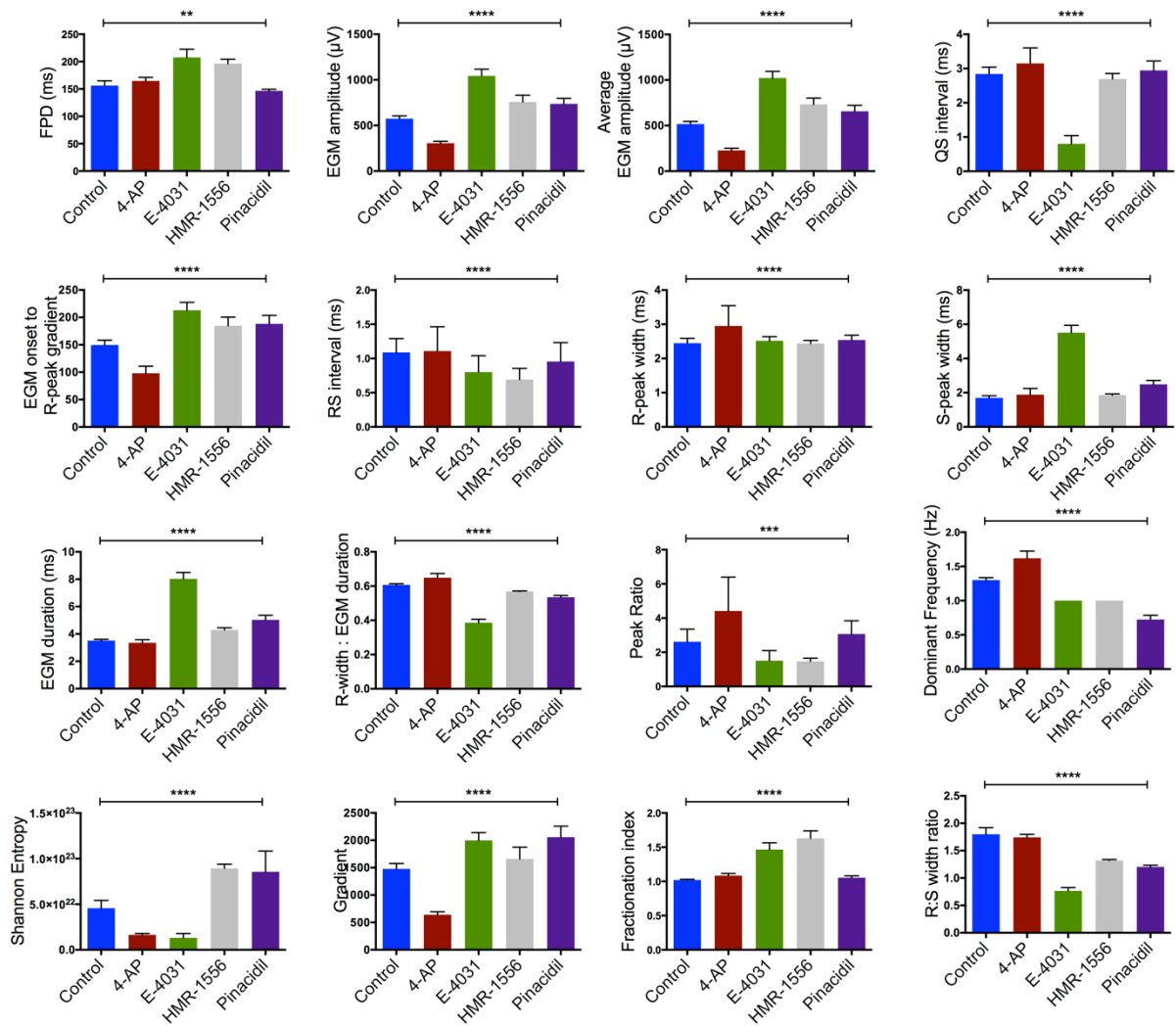


Figure 4.27: Features mainly changed in an indiscriminate manner following drug exposure. With the exception of FPD, all features were quantified by automated feature extraction. The feature may be changed solely by one drug (such as EGM duration), by two drugs (such as EGM amplitude) or by all of them (such as FPD). FPD changes agree with ion channel blocker activity. Kruskal-Wallis test analysis. All bar charts represent mean±SEM; **p<0.01; ***p<0.001; ****p<0.0001. (Control: n=190 EGMs; 4-AP: n=72 EGMs; E-4031: n=73 EGMs; HMR-1556: n=27 EGMs; pinacidil: n=73 EGMs). The FPD bar graph were made using manually annotated data.

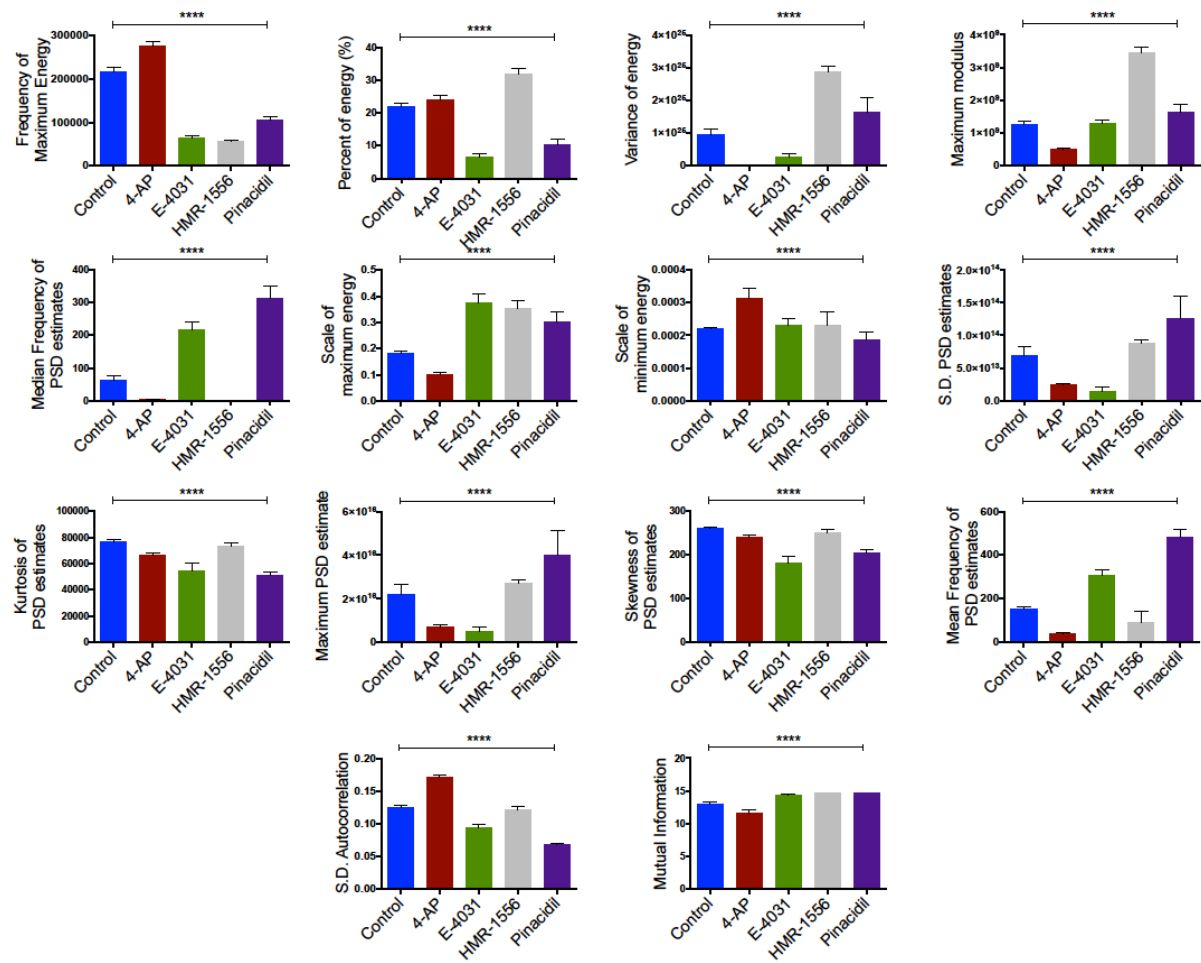


Figure 4.28: Features changed in an indiscriminate manner following drug exposure. Features changed in this way would not be informative enough when attempting to distinguish K^+ current enhancement from blockade. Kruskal-Wallis test analysis. All bar charts represent mean \pm SEM; **** p <0.0001. (Control: n=190 EGMs; 4-AP: n=72 EGMs; E-4031: n=73 EGMs; HMR-1556: n=27 EGMs; pinacidil: n=73 EGMs).

4.4 Discussion

With the characterisation of EGM morphology following the modification of ion channel activity in cell monolayers, it was demonstrated that EGM morphology alterations can be extremely variable. These modifications are mainly dependent on the role of the function which is disrupted. Even though the average morphology changes in overall, not all features are significantly modified. Furthermore, there is always a different group of EGM characteristics that is related to these changes making it the EGM signature following abnormalities at cellular level.

As AP morphology is better characterised following ion channel modifications, specific AP features were investigated along their EGM correlates. That was the method in order to confirm the EGM effects due to changes in cellular function. Although there was discernible relationship between AP and EGM morphology features, this could not be achieved for the majority of characteristics analysed in this study, due to lack of knowledge on them. All but time-domain features were particularly under that category.

All modifications and comparisons among experimental conditions presented in this chapter can justify the features selected during the automated feature selection process applied for the Chapter 7 findings and partially explain the machine learning results that followed in the same Chapter.

4.4.1 Na⁺ channel blockade effects on EGM morphology

The effects of Na⁺-channel deficiency on EGM morphology were assessed by exposing MO, NC cultures and NRVM/fibroblast co-cultures to increasing lidocaine concentrations. Lidocaine is a well-known class I antiarrhythmic that inhibits ionic currents through selectively binding on voltage-gated Na⁺ channels (254,255). The highest dose of lidocaine in each cell group was decided based on the fact that one-to-one capture was lost beyond these concentrations. A similar approach was used by Anderson *et al.* (256). That team suggested that the prolongation of QRS complex when combined with loss of one-to-one capture and increasingly distorted patterns during fixed-rate pacing were indications of progressive Na⁺-channel block (256). This study also showed that conduction velocity and EGM duration were inversely affected by lidocaine activity, leading to a strong relationship between upstroke duration and EGM duration. These findings agree with previous publications about conduction velocity reduction and upstroke duration increase due to Na⁺-channel blockade (256,257).

However, upstroke duration and EGM duration did not change when there was an increased amount of fibroblast in cell culture, i.e. NC, +20% FB, +40% FB, while conduction velocity was influenced both by Na⁺-channel block and the increased presence of fibroblasts. EGM amplitude was affected in all types of cell monolayer though in different ways suggesting that it is a sensitive parameter for the assessment of fibroblast burden and low-level Na⁺ channel block. Thus, EGM amplitude could be used as a marker for evaluating the severity of conditions related to simultaneous Na⁺-channel deficiency and irregular levels of fibroblast amounts. In the clinic, this could be useful for the improvement of the recent approach of therapeutic ablation in Brugada Syndrome (BrS) (258,259). EGM amplitude could

help assess the severity of the arrhythmic substrate after identifying responsible regions using other markers, such as fractionation (6).

EGM amplitude and fractionation may not be enough though for the accurate evaluation of a condition (7). Therefore, the fact that a relatively large number of EGM features were observed to notably change following the addition of high lidocaine doses in each type of cell culture suggests that other EGM features, beyond the traditional ones, need to be considered.

4.4.2 Transient outward K^+ current effects

Transient outward K^+ current, which is involved in the early repolarisation phase, is the net result of K^+ flux through two different types of channel. This channel activity is responsible for the fast I_{to} ($I_{to,f}$) current, mediated by channels that recover inactivation in less than 100ms, and the slow I_{to} ($I_{to,s}$) current, in which the recovery takes a few seconds (260,261). The expression of $I_{to,f}$ and $I_{to,s}$ channel proteins may vary with species, developmental stage and region of the heart (260). $I_{to,f}$ is usually more prominent in ventricular epicardial myocytes rather than endocardial cells. $I_{to,f}$ and $I_{to,s}$ have similar contributions to total I_{to} in neonatal rodent ventricular myocytes (262,263), as the NRVMs used in this study.

Changes in the early repolarisation phase affect the AP morphology. In specific, increased I_{to} current is correlated to AP shortening (262) and a decrease in I_{to} density, which is associated with aging, leads to AP prolongation (264). This is in agreement with the findings of this study, as it was shown that a reduction in I_{to} current in MO and NC cell cultures using 4-aminopyridine prolonged the APD. FPD, which is poorly studied in cell cultures, was found to be directly correlated to APD_{90} and to be influenced by I_{to} on the same way. The administration of 4-aminopyridine on MO and NC monolayers resulted to APD_{90} and FPD prolongation and a strong linear correlation between these two parameters. These results make a connection between optical mapping and EGM data collection showing that EGM morphology changes could be used for the prediction of adjacent AP modifications and subsequently functional changes at cellular level.

It was also observed that a large number of EGM features changed following the administration of 4-aminopyridine. 4-aminopyridine is a non-specific blocker of I_{to} and especially selective for the Ca^{2+} -independent (59,265) and 4-aminopyridine sensitive component often termed as I_{to1} (266). This fact may explain any unexpected alterations, such as the significant increase of EGM amplitude variation in NC cell cultures and the fractionation of data obtained from NRVM/+20% FB co-cultures. Another interesting effect of the 4-aminopyridine activity was the alteration of RR-interval in NC (reduction)

and NRVM/+20% FB co-cultures (increase). I_{to} downregulation and AP prolongation, both caused by 4-aminopyridine, enhance arrhythmia vulnerability when they are both present in disease states (260). In addition to re-entry predisposing changes in refractoriness dispersion (267), triggered activity is often associated with increased AP duration (260). The Ca^{2+} influx enhancement, due to the 4-aminopyridine activity, leads to a greater cell and sarcoplasmic reticulum (SR) Ca^{2+} load. The SR Ca^{2+} overload influences the SR Ca^{2+} release facilitating arrhythmogenesis (260,268). This could explain the fact that RR-interval was diminished following I_{to} current blockade.

4.4.3 L-type Ca^{2+} channel blockade effects

Nifedipine is a well-studied L-type Ca^{2+} channel blocker (269). The present study demonstrated through optical mapping that nifedipine leads to decrease of APD_{90} in MO and NC cell cultures, as a result of the loss of the plateau phase. This indicates the successful L-type Ca^{2+} channel blockade and allows the assumption that the simultaneous decrease of FPD on unipolar extracellular EGMs is due to the same fact.

L-type Ca^{2+} channel modulations, more often due to gene mutations, are responsible for QT interval shortening and the short QT syndrome (270). This clinical condition is related to families characterised by sudden cardiac death, atrial fibrillation and Brugada type I ECG pattern (57). There is poor knowledge on the EGM effects of L-type Ca^{2+} channel deficiency or blockade on cultured ventricular monolayers and since the dual modality experiments showed that I_{CaL} blockade could be predicted by analysing the FPD, it would be interesting to confirm QT shortening on NRVM/cultured fibroblast co-cultures. However, the inability for accurate T-wave detection, as described in Chapter 3, was a prohibiting factor. It was also observed that a relatively small number of features were altered. This could be a result of the specific pharmacological activity of nifedipine among a variety of L-type Ca^{2+} antagonists, such as verapamil and diltiazem (271).

4.4.4 The influence of rapidly activating delayed rectifier K^+ current on EGM morphology

The ventricular action potential of most species is terminated by the delayed rectifier K^+ current (I_K). It consists of a rapidly activating (I_{Kr}) and slowly activating (I_{Ks}) component (272). E-4031, a class III anti-arrhythmic agent (273), has been shown to prolong APD in guinea pig ventricles and papillary muscles derived from adult rats and guinea pigs (274) by blocking ERG channel, the rat analogue of

HERG channel (275). In a similar way, the optical mapping results showed that E-4031 significantly prolonged APD₉₀. This effect took place simultaneously with the FPD prolongation and a direct correlation between these AP and EGM characteristics was found. This agrees with previous publications reporting that the blockade of HERG channel by class III anti-arrhythmics is responsible for APD prolongation alongside prolonged QT interval, as measured by body surface ECG, and subsequent long QT syndrome (LQTS) or Torsades de Pointes (TdP) (276).

Loss of one-to-one capture, indicating fibrillation event, was observed on EGM data collected from NC and +40% FB cell cultures. This change was accompanied with extended RS interval in NC cultures and diminished EGM amplitude and enhanced voltage variability within electrode recordings in +40% FB monolayers. There are a number of assumptions that can be made for the reason of these effects. Even though it is known that ERG channels are responsible for the generation of I_{Kr} current in rats (275), it does not seem that the ERG subunits are expressed in fibroblasts extracted from neonatal rat ventricles (277). This suggests that I_{Kr} may be absent from neonatal rat ventricular fibroblasts, and therefore the total I_{Kr} current of NC and +40% FB cell cultures could be lower than normal in any experimental model of this study. As a result, the I_{Kr} density may be lower in these types of cell culture and the effect of E-4031 is more severe. However, the variable time-domain feature modifications may be due to the different electrophysiological characteristics between cultured and naturally occurring fibroblasts. Fibroblast culture produces phenotypic changes resulting in activation of pathological signalling pathways and the expression of myocyte-like characteristics, such as increased connexin expression and electrical coupling (278,279). Thus, the electrophysiological mechanisms must be different between NC and +40% FB cultures.

4.4.5 Slowly activating delayed rectifier K⁺ current related changes

The slowly activating delayed rectifier K⁺ current (I_{Ks}) is one of the K⁺ currents, the other being I_{Kr}, predominantly contributing to repolarisation of the cardiac action potential (280,281). The inhibition of these currents, resulting to AP prolongation, is considered to be antiarrhythmic and proarrhythmic depending on conditions (281,282). Even though the effects of I_{Kr} blockade are well investigated, there is poor knowledge about the ways that I_{Ks} blockade influences cellular electrophysiology, due to the lack of available highly selective I_{Ks} blockers (283). HMR-1556 ((3R,4S)-(+)-N-[3-hydroxy-2,2-dimethyl-6-(4,4,4-trifluorobutoxy)chromatin-4-yl]-N-methylmethanesulfonamide is a novel chromanol developed approximately a decade ago as a selective I_{Ks} blocker (283–285). The effects of HMR-1556 in NRVMs were previously untested. The range of doses used in this study was broader than that used for the initial characterisation of the drug in isolated guinea pig ventricular myocytes

(285). Despite this fact, no changes could be provoked either on APD₉₀ or FPD measured on myocyte only and NC cell cultures in response to HMR-1556. This was unexpected considering the specificity of HMR-1556 for I_{Ks} current (283,285) and the role of I_{Ks} in repolarisation in NRVMs paced at 1Hz (286). This effect could be explained by HMR-1556 applications in other experimental models, where it was shown that the combined blockade of I_{Kr} and I_{Ks} was necessary for APD prolongation following the administration of HMR-1556 (287).

Apart from FPD, other components of EGM morphology were also affected by HMR-1556. These features were influenced in variable ways depending on the type of K⁺ blockade (I_{to}, I_{Kr}, I_{Ks}). This suggests that EGM feature extraction, as applied in this study, could be used in order to make refined discriminations about the function of specific currents. In addition, the pattern of EGM morphology changes was very similar between NC and +20% FB cell cultures, but completely different to the one observed in EGMs recorded from myocyte only cell cultures. The increased presence of fibroblasts in NC and NRVM/FB co-cultures could explain this, but there is no clear biological explanation for this observation warranting further investigation nonetheless. I_{Ks}-induced modifications are better investigated using ECG on the context of LQTS and much less work has been carried out on how EGM morphology may be affected (288). The results presented in this study are a first step towards this.

4.4.6 K⁺ current enhancement alters EGM morphology

Cardiac K_{ATP} channels are closed under normal conditions and their contribution to the AP is limited. The imposition of ischaemia and other metabolic stressors provoke the opening of ion channels to provide outward current that leads to AP abbreviation. The enhancement of I_{KATP} in this study was achieved by using pinacidil, a drug which has been showed to influence the transient outward current in some models (289). As a result, the dual modality data showed that APD₉₀ and FPD were reduced. This agrees with the current knowledge that the enhancement of transient outward K⁺ current abbreviates the AP. The AP plateau (phase 2) becomes shorter limiting Ca²⁺ influx to reduce contractility and preserve cell energy supplies (290). Moreover, previous patch clamp studies showed that pinacidil enhanced I_{KATP} in NRVMs resulting to the abbreviation of FPD (291) in a similar way as presented in this study.

Further to the initial characterisation of FPD, any alterations to non-conventional EGM morphology features were assessed. A number of features were changed following the administration of pinacidil and in some cases these modifications were opposing to K⁺ current blockade effects. This is another indication within this study that the performed EGM morphology analysis has the ability to distinguish

functional modulations. ECG analysis in the time-frequency domain is well established, but there is much less experience on analysing EGMs in this way (292). These findings characterise the differences in EGM morphology alterations between K^+ current blockade and enhancement.

There were also features affected by drug exposure, but not enough to be informative for distinguishing K^+ current enhancement from blockade. Finally, there were features altered in the same direction by pharmacological agents with opposing effects, which is an indication of low specificity. For instance, pinacidil is known to reduce I_{to} in relatively low concentrations in canine ventricular myocytes (13). Since this is an effect of mimicking the 4-AP activity, it could also explain the fact of similar effects between drugs with opposing activity.

4.4.7 Investigating the effects of double current blockade

Ion channel remodelling occurs in a variety of cardiac diseases and rhythm disturbances. These alterations may be part of the homeostatic adaptive response to the primary abnormality, but they can also result in secondary cardiac dysfunctions (293). In many occasions multiple ion channel dysfunctions may take place simultaneously, such as in congestive heart disease, myocardial infarction and atrial fibrillation (293). I_{Na} and I_{CaL} currents can be concurrently downregulated, but this is also species specific (293). It has been observed that I_{Na} may not be reduced in AF patients due to no changes in I_{Na} density (294) or the expression of α -subunits of Na^+ channels at the mRNA level (295). There is lack of experience on the effects of double ion channel blockade on AP and EGM morphology, especially on NRVM and this study used myocyte only and NC cell cultures for proof-of-concept investigations.

It was observed that APD_{90} and FPD were both abbreviated in myocyte only, but not in NC cultures, following the concurrent I_{Na} and I_{CaL} blockade using lidocaine and nifedipine. This could be the result of I_{CaL} blockade. Ca^{2+} reduction is an adaptation mechanism in atrial myocytes during AF, as a result of increased atrial Ca^{2+} loading during tachycardia (296), in order to minimise potentially lethal Ca^{2+} overload (297). The I_{CaL} reduction may be one of the AF consequences, but it is also one of its maintenance contributors. By blocking the I_{CaL} current in this study, I managed to reproduce the APD reduction which is typical of atrial tachycardia remodelling (293,298). On the other hand, the reason that I_{Na} reduction did not manage to prolong upstroke duration and EGM duration or lead to conduction slowing is unclear. Conduction slowing was not statistically significant. However, there was either rotational activity on cell monolayers on highest lidocaine/nifedipine doses or loss of cellular

electrophysiology activity. Previous studies showed that when I_{Na} is decreased, especially due to loss-of-function mutations, it can promote AF by slowing conduction velocity and causing reentrant activity (54).

RR interval was the only time-domain characteristic to be modified in myocyte only cell cultures, while the rest of features belonged to time-frequency domain and frequency-domain analysis. However, a different group of features was influenced in NC monolayers which included the voltage elimination, which could be the result of I_{Na} reduction (293). No changes in upstroke duration or EGM duration were found though which is contradicting since I_{Na} regulates the depolarisation (293).

4.5 Limitations

There is no doubt that T-waves are routinely used in the clinic for the assessment of clinical conditions, such as Torsades de Pointes (TdP) and long QT syndrome (LQTS) (276,288,299). It is also the main way of studying the effects of ion channel remodelling in repolarisation (281,288). Therefore, one of the major limitations of this study arose from the absence of automated detection of T-waves, and the subsequent analysis of T-wave related features, such as FPD, T-wave amplitude and duration. This would allow us to investigate how the QT interval is modified under functional modulations, not only in cell monolayers but also in tissue slices. Apart from the dual modality experiments conducted during this study, the effects on repolarisation were assessed using EGM measures only. This may be unimportant though given the close correlation noted between APD_{90} and FPD in this investigation and elsewhere (40). The reasons behind this limitation have been discussed in Chapter 3.

The generation of the action potential and the regional differences that are observed throughout the heart are the result of the greater complexity of ion channel function consisting of a network of ion channels distributed on the cell membrane (300). Even though the main ion currents were studied in isolation for the purposes of this study, there are a number of additional ones which still remain to be investigated for their influence on EGM morphology. These currents include I_{Kur} , I_{KAch} , I_{K1} , NCX and I_f (300). Since each one of these currents play a key role in the generation of a normal action potential and abnormalities on them are related to cardiac diseases, it would be interesting how they affect EGM morphology both in isolation and in combination with other ion currents.

The majority of ion channel studies are focused on AP changes and regarding EGM morphology, these are limited to basic time-domain characteristics. However, the EGM morphology analysis showed that a variety of features belonging to varied types of data analysis can be affected. Our current knowledge

about the role of ion channel deficiencies on cardiac arrhythmias makes it difficult to suggest any mechanisms which are responsible for changes in time-frequency domain and frequency-domain analysis feature. Even though modulation-specific patterns were found, it would be interesting to understand the reasons behind each modification.

Dual modality experiments combining optical mapping and MEA recordings at multicellular preparations were carried out in this study. This provided direct and concurrent validation of how AP changes at a single cell level can manifest on EGM morphology. However, these experiments were not conducted on each type of cell culture treated with each modulator, due to difficulties in preparing cell monolayers that could be stimulated and the variability of the quality of samples.

4.6 Conclusions

The effects of ion current modifications on EGM morphology, either in isolation or combined with cellular scar, were systematically demonstrated in this chapter. Some of the EGM features were primarily correlated to AP morphology changes, in order to validate the effect of the pharmacological agents used in this study. It was found that a number of EGM features changed in each case, and while there may be similarities in a few cases, the extent to which features were modified was unique for each type of modification. Therefore, these features could be used to distinguish the groups when a modification is unknown. It can be suspected that each drug used in this thesis has relatively specific effects compared to the rest and any morphology changes constitute the EGM signature characterising a specific modulation. In overall, interesting insights were gained on the way that EGM morphology can be affected and this knowledge could promote further studies for the understanding of the underlying mechanisms that occur on ion current abnormalities.

Chapter 5

EGM morphology effects of gap junctional block
in multicellular preparations

5.1 Introduction

Gap junctions are transmembrane channels that facilitate intracellular coupling and allow the transfer of ions and small molecules between cells by directly linking the cytoplasmic components of neighbouring cells (43–45). The involvement of gap junctions on the electrical coupling between cells plays a key role in allowing the orderly spread of electrical excitation across the myocardium. Gap junction remodelling, which involves alterations in expression, function and distribution, is known to affect cardiac impulse propagation and is also related to cardiac pathologies and cardiac arrhythmias (43,44). Gap junction remodelling is related to changes in gap junction function, such as the phosphorylation state, or structural modifications, such as changes in expression levels or the cellular localisation of connexin. Multiple studies have showed that the altered gap junction localisation and expression levels contribute in arrhythmogenesis both in human and animal models (44,301,302). Cx43 may change its cellular localisation from the intercalated disc at the ends of myocytes to the lateral membrane away from the intercalated disc. This phenomenon is known as lateralisation (303). Structural remodelling has been also associated with conduction slowing in dilated cardiomyopathy (304), ischaemia (305), heart failure (306) and atrial fibrillation (307).

Pharmacological gap junction modulators are potential anti-arrhythmic agents, because gap junction abnormalities have been found to be part of arrhythmogenic mechanisms (308). These drugs modulate the gap junctional intercellular communication either by uncoupling or coupling. The uncoupling involves the enhancement of the intercellular electric resistance, while coupling includes the opposite mechanism. A number of agents are currently available for gap junction modulations. Lyophilic agents, such as heptanol, octanol, myristoleic acid and carbenoxolone act as uncouplers (309–311). AAP10 and rotigaptide are known for enhancing gap junctional intercellular communication (312,313).

According to previous studies on isolated rabbit ventricular myocytes, carbenoxolone, and consequently conduction slowing, is responsible for voltage reduction and a decrease in the steepest negative slope of unipolar extracellular EGMs (314). Fractionation has been also related to modulated intercellular communication due to changes in atrial Cx40 and Cx43 distribution (315), which has been suspected to be another way of gap junction remodelling (43). Moreover, we have shown that carbenoxolone leads to abbreviated APD₉₀ and consequently FPD in myocyte only NRVM cultures (40). It is unclear though how EGM morphology obtained from isolated cells may be further modified due to the gap junctional block. A better characterisation of EGM morphology following gap junction uncoupling is necessary, in order to accurately distinguish this type of remodelling from ion current abnormalities, as discussed in Chapter 4, using EGM features in addition to the conventional ones of voltage and fractionation.

As it has already become clear through *in vitro* and *in vivo* studies, and as it was presented in Chapter 1, there is a limited understanding of the underlying mechanisms leading to EGM modifications. *In vitro* and *ex vivo* models are preferred though for the present investigation because of the degree of manipulation which is possible, compared to an *in vivo* experimental setting. On one hand, NRVM monolayers are well established and characterised (172). On the other hand, thin tissue slices of the human myocardium retain structural and functional properties of native myocardium and remain viable for several hours *ex vivo* (20). In addition, conduction velocities in such *in vitro* and *ex vivo* models are similar and representative of *in vivo* findings (20).

In this part of the thesis carbenoxolone was used as a tool to investigate the role of gap junction uncoupling on specific modifications of EGM morphology. The hypotheses to be addressed were:

1. APD modifications due to gap junction uncoupling can be correlated to modified EGM features.
2. Administration of carbenoxolone in NRVM monolayers will lead to conduction slowing and fractionated and low voltage EGM due to reduction in intercellular communication.
3. Conduction slowing due to gap junction uncoupling is presented through additional feature alterations.
4. Administration of carbenoxolone in tissue slices will have similar effects on EGM morphology as in cell cultures.

For this purpose, the effect of gap junction uncoupling was investigated by administering CBX in the different types of cell culture, presented in section 2.2.1, and carrying out the dual modality technique showed in section 4.2.4. EGM and action potential recordings were analysed and correlated to each other. The range of 5 μ M – 100 μ M CBX was investigated in each culture – higher doses were toxic for NRVMs. The EGMs, recorded in the baseline level and following gap junction uncoupling by the CBX concentration closest to the IC₅₀ for each culture, were analysed using the automated feature analysis presented in Chapter 3.

5.2 Results

5.2.1 The effects of gap junction uncoupling on cellular EGM morphology

Gap junction uncoupling was carried out on myocyte only NRVM cultures using CBX. The IC_{50} , which was used during concurrent optical mapping and MEA recordings, was $15.8\mu\text{M}$. This was calculated as the average of IC_{50} concentrations at 1Hz and 5Hz ($16.7\mu\text{M}$ and $14.9\mu\text{M}$ respectively) (Figure 5.1A). Electrical restitution occurred by stimulating cells at progressively reduced pacing cycle lengths from 1000 to 200ms both before and after the administration of CBX. Both APD_{90} (control: 1Hz $190.78\pm 48\text{ms}$, 5Hz $125.7\pm 3\text{ms}$, $n=6$ arrays; CBX: 1Hz $151.5\pm 37\text{ms}$, 4Hz $115.7\pm 17\text{ms}$, $n=5$ arrays; $p<0.01$; mean \pm SEM) and FPD (control: 1Hz $191.5\pm 47\text{ms}$, 4Hz $113.4\pm 18\text{ms}$, $n=6$ arrays; CBX: 1Hz $149.8\pm 46\text{ms}$, 4Hz $88.3\pm 27\text{ms}$, $n=5$ arrays; $p<0.001$; mean \pm SEM) were affected not only by gap junctional block, but also by pacing rate (APD_{90} : $p<0.001$; FPD: $p<0.0001$) (Figure 5.1B-C). Linear regression analysis showed a direct correlation between APD_{90} and FPD ($r^2=0.82$, $Sy.x=18.7$, slope = 1.14 ± 0.07 , $p<0.0001$, $n=59$) and there was no significant change in correlation coefficients and slopes when this relationship was studied before or after gap junction uncoupling (control: $r^2=0.85$, slope = 1.079 ± 0.09 ; CBX: $r^2=0.72$, slope = 1.1 ± 0.14 ; $p=0.06$) (Figure 5.1D). The Bland-Altman plot revealed that FPD was systematically 38ms shorter after the administration of CBX (Figure 5.1E). The differences did not seem to be affected by the magnitude of average FPD and approximately 7% (2 out of 28) of the differences were out of the limits of agreement for FPD.

The automated analysis of EGM morphology showed that 20 features changed (Figure 5.2A) due to the administration of $20\mu\text{M}$ CBX, the closest concentration to the IC_{50} from those tested in the range of $5\mu\text{M}$ – $100\mu\text{M}$. Significant changes mostly involved reductions of feature characteristics, such as the average EGM amplitude (control: $392.4\pm 8\text{ms}$; CBX: $85\pm 2\text{ms}$; $p<0.0001$; mean \pm SEM). Among the features that increased was fractionation index (control: 1.2 ± 0.04 ; CBX: 2.7 ± 0.1 ; $p<0.0001$) and RS interval (control: $0.55\pm 0.02\text{ms}$; CBX: $1.74\pm 0.08\text{ms}$; $p<0.0001$; mean \pm SEM). In overall, all extracted features were modified to some extent (Figure 5.2B).

Concurrent optical mapping and MEA experiments were carried out on NC cell cultures and NRVM/+20% fibroblasts co-cultures, in combination with gap junction uncoupling. For that purpose, the IC_{50} of CBX was used, which was extracted from dose response curves based on data collected from each cell group (NC: $38.8\mu\text{M}$; Low %FB: $2.65\mu\text{M}$) (Figure 5.3A-B). Rate adaptation of APD_{90} and FPD

before and after the gap junction uncoupling at NC monolayers was seen ($p=0.0015$ and $p=0.0008$ respectively) (Figure 5.3C-D).

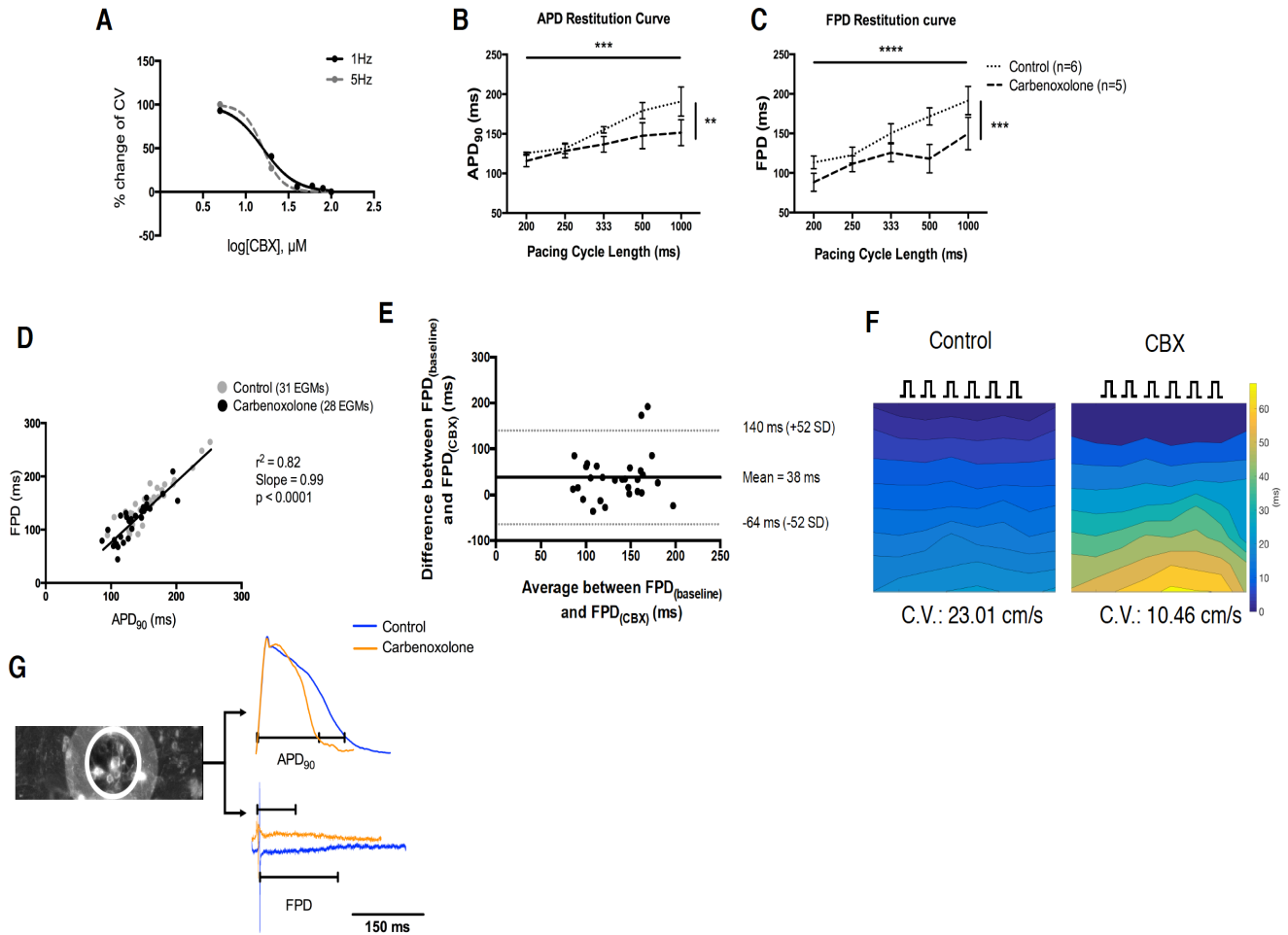


Figure 5.1: (A) Dose response curve for the effect of CBX on CV at 1Hz and 5Hz, in order to extract the average of IC₅₀ at both frequencies. (B) APD₉₀ restitution curve showing that it is affected both by pacing rate ($p=0.007$) and gap junction uncoupling ($p=0.0002$). (C) Restitution curve presenting how FPD was affected by pacing rate ($p<0.0001$) and CBX activity ($p=0.0008$). Control: $n=6$ arrays; CBX: $n=6$ arrays both for (B) and (C). APD₉₀ and FPD plots were made using manually annotated data. (D) Linear relationship between FPD and APD₉₀ ($r^2=0.82$; $Sy.x=18.7$, slope=0.99; $p<0.0001$, $n=59$). (E) Bland-Altman plot for FPD at baseline and following gap junction uncoupling using CBX in NRVM cell monolayers. The middle bold line is the average difference between the baseline and gap junction uncoupling. Two additional lines are the upper and lower bounds of the limits of agreement. (F) Isochronal activation maps for the wavefront propagation before and after gap junction uncoupling in myocyte only cell cultures. (G) Superimposed action potentials (top) and EGMs (bottom) obtained from the same electrode (left) before and after gap junction uncoupling. Mean \pm SEM; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$.

However, there was no change in APD₉₀ and FPD due to the CBX activity ($p=0.48$ and $p=0.57$ respectively). There was also no observed difference in APD₉₀ in NC and +20% FB cell cultures (NC: 1Hz control – 170.8 ± 6.3 ms, CBX - 208.5 ± 24.7 ms, $n=4$ arrays; +20% FB: 1Hz control – 206ms, CBX – 210.9ms, $n=1$ array, $p=0.48$; mean \pm SEM) (Figure 5.3E). Furthermore, no FPD change was observed in both cell groups (NC: 1Hz control – 167.2 ± 8.7 ms, CBX - 185 ± 25.5 ms, $n=4$ arrays; Low% FB: 1Hz control – 193ms, CBX – 231ms, $n=1$, $p=0.39$; mean \pm SEM) (Figure 5.3F). Linear regression analysis showed a direct correlation between APD₉₀ and FPD measured on data obtained from both types of cell culture ($r^2 = 0.65$, $Sy.x=21.1$, slope = 0.97 ± 0.07 , $p<0.0001$, $n=104$) and there was a linear relationship even within each group (NC: $r^2 = 0.56$, $Sy.x=19.1$, slope = 0.88 ± 0.1 , $p<0.0001$, $n=90$; Low %FB: $r^2 = 0.49$, $Sy.x=28.8$, slope = 0.82 ± 0.2 , $p = 0.005$, $n=14$) (Figure 5.3G). Conduction velocity was affected both by the type of cellular scar ($p=0.012$) and gap junction uncoupling (NC: 1Hz control – 22.22 ± 1.2 cm/s, CBX – 7.7 ± 1.8 cm/s, $n = 4$ arrays; Low %FB: 1Hz control – 5.86cm/s, CBX: 2cm/s, $n = 1$; $p = 0.027$; mean \pm SEM) (Figure 5.3H).

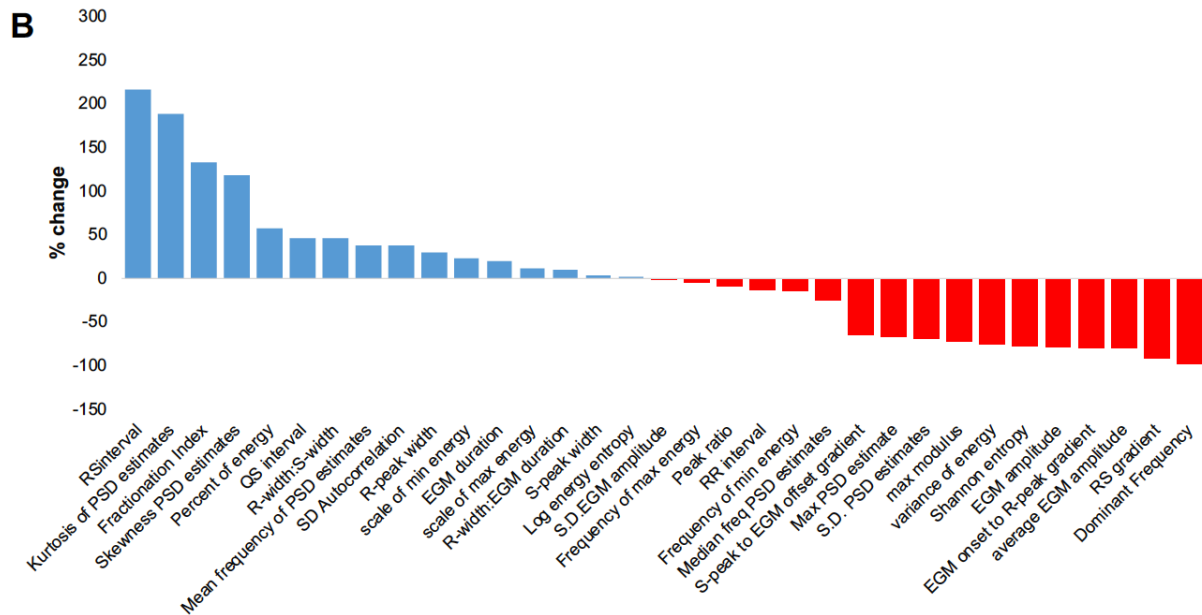
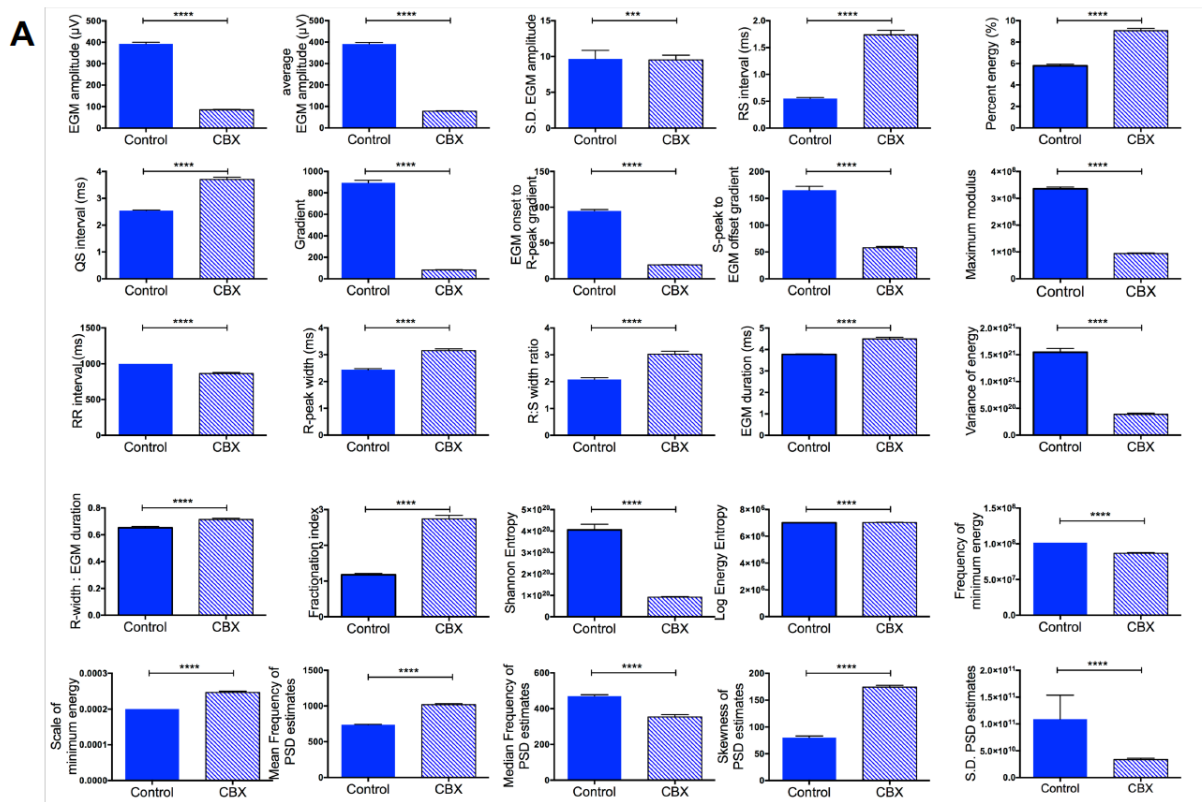


Figure 5.2: EGM morphology modifications due to gap junction uncoupling when 40 μ M CBX were administered on myocyte only cultures. (A) Features significantly changed due to CBX. Mann-Whitney test analysis. (B) Quantification of EGM morphology changes following gap junction uncoupling using 40 μ M CBX as compared to the average control EGM morphology. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (Control: n = 481 EGMs; CBX: n =446 EGMs).

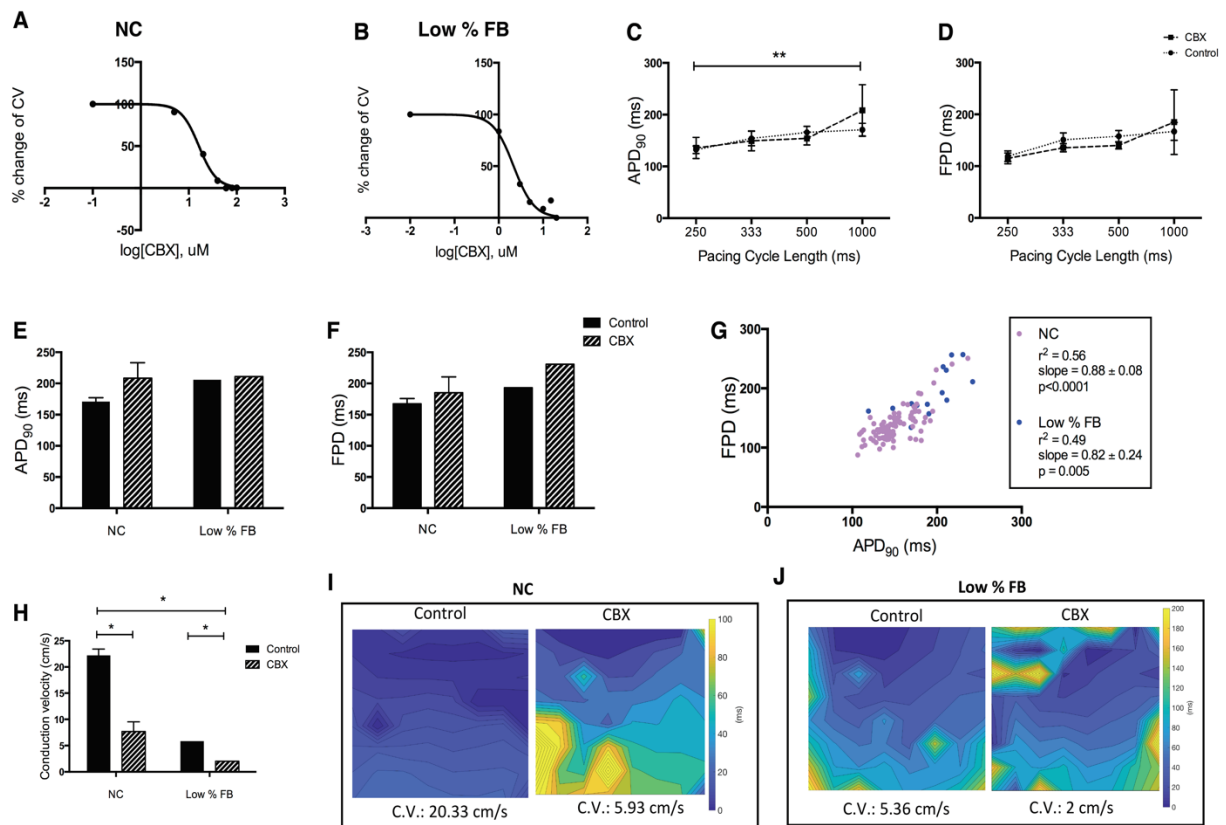


Figure 5.3: (A-B) Dose response curves showing the effect of a range of CBX doses on the conduction velocity measured on natural cell composition (A) cultures and co-cultures of NRVMs with low amount of cultured fibroblasts (B). (C-D) Restitution curves showing that APD₉₀ (C) was affected by the pacing cycle length only ($p=0.002$), and gap junction uncoupling did not cause significant changes either on APD₉₀ or FPD (D), obtained from NC culture data. (E-F) No significant changes due to CBX activity were observed on the APD₉₀ (E) or FPD (F) obtained from 1Hz recordings. APD₉₀ and FPD plots were made using manually annotated data. (G) Linear regression analysis showed a trend towards a linear relationship between APD₉₀ and FPD measured from NC and low amount of fibroblasts cell cultures. (H) Two-way ANOVA showed that conduction velocity was significantly reduced due to gap junction uncoupling in both groups. (I-J) Isochronal activation maps before and after gap junction uncoupling in NC cell cultures (I) and co-cultures of NRVM with low amount of fibroblasts (J). Mean \pm SEM. * $p < 0.05$; ** $p < 0.01$.

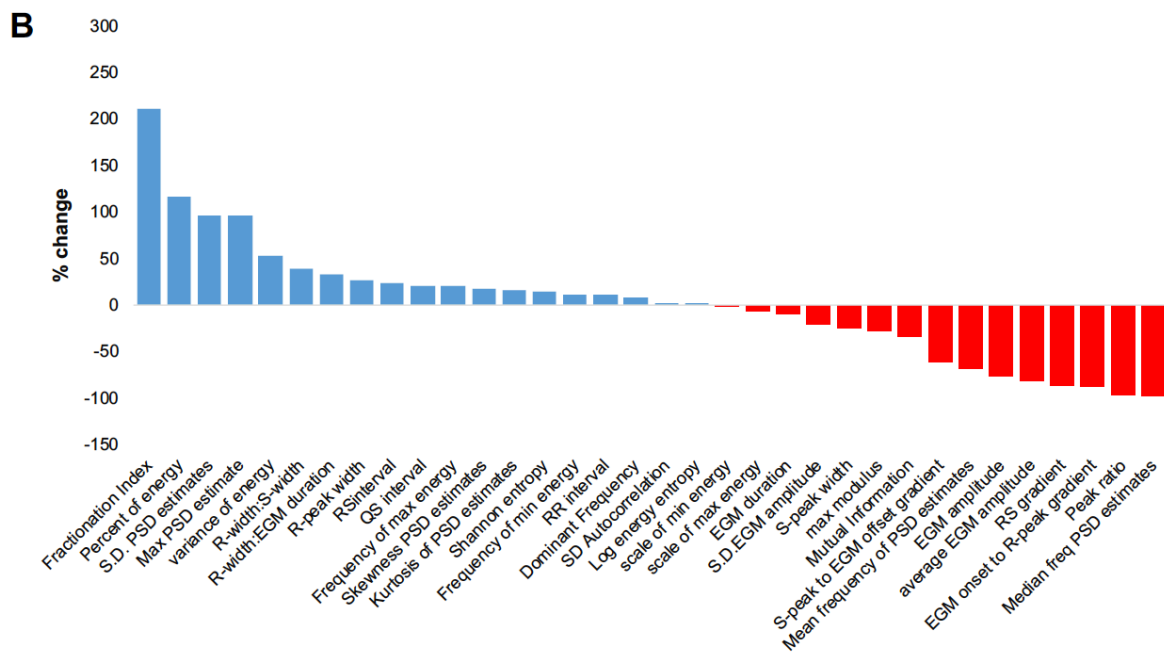
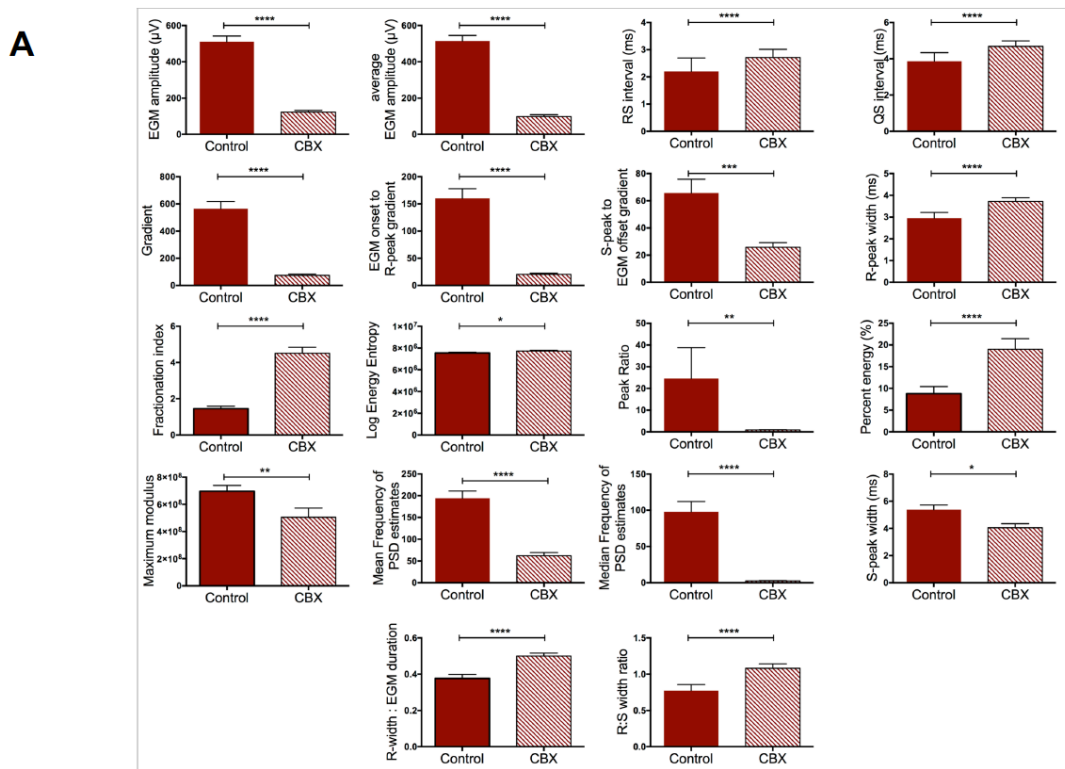


Figure 5.4: Modifications on EGM morphology due to gap junction uncoupling when 40 μ M CBX were administered on NC cultures. (A) Features significantly changed due to CBX activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) Quantification of EGM morphology changes following gap junction uncoupling as compared to the average control EGM morphology. (Control: $n = 64$ EGMs; CBX = 43 EGMs).

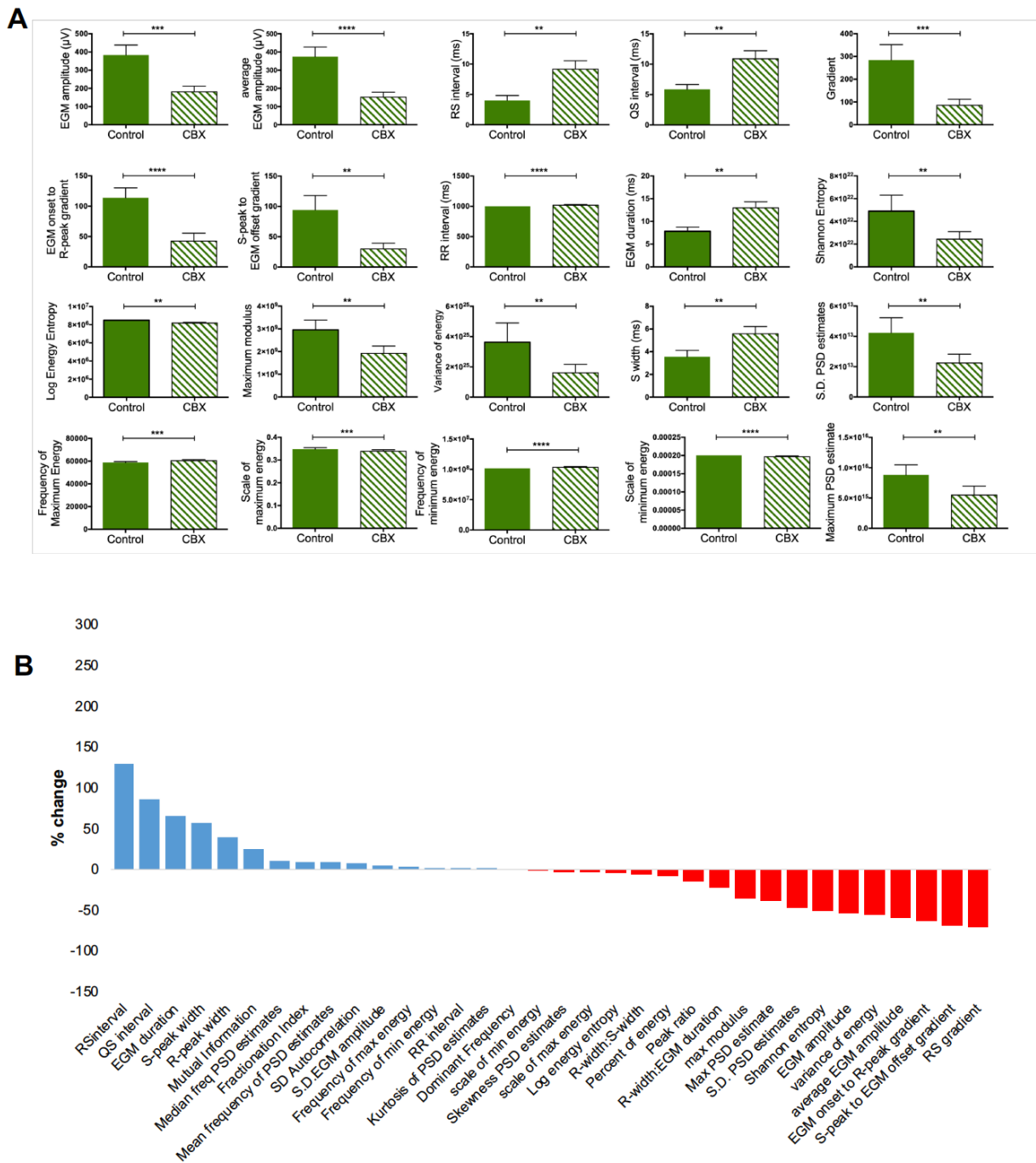


Figure 5.5: EGM morphology modifications due to gap junction uncoupling when $3\mu\text{M}$ CBX were administered on NRVM/+20% fibroblast co-cultures. (A) Features significantly changed due to CBX. Mann-Whitney test analysis. (B) Quantification of EGM morphology changes following the enhancement gap junction uncoupling using $3\mu\text{M}$ CBX on NRVM/20% fibroblast co-cultures when compared to the average control EGM morphology. All bar charts represent mean \pm SEM; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$. (Control: $n = 29$ EGMs; CBX: $n=24$ EGMs).

Gap junction uncoupling in NC cultures using 40 μ M CBX significantly affected 18 EGM features (Figure 5.4A). In this group, there was an equal amount of features that increased, such as fractionation index (control: 1.45 \pm 0.1; CBX: 4.5 \pm 0.3; p <0.0001; mean \pm SEM), which changed by +210% compared to control data, and RS interval (control: 2.2 \pm 0.5ms; CBX: 2.7 \pm 0.3ms; p <0.0001; mean \pm SEM), or decreased, such as the average EGM amplitude (control: 514 \pm 31 μ V; CBX: 98.4 \pm 9.7 μ V; p <0.0001; mean \pm SEM). Regarding the data obtained from co-cultured NRVMs with +20% cultured fibroblasts, 22 features were modified due to gap junction uncoupling using 3 μ M CBX (Figure 5.5A). These also included the average EGM amplitude (control: 374 \pm 53 μ V; CBX: 152 \pm 27 μ V; p <0.0001; mean \pm SEM) and EGM duration (control: 7.8 \pm 0.9ms; CBX: 13 \pm 1.3ms; p = 0.004), but not fractionation index (control: 3.2 \pm 0.3; CBX: 3.5 \pm 0.27; p = 0.35), which was modified in other cell culture groups due to intercellular uncoupling (Figure 5.5A).

5.2.2 Effects on upstroke heterogeneity and fractionation at the single cell level

Dual modality recordings obtained before and after the administration of 15.8 μ M (IC₅₀) CBX were analysed for individual myocytes belonging on the same group of cells on top of an electrode. The purpose of this was the investigation of the possible AP upstroke timing heterogeneity, alongside EGM fractionation. It was observed that the time point of the upstroke peak was similar for all traces derived from the same electrode at the baseline state (Figure 5.6C). However, the timing of the upstroke peak was spread across an electrode following gap junction uncoupling (Figure 5.6D). The larger distribution of the upstroke peak time point was concurrent with an increase of EGM fractionation. Linear regression analysis using data obtained from 7 cells before and after the addition of CBX showed that fractionation was directly correlated to the time difference (ΔT) between upstroke peak time points of two cells located across an electrode ($r^2=0.74$, $S_{y.x}=0.52$, slope=0.57, p <0.0001, $n=14$ electrodes - with approximately 3 cells tested on each electrode) (Figure 5.6B).

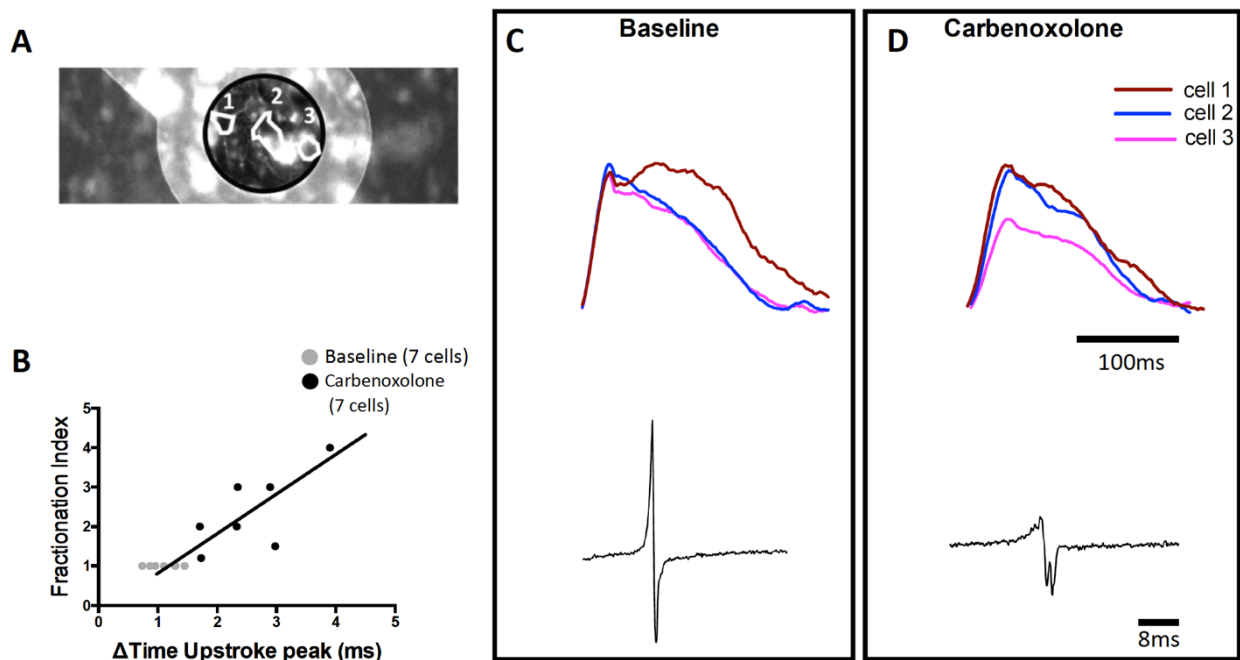


Figure 5.6: Single cell level correlation of action potential dispersion and EGM fractionation. (A) NRVM monolayer obtained by optical imaging, where the outlines of three neighbouring myocytes are highlighted. The cells are placed in a series across the electrode. (B) Linear regression graph presenting the correlation between fractionation and the time dispersion of upstroke peaks for cells in positions 1 and 3 on the electrode before and after the administration of $15.8\mu\text{M}$ CBX ($r^2=0.74$, $Sy.x=0.52$, $\text{slope}=0.57$, $p<0.0001$). (C) (Top) Action potentials obtained from three single myocytes located in positions 1-3, as shown in (A), at the baseline state. Upstroke duration is the same for all cells. (Bottom) EGM obtained from the same group of cells before gap junction uncoupling. (D) (Top) Superimposed action potentials obtained from the same cells as in (C) after the addition of CBX. Upstroke duration increases alongside the AP propagation. (Bottom) The corresponding EGM showed decreased amplitude, increased fractionation and prolonged EGM duration due to gap junction uncoupling. Adapted from Chowdhury, Tzortzis *et al.* (40).

5.2.3 The effects of gap junction uncoupling on human ventricular slice-derived EGM morphology

Simultaneous MEA and optical mapping recordings were obtained between 0.2Hz to 1Hz from *ex vivo* adult human cardiac slices. These slices were derived from the epicardial (n=8) and the endocardial (n=8) left ventricular wall. The intact slice results reconfirmed those of the cell monolayer models about the direct relationship between APD_{90} and FPD both in the epicardial and the endocardial samples at the baseline status ($r^2=0.64$, $p<0.0001$) (Figure 5.7A-B). In addition, the effects of gap junction uncoupling by CBX were studied. Dose response experiments for CBX were not carried out on slices and the same IC_{50} of CBX, as the one used for experiments on myocyte only cell cultures, was used which was $15.8\mu\text{M}$. Cx43 is the main connexin expressed in adult human ventricles and cultured

NRVMs (316,317). The slices used for this purpose (n=4) were derived from the endocardial side of the same left ventricular sample. Adjacent slices were alternatively used as control samples and CBX treatment (Figure 5.7A). The collagen coverage, as a marker of the size of fibrotic regions, was $22.6 \pm 0.4\%$ of the total area in slices used as control (n = 2) and $23.3 \pm 2\%$ in slices perfused with CBX (n = 2) (Figure 5.7B). There was no significant difference in the amount of collagen area between the groups of slices. The analysis showed that gap junction uncoupling modified APD₉₀ (control: 0.2Hz 883.6 ± 77 ms, 1Hz 522.6 ± 194.6 ms; CBX: 0.2Hz 481.3 ± 122.8 ms, 1Hz 445.5 ± 50.6 ms; n=8; p<0.0001; mean \pm SEM) and restitution was seen due to stimulation frequency (p<0.01). There was also a significant interaction between these factors (p<0.05, ordinary two-way ANOVA) (Figure 5.8D). FPD was affected by pacing frequency (control: 0.2Hz 603.9 ± 33.2 ms, 1Hz 522.5 ± 135.1 ms; CBX: 0.2Hz 622.5 ± 46.3 ms, 1Hz 455.7 ± 45 ms; n=4; mean \pm SEM), but not by gap junction uncoupling (Figure 5.8E). A linear relationship between APD₉₀ and FPD was observed ($r^2=0.64$, slope = 0.39, p<0.0001). This relationship was not significantly different between the before and after gap junction uncoupling data (control: $r^2=0.63$, slope = 0.51; CBX: $r^2=0.77$, slope = 0.68; p = 0.24; mean \pm SEM) (Figure 5.8F). Linear regression analysis showed that there was no correlation between conduction velocity and EGM duration ($r^2=0.14$, slope = -0.14 ± 0.08 , p=0.1) (Figure 5.7G), despite the decrease in conduction velocity (control: 12.92 ± 1.5 cm/s; CBX: 7.9 ± 0.7 cm/s; p=0.008; mean \pm SEM). EGM duration remained stable (control: 39.5 ± 4 ms; CBX: 33.2 ± 2.7 ms; p=0.23; mean \pm SEM).

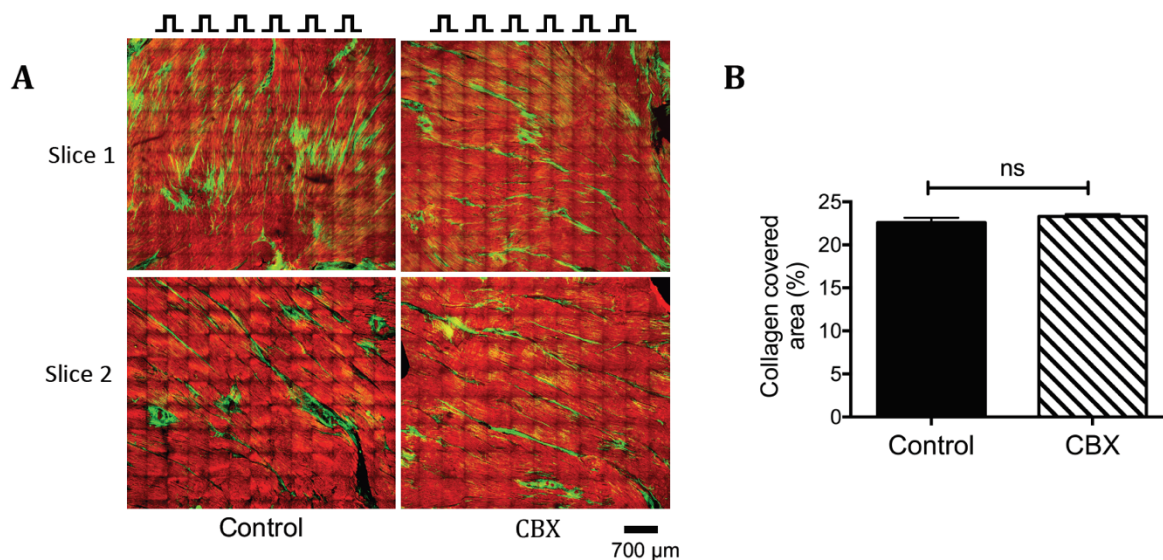


Figure 5.7: Quantification of collagen coverage on slices derived from human left ventricle. (A) Adjacent slices stimulated for obtaining EGM data before (control) and after (CBX) the perfusion with CBX. Two slices were used per condition. (B) The bar graph presents no difference on the size of collagen coverage between slices used in experiments (Control: n=2; CBX: n=2). Mann-Whitney t-test. Bar charts represent mean \pm SEM. ns = non-significant.

A thorough analysis of EGM morphology with or without gap junctional block using 15.8 μ M CBX showed that 25 features were affected (Figure 5.9A). These were features belonging to all types of EGM analysis, i.e. time-domain, frequency-domain, time-frequency. Voltage was increased (control: 2423 \pm 272.5 μ V, n=55 EGMs; CBX: 3429 \pm 328 μ V, n=40 EGMs; 2 slices each; p = 0.006; mean \pm SEM), EGM duration was abbreviated (control: 40.1 \pm 0.9ms; CBX: 17.6 \pm 1ms; p<0.0001; mean \pm SEM) and fractionation index was also reduced (control: 4.27 \pm 0.4; CBX: 2.6 \pm 0.3; p<0.0001; mean \pm SEM). Gap junction uncoupling affected EGM morphology the opposite way when CBX was applied on NRVM cultures, as presented in section 5.2.1. The EGMs used in this analysis were collected throughout the slice surface and there was no relationship with specific scarred or non-scarred regions. Comparing the average EGM morphology following CBX administration with the control morphology, it can be seen that half of features increased and the rest decreased (Figure 5.9B). Dominant frequency was the only characteristic that remained unaffected.

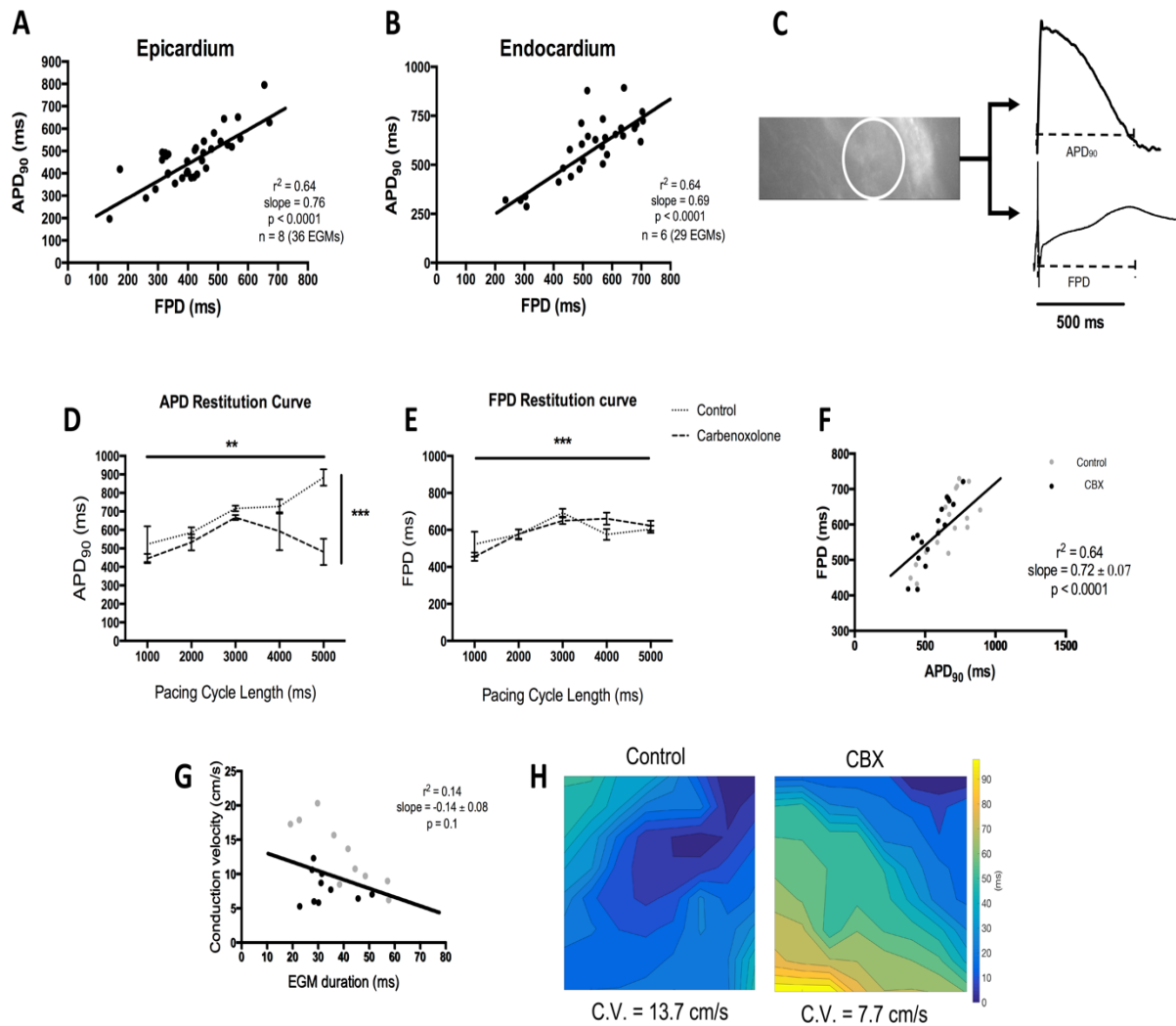


Figure 5.8: (A-B) Correlation between APD_{90} and FPD in human ventricular slices obtained from the epicardium ($n=36$ EGMs, obtained from 8 slices in total) (A) and the endocardium ($n=29$ EGMs obtained from 6 slices in total) (B). MEA and optical mapping recordings were obtained simultaneously, and the results show that APD_{90} and FPD have a linear relationship (epicardium: $r^2=0.64$; slope= 0.76 ; $p<0.0001$, endocardium: $r^2=0.64$; slope= 0.69 ; $p<0.0001$). (C) Action potential (top) obtained from human epicardial slice and the concurrent EGM (bottom), both derived from the same electrode on the left. (D-E) Restitution curves showing that APD_{90} changed significantly due to the pacing rate ($p<0.01$) and the gap junction uncoupling ($p<0.001$) (D), while FPD was affected by the pacing rate only ($p<0.001$) (E). APD_{90} and FPD plots were made using manually annotated data. (F) Linear regression analysis for the relationship between APD_{90} and FPD before and after the administration of CBX. (G) No correlation was observed between conduction velocity and EGM duration, as measured before and after gap junction uncoupling ($r^2=0.14$; slope= -0.14 ± 0.08 ; $p=0.1$). (H) Isochronal maps of the wavefront propagation before and after the CBX administration in adjacent tissue slices. All bar charts represent mean \pm SEM; ** $p<0.01$; *** $p<0.001$.

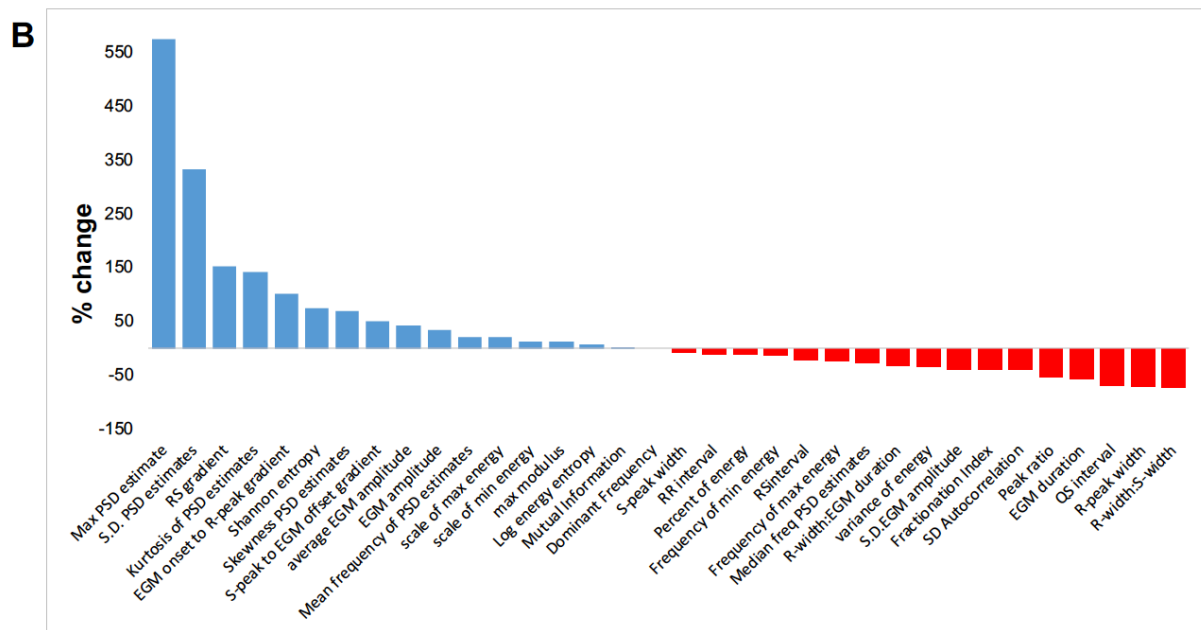
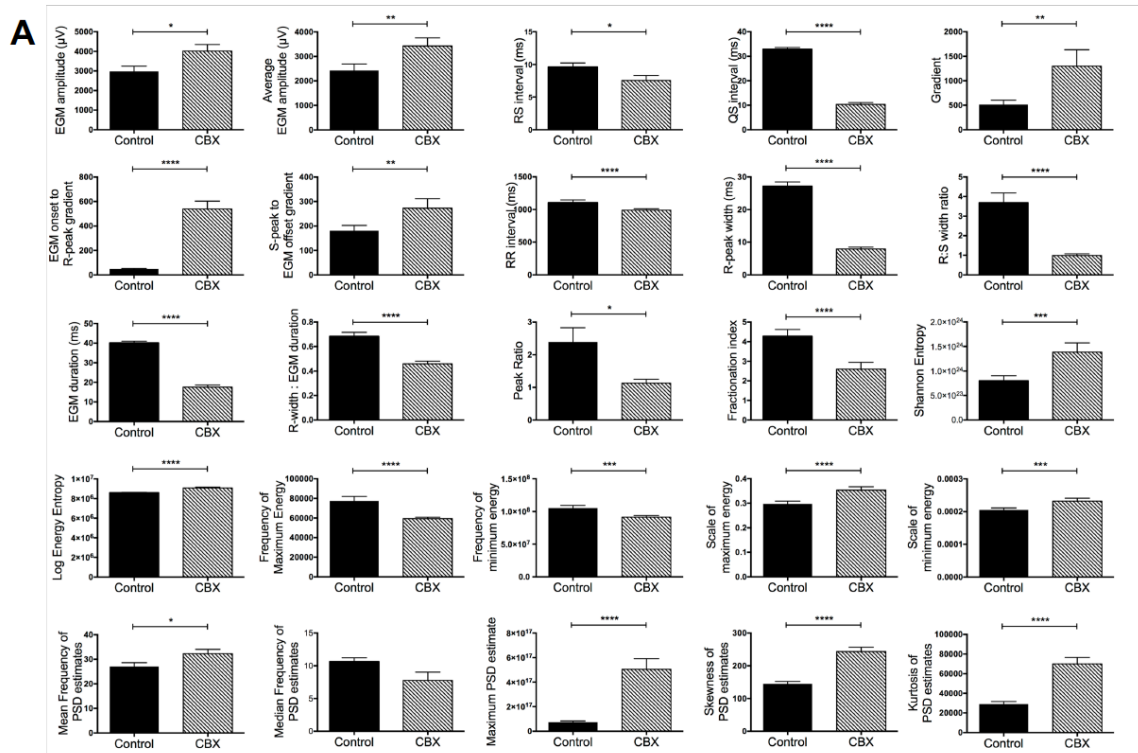


Figure 5.9: (A) Modified EGM features due to gap junction uncoupling when human left ventricle tissue slices were perfused with $15.8\mu\text{M}$ CBX. Mann-Whitney test analysis. (B) Quantification of EGM morphology changes following gap junction uncoupling using $15.8\mu\text{M}$ CBX on human left ventricle tissue slices when compared to the average control EGM morphology. All bar charts represent mean \pm SEM; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$. (Control: 2 slices, $n = 55$ EGMs; CBX: 2 slices, $n=40$ EGMs).

5.3 Discussion

The main findings of this part of the thesis are that CBX shortened both APD₉₀ and FPD in MO cell cultures and tissue slices, while the direct relationship between them was maintained following gap junctional block. EGM morphology changed dramatically in all experimental models, including new features in addition to the traditional ones of voltage and fractionation. However, despite the fact that CBX led to conduction slowing in tissue slices, the EGM feature alterations were made towards the opposite direction than the changes that were observed *in vitro*. Based on these results, the initial hypotheses can be accepted apart from the one stating that gap junction uncoupling should affect EGM features on a similar way in *in vitro* and *ex vivo* models.

5.3.1 *In vitro* gap junction uncoupling influence on EGM morphology

CBX, a well investigated anti-ulcer drug, has been also studied for its ability to block gap junctions (318). Gap junction uncoupling due to CBX leads to conduction slowing and increased fractionation, as shown from data obtained from the left ventricle of hypertrophic cardiomyopathy patients (15), right atrial and right ventricular activation mapping on humans in sinus rhythm (16) and isolated rabbit left ventricular myocytes (12). The findings from MO and NC cell cultures were in agreement with this. However, no increased fractionation was observed in NRVM/+20% FB co-cultures. A general observation that has to be acknowledged though was the fact that the blockade of intercellular electrical coupling using CBX changed all computed features.

The optical mapping data presented significantly abbreviated APD₉₀ in MO cultures, while this was not the case in NC and +20% FB cell cultures, where both characteristics remained stable. There is limited understanding about the effects of gap junction uncoupling on AP morphology and duration. It has been suggested that CBX can either prolong APD in perfused adult rat hearts (319) or it had no effect on the AP of rabbit ventricular myocytes (314,320). Cx40 and Cx43 knock-out experiments on neonatal murine ventricular myocytes caused gap junction uncoupling and APD shortening (321). It could be suspected that these were species- and age-specific results. Since there is poor knowledge on the effects of gap junctional block on NRVMs, it can be suggested that this type of functional manifestation shortens APD₉₀. It could be suspected that the CBX effects are species specific, but without further mechanistic explanations to be given, as the effects of gap junction uncoupling on AP are unclear.

Data obtained from concurrent optical mapping and MEA recordings were analysed not only in groups of cells, but also at the single cell level. AP and EGM morphology were studied before and after the administration of CBX on individual cells. It was observed that the gap junction uncoupling by CBX leads to APD dispersion. These findings agree with previous studies which showed that the *in vitro* intercellular uncoupling can cause intercellular activation delays and multiphasic optically recorded AP upstrokes (41,322). The findings presented in this chapter showed that there was no fractionation before the administration of CBX, but this was detected following gap junction uncoupling. There was an activation delay in the experimental model which correlated with fractionation. Gap junction uncoupling is associated with the subsequent discontinuity in AP propagation (320), which may explain the fractionation identified on the data. As it was also presented in Chowdhury *et al.* (40), the ΔT of action potential upstroke was correlated with EGM fractionation proposing that the discontinuous propagation in cellular preparations, where there are normally linear propagating wave fronts, can lead to EGM fractionation.

5.3.2 *Ex vivo* gap junctional block effects

Myocardial slices obtained from adult human ventricles have been an increasingly used model in cardiac electrophysiology, because they retain the multicellularity, complex architecture and electrophysiology of adult cardiac tissue compared to *in vitro* experimental models. Their thin structure allows oxygen and metabolic substrate diffusion into the inner cells maintaining the tissue viability *in vitro* (20). Tissue slices have been produced from a variety of mammals, such as rabbits (323), guinea pigs (323,324), mice (325,326), rats (326,327), dogs (20,328) and human biopsies (20,40,328,329). These have been studied for the electrophysiology as well as for fibrosis.

Epicardial left ventricular slices derived from human failing heart were used in this study. Tissue anisotropy has to be taken on account though during the analysis of these data. The fibre orientation, which is observed in Figure 5.7, seems to be variable between samples, even though it was not quantified. It is well established that the difference in conduction velocity between the longitudinal and transverse direction in a bundle of parallel muscle fibres is a pro-arrhythmic factor. Conduction is faster in the longitudinal compared to the transverse direction and thus, conduction velocity is considered to be anisotropic (330). Apart from fibre orientation, previous studies associated conduction velocity with EGM amplitude and fractionation, which can be a weak correlation, as there are counter-intuitive findings (331,332). However, there is the general notion that transverse conduction to myocardial fibres is related to increased fractionation (332). Since fibre orientation on tissue slices could not be identified

using MRI, it would be arbitrary to make any conclusions about tissue anisotropy. Therefore, it can only be assumed that it is an additional factor affecting EGM features.

A dual modality system combining optical mapping and MEA recordings (40) was used for concurrent AP and EGM recordings from the same area of a slice. It was shown that a direct correlation between FPD and APD₉₀ exists, even though FPD showed a trend towards abbreviation while APD₉₀ was significantly shortened. The linear relationship between these characteristics was present before and after gap junction uncoupling. This further validated the results obtained from our *in vitro* experiments in a more complex *ex vivo* intact tissue model and demonstrated the translational benefits of combining optical mapping with MEA recordings.

Regarding the automated EGM morphology analysis, a large number of features were modified following gap junction uncoupling. Among these changes, it was observed that EGM amplitude increased while EGM duration and fractionation both decreased. The decreased intercellular coupling leads to conduction slowing, which was also observed in the experiments presented in this chapter. That means action potential needs more time to pass across an electrode and therefore, it can be concluded that gap junction uncoupling results to prolonged EGM duration (40). In addition, as EGM amplitude is associated to the degree of depolarisation, the amplitude is theoretically expected to be lower after gap junction uncoupling has occurred, due to the reduced ion flux and the discontinuous action potential propagation (40). The fact that conduction slowing was caused after the addition of CBX is a persuasive fact that gap junction uncoupling occurred. It has been shown that CBX has the ability to reduce conductance without affecting excitability, even though the exact molecular mechanism is not clear (314). A study on isolated ventricular and atrial rabbit myocytes showed that CBX can lead to Cx43 de-phosphorylation and it does not affect any ion currents (314). In addition, the particular slices which were perfused with CBX solution were derived from a failing heart diagnosed with DCM. DCM has been linked to reduced ventricular Cx43 expression levels in ventricles, gap junction heterogeneity and re-organisation, probably due to the decreased Cx43 protein levels, leading to conduction slowing and a propensity to arrhythmias (333,334). However, the counter-intuitive effects of CBX in the experimental model of the present study cannot be explained by these changes. An explanation to the increased EGM amplitude could be the increased intracellular resistance due to gap junction blockade. This effect would limit the spread of local circuits generating AP. These local circuits normally discharge the cell membrane capacitance in the unexcited cell regions in front of an AP wavefront depolarising the membrane potential to threshold in this region. Thus, the increased charge density due to the increased intracellular resistance would lead to faster depolarisation of the unexcited cell membrane near the AP wavefront, as the local charge density is greater. As a consequence, the conduction velocity may be diminished. This overall change manifests itself as a greater dV/dt for the

AP upstroke and therefore, increased EGM amplitude in extracellular recordings, such as the ones conducted for this research. This theory is supported by the research carried out by Rohr & Kleber (41, 478) and Jabr *et al.* (479).

By observing the slices used in this study, it can be easily noticed that they had a discontinuous structure due to the presence of patchy and interstitial fibrosis. Fibrotic areas were identified by the fluorescence signal emitted from collagen I molecules and by applying SHG microscopy. According to the theory developed by Rohr *et al.* (335), the modification of intercellular electrical coupling in a sample with well coupled cells, which also contains structural heterogeneity and discontinuities, can have completely different effects on action potential propagation than what normally happens in linear or continuous excitable structures (43,335). The action potential propagation can fail in a large mass of tissue receiving a limited amount of depolarising current from a small group of cells. Gap junction uncoupling under these conditions may overcome the conduction block and the tissue may bear pro-arrhythmic effects since there will still be slow conduction (43,335). The reduction of intercellular coupling should be of magnitude sufficient to decrease coupling (43). Such a protective response, which is important for protecting conduction in a diseased ventricle, could also explain the counter effects observed in this work. However, further investigation on this matter has to be carried out, as a limited number of samples were used for this study.

5.4 Conclusions

Gap junction uncoupling was shown to reduce conduction velocity *in vitro* and *ex vivo* and lead to dramatic changes on EGM morphology. Since the effects of CBX on AP morphology had been poorly studied, it was interestingly shown that the use of a gap junction uncoupler reduces AP duration, due to its effects on the repolarisation phase, as also confirmed by adjacent changes on FPD. In corroboration with our findings from gap junction uncoupling on cell cultures, we have also demonstrated that the blockade of intercellular coupling leads to inverse and counterintuitive results on *ex vivo* data. These were shown to be mainly limited to time-domain analysis features. An explanation for this could be the high amount of fibrosis, as quantified by the size of collagen coverage, which is responsible for a compensation mechanism, when intercellular uncoupling concurrently occurs in highly heterogeneous tissue regions. The findings of this section will provide a better understanding during the development of a predictive model presented in later chapters of this thesis. Features selected during the development of a predictive model can become biologically meaningful and the reasons behind the better or worse predictability of gap junction uncoupling, when compared to other types of cellular electrophysiology abnormalities, can be revealed.

Chapter 6

Evaluation of EGM morphology modifications
due to structural remodelling

6.1 Introduction

6.1.1 Structural remodelling affects cardiac electrophysiology

In addition to the functional parameters of cardiac electrophysiology, i.e. ion currents and intercellular communication, presented in Chapters 4 and 5, the focus of the present section is the role of structural factors on influencing the EGM morphology. Our knowledge about the effects of fibrosis on morphological characteristics comes from clinical studies, but also animal and computational studies. CFAEs have been thought to represent continuous re-entry or fibrillation waves into a certain area of the left atrium. These areas have been targeted as part of AF substrate remodelling, and despite the fact this strategy was initially adopted because of promising outcomes (106), its efficiency is questionable due to poor reproducibility of clinical outcomes across other groups (336,337). The underlying pathophysiological mechanism of atrial EGM fractionation is inadequately understood, and this may be part of the divergent and limited success rates of catheter ablation strategies targeting CFAEs.

Clinical success has been improved when targeting low voltage endocardial EGMs, which are often considered to be a surrogate marker of scar (338,339). Structural remodelling that occurs in AF, leading to extended scar regions, creates a substrate for re-entry and conduction slowing sustaining rotor activity. Localised rotational activity detected with multipolar endocardial mapping tends to co-localise with areas of low voltage EGMs (340). However, even this theory has inconsistencies. The *in vivo* detection of scar areas is usually carried out using LGE-MRI and there is the general notion that a chaotic electrical activity and fractionation are associated with fibrotic change and translate to reduced EGM voltage. However, there is a consistent reported inverse and counterintuitive relationship between EGM characteristics (voltage and fractionation) and atrial fibrosis in AF patients (332). Jadidi *et al.* (7) showed that the majority of CFAE sites detected in their clinical study were not related to fibrotic areas, but healthy atrial tissue with unidentifiable fibrosis by MRI. There was also a proportion of CFAEs detected in areas of patchy fibrosis, where the electrical activity was similar to non-fibrotic areas, and a smaller proportion of CFAE sites were detected in the border zone of or within dense fibrosis regions (7).

The fact that EGM morphology cannot be easily correlated to clinical phenotypes has led to intense study of the underlying mechanisms at a cellular level. Cardiac scar tissue is a dynamic living structure consisting of cellular and acellular components (341,342). The extracellular matrix (ECM), which has increased presence in scar, is interlaced with phenotypically various types of cells: fibroblasts and interstitial fibroblast-like cells, endothelial cells, surviving cardiac myocytes, immune cells, neurons

and adipocytes (343,344). The impact of scar on electrical activity is still a topic under investigation (345). Fibrosis exhibits variable degrees of density leading to the separation of myocardium strands and forcing excitation waves to take anisotropic, circuitous paths (33). This is one of the suggested mechanisms of arrhythmogenesis. However, it is still unclear how scar is involved in either the active generation or the passive maintenance of abnormal electrical conduction episodes (32). It has been shown that fibroblasts interact with myocytes via structural, electrical and paracrine pathways, which have been assessed for their role in arrhythmogenesis. Pedrotty *et al.* (346) have suggested a mechanism according to which NRVM growing in cell culture medium derived from cultured of neonatal rat fibroblasts show prolonged APD and significant conduction slowing. These electrophysiological changes were proposed to be caused by paracrine factors which are normally secreted by cardiac fibroblasts. This agrees with findings from Guo *et al.* (347), who reported that cardiac myocytes provided with cultured fibroblast-derived medium can result to APD prolongation and downregulation of the $K_v4.3$ subunit gene, which is implicated to the I_{to} current.

Repolarisation dispersion plays a key role for the initiation and maintenance of arrhythmogenesis and combined with conduction slowing promotes re-entrant tachycardia initiation (301). However, it seems that the duration of repolarisation depends on the presence of scar and scar border zone in cases of heart failure, and thus APD becomes heterogeneous. In addition to the *in vitro* research, data obtained from VT patients resistant to medical therapy or recurrent implantable cardioverter-defibrillator shocks showed that the activation recovery interval (ARI) was prolonged in areas of myocardial scar compared to areas with normal voltage (348). The transmural ARI was also decreased in scar regions. These effects were focused on scar areas and they were not present in the rest of tissue. An explanation for these findings could be the downregulation of Cx43 expression in scar regions or the upregulation of I_{Ca} , which is an adaptive mechanism for Ca^{2+} overloading in a cell for contractility enhancement (348). Also, APD prolongation and dispersion of the refractory period have been detected in remodelled LV of human failing hearts (349). Thus, it would not be surprising if scar tissue consequently affected the cardiac AP recorded in the experimental models presented in this chapter.

Apart from the production of ECM components and enzymes, regulated by signalling molecules with autocrine and paracrine effects (31,32,278,350), fibroblasts also express ion channels and stretch-activated channels (351). It has been suggested that fibroblasts can be electrically coupled to cardiac myocytes via gap junctions (32,80,350–352). Thus, fibroblasts are part of tissue conductivity, they alter cardiac electrophysiology and they can possibly act as passive inexcitable conductors over a maximum distance of 300 μ m (31). Therefore, this mechanism could explain whether fibroblasts are directly involved in arrhythmogenesis and if they can influence EGM morphology.

6.1.2 The role of myofibroblasts

Myofibroblasts are phenotypically transformed fibroblasts and express α -smooth muscle actin (α SMA), as well as integrins, fibronectin and connexins 43 and 45 (Cx43/45). α SMA is not found in fibroblasts and allows contractility similar to that of smooth muscle cells. Myofibroblast proliferation and migration mainly occurs in myocardial remodelling in dilated, ischemic, hypertrophic, hypertensive cardiomyopathies (353) and arrhythmogenic right ventricular cardiomyopathy (354). Myofibroblasts also contribute in atrial fibrosis, as part of the maladaptive atrial response in AF (355). The enhanced myofibroblast proliferation under these situations is accompanied by collagenous matrix accumulation leading to fibrosis (353). The primary drivers of this phenotypic shift of fibroblasts are altered mechanical activation and inflammatory cytokines, specifically the transforming growth factor-beta1 (TGF- β 1) and platelet-derived growth factor (PDGF) (356,357).

The myofibroblast-to-myocyte interaction has been reported in multiple *in vitro* studies. Myofibroblasts are not excitable, like fibroblasts, and are not directly involved in conduction in the post-myocardial infarction heart. They are believed to block conduction, due to their presence among myocytes, thus increasing the distance between adjacent myocytes and decreasing the intercellular coupling via gap junctions. However, it has been suggested that there is gap junction uncoupling between myofibroblasts and myocytes (358). This is supported by the expression of Cx43 and Cx45 in myofibroblasts and the presence of this intercellular coupling between myofibroblasts (278) and neonatal myocytes when they are cultured together. It is still unclear though if this coupling exists *in vivo* and how important it is for cardiac electrophysiology (358). Myofibroblasts also have relatively high membrane resistance with no significant difference in the electrical activity between two coupled myocytes and a myofibroblast coupled with another cardiac myocyte (359). On the other hand, fibroblasts also couple with myocytes *in vitro* and *in situ*, but connexins appear to be located at punctuate contacts. These are orders of magnitude smaller intercellular couplings compared to the ones in myocyte-myocyte couplings (360). However, this type of intercellular communication is able to depolarise the myocyte membrane potential sufficiently in order to cause automaticity and changes in conduction velocity, mainly when they fill interstitial spaces and not when fibroblasts are inserted on top of myocytes. Fibroblasts can also change conduction by creating zig-zag conduction pathways and they break the intercellular gap junction coupling between myocytes (361).

As it appears from what has been presented so far in this chapter, the intercellular interaction and a balance of structural characteristics are essential for the normal AP propagation and the sustainability of health cardiac electrophysiology. Thus, the rest of this chapter will investigate the effects of variable

structural characteristics on EGM morphology. For this purpose, different experimental models were used. The hypotheses to be investigated were:

1. The cellular scar, added in cell cultures, is responsible for prolonged APD, and subsequently FPD.
2. Myofibroblasts can alter EGM parameters.
3. Fibrosis burden in tissue slices correlates inversely with APD and FPD.
4. Cellular scar and tissue slice fibrosis modify not only the conventional features of voltage and fractionation, but also more complex EGM parameters.
5. The amount of cellular scar and tissue fibrosis is related to the level of feature modifications.
6. *In vivo* scarred regions are responsible for changing the EGM morphology compared to non-scarred areas.

The aim of this chapter was to address these hypotheses by: (1) focusing on the effects of fibroblasts, which is the cellular component of scar, and fibrosis using *ex vivo* and *in vivo* EGM data, (2) focusing on the electrophysiological properties of myofibroblasts and how these affect EGM morphology, (3) carrying out concurrent *in vitro* and *ex vivo* optical mapping and EGM recordings, and (4) fully characterising the EGM morphology that occurs in each experimental model. The information obtained from this body of work, after answering the above hypotheses, will help better understand the features selected during the supervised machine learning process for predicting the conditions presented in this chapter.

6.2 Results

6.2.1 *In vitro* effects of cellular scar on EGM morphology

Control EGM data, collected at 1Hz from the different types of cell cultures presented in section 2.2.1, were analysed. The percentage of total area of cell cultures covered by vimentin was considered to correspond to the non-myocyte cellular component of the cell culture. Vimentin is used as a fibroblast marker by convention, even though it is also found in myofibroblasts. Our results showed that MO cell cultures had $35.7 \pm 2\%$ coverage ($n = 8$ MEAs; mean \pm SEM), NC cell cultures had $63.9 \pm 4.5\%$ ($n = 28$ MEAs), the co-cultures of NRVM with +20% of cultured fibroblasts had $72.7 \pm 3.5\%$ coverage ($n = 16$ MEAs) and the co-cultures of NRVM with +40% of cultured fibroblasts had $76.2 \pm 3.4\%$ coverage ($n =$

16 MEAs) (Figure 6.1). MO cultures had significantly less vimentin signal, and subsequently fibroblasts, than the NC cell monolayers ($p < 0.01$), the +20% FB cell cultures ($p < 0.01$) and the +40% FB monolayers ($p < 0.0001$) (Figure 6.1).

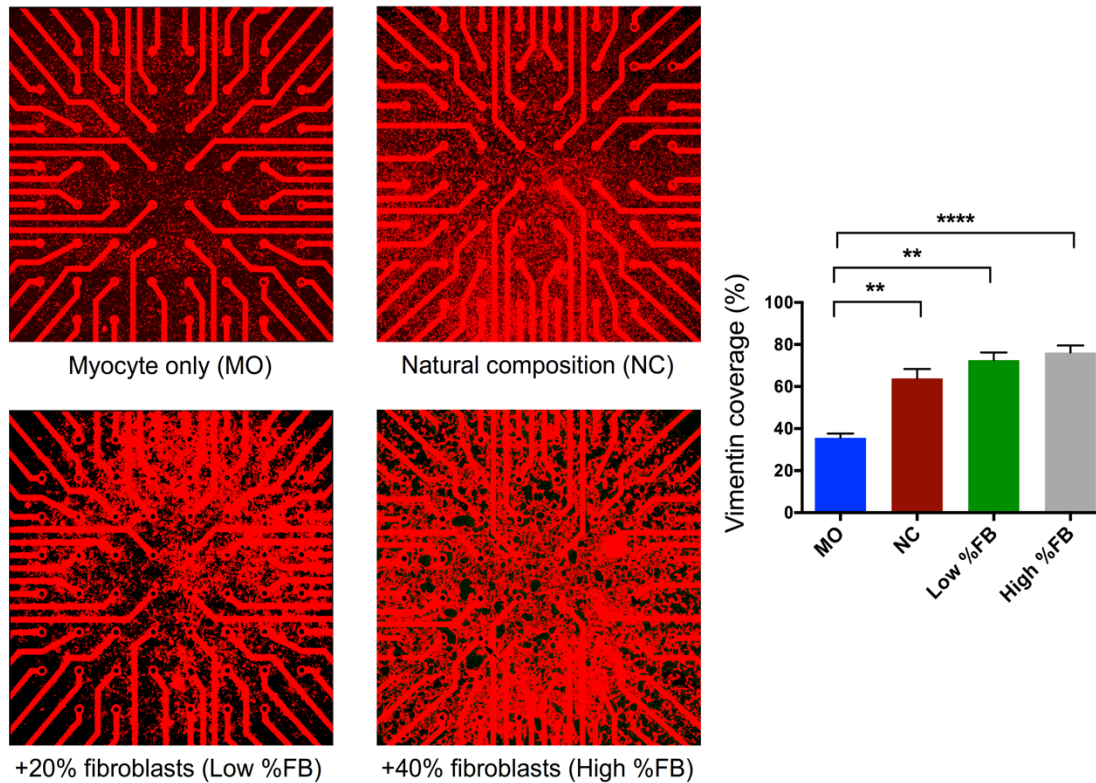


Figure 6.1: Immunostaining for fibroblast quantification. (*Left*) MEA plates immunostained for vimentin following electrical recordings to accurately quantify fibroblast burden. Representative examples of different burdens of fibroblast in each group, the red representing the vimentinstaining of fibroblasts. (*Right*) Quantification of fibroblast burden in each experimental group as a percentage of total area of MEA plate covered by fibroblasts. Myocyte only (MO) ($n=8$), Natural Composition (NC) ($n=28$), +20%FB ($n=16$), +40%FB ($n=16$). Kruskal-Wallis test analysis. All bar charts represent mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Concurrent optical mapping and MEA experiments were carried out on the different types of cell culture preparations, as shown in section 4.2.4. It was observed that APD₉₀ was increased in cell cultures with higher amount of fibroblasts, compared to MO monolayers, and it was also dependent on the amount of fibroblasts (MO: 167.3 \pm 7.8ms; NC: 211.2 \pm 9.7ms; +20% FB: 231.3 \pm 19.6ms; +40%FB: 209 \pm 8.8ms; $p = 0.002$; mean \pm SEM) with significant difference between MO and NC ($p < 0.05$) and between MO and +20%FB ($p < 0.05$) (Figure 6.2A). A similar pattern of FPD manifestation was also observed (MO: 151.2 \pm 7.4ms; NC: 205 \pm 9.2ms; +20%FB: 220 \pm 18ms; +40%FB: 202.3 \pm 9.6ms; $p < 0.0001$; mean \pm SEM) with significant change between MO and NC cultures ($p < 0.0001$) and MO and +20%FB monolayers

($p < 0.05$) (Figure 6.2B). Linear regression analysis showed a direct correlation between APD_{90} and FPD measured on data obtained from all cellular experimental models ($r^2 = 0.85$, $Sy.x=18.2$, slope = 1 ± 0.05 , $p < 0.0001$, $n=59$) and there was a linear relationship even within each group (MO: $r^2 = 0.78$, $Sy.x=20.1$, slope = 0.99 ± 0.1 , $p < 0.0001$, $n=34$; NC: $r^2 = 0.81$, $Sy.x=17.8$, slope = 0.9 ± 0.1 , $p < 0.0001$, $n=16$; +20% FB: $r^2 = 0.95$, $Sy.x=9.9$, slope = 0.89 ± 0.1 , $p = 0.004$, $n=5$) (Figure 6.2C). However, there was no strong APD_{90} /FPD relationship within the +40% FB cell cultures ($r^2 = 0.72$, $Sy.x=13$, slope = 0.92 ± 0.3 , $p = 0.07$, $n=5$).

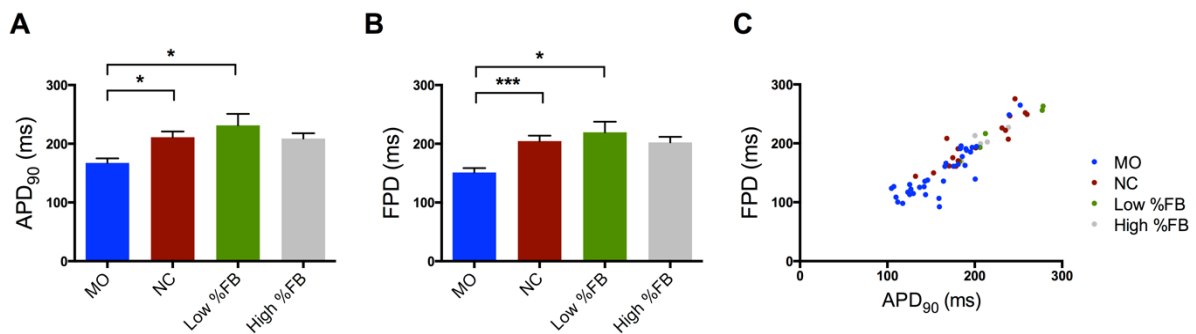


Figure 6.2: Effects of fibroblast burden on APD_{90} and FPD. (C) Linear regression analysis for the relationship between APD_{90} and FPD. The direct relationship between these features was maintained in all types of cell culture. All bar charts represent mean \pm SEM; * $p < 0.05$; *** $p < 0.001$. APD_{90} and FPD plots were made using manually annotated data.

Following the automated EGM morphology detection and quantification for measuring 35 features, it was found that 9 of them were significantly affected by the type of cell culture (Figure 6.3). The EGM amplitude was significantly higher for the NC and +20% FB cell cultures compared to the MO cultures (MO: $546.2 \pm 70 \mu V$; NC: $1312 \pm 199 \mu V$; +20% FB: $1217 \pm 129 \mu V$; $p = 0.005$; mean \pm SEM). Regarding fractionation, that was the same among the MO, NC and +20% FB groups, but significantly higher for the +40% FB group when compared to the MO cell cultures (MO: 1.04 ± 0.01 ; +40% FB: 2.3 ± 0.7 ; $p = 0.016$; mean \pm SEM). Apart from the time-domain features, it was observed that maximum modulus was significantly increased at +20% FB and +40% FB cell cultures compared to the MO cultures (MO: $1.14e+009 \pm 3.6e+008$; +20% FB: $1.67e+009 \pm 2.1e+008$; High %FB: $1.75e+009 \pm 2.5e+008$; $p = 0.006$; mean \pm SEM). Among other features the kurtosis and skewness of PSD estimates were significantly reduced in +40% FB data compared to the MO and +20% FB EGMs (Skewness – MO: 242.2 ± 17 ; +20% FB: 245 ± 11.5 ; +40% FB: 160.7 ± 13 ; MO/+40%FB – $p < 0.01$; +20%FB/+40%FB –

$p < 0.05$ / Kurtosis – MO: $71,197 \pm 6940$; +20%FB: $70,511 \pm 5425$; +40%FB: $34,639 \pm 5141$; MO/+40%FB – $p < 0.05$; +20%FB/+40%FB – $p < 0.05$; mean \pm SEM) (Figure 6.3).

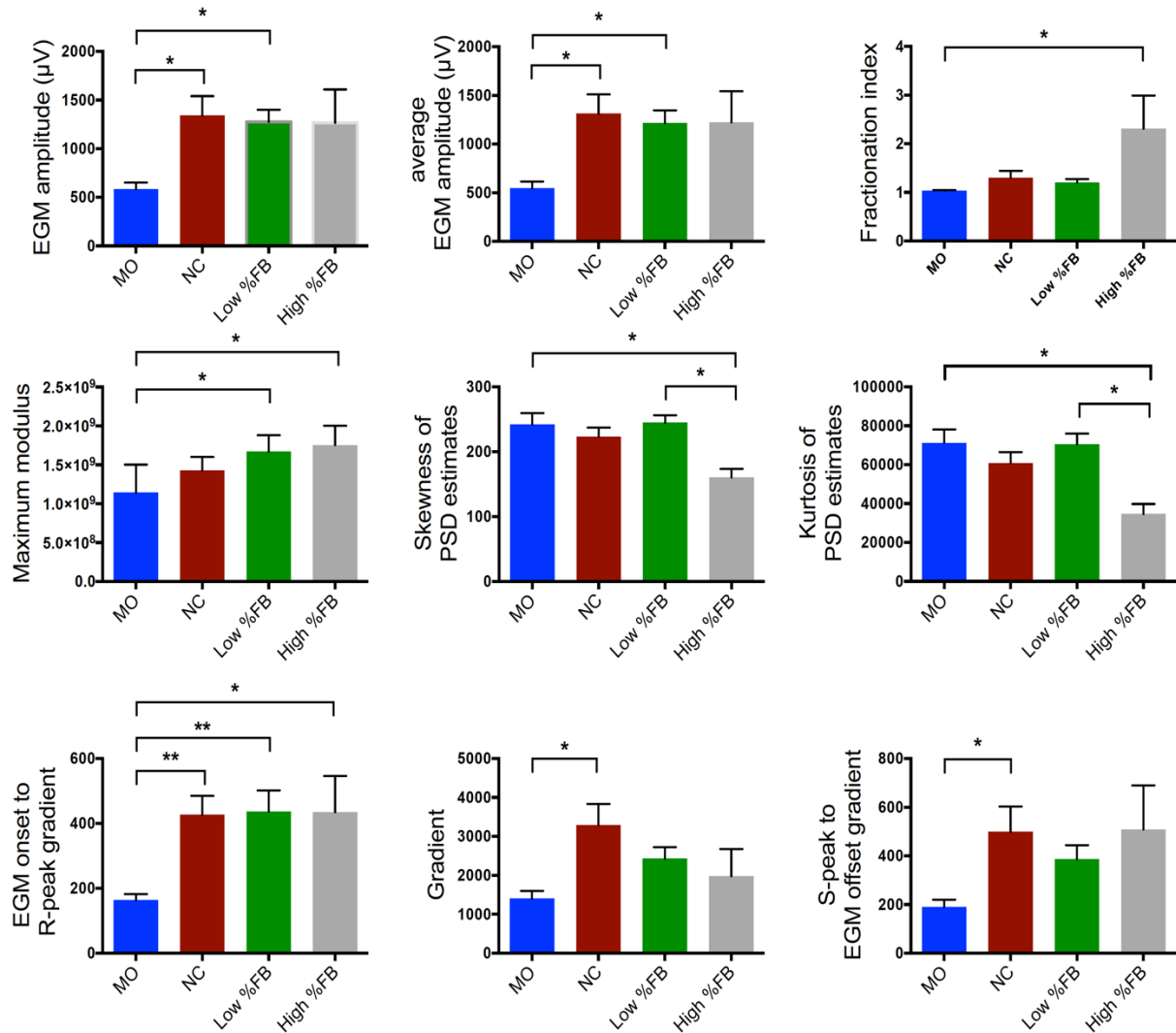


Figure 6.3: EGM properties showing significant differences among cell groups with various amounts of fibroblasts. Kruskal-Wallis test analysis. All bar charts represent mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

6.2.2 EGM morphology changes compared across modulations

Following the analysis of EGM morphology modification per modulation, an overall comparison of feature changes was carried out. Figure 6.4 presents a clustered heatmap, created using Clustergrammer (362), showing the amount of feature alterations in each abnormality based on the average EGM

recorded in myocyte only cultures. The heatmap presents clusters of similar feature modifications on a variety of biological experimental modulations. The range of modifications was between -99 to +683%. It was observed that the majority of features, belonging to EGMs obtained from myocyte only cultures with ion channel blockade or gap junction uncoupling, were dramatically decreased following the modification. However, there were exceptions on this, such as the I_{Kr} , I_{Ks} and K_{ATP} data. Moreover, there is variability among K^+ channel blockers depending on their target. Time-frequency features increased following the opening of K_{ATP} channels, as opposed to any K^+ channel blockade. Time-domain features, such as EGM amplitude, RS/QS-interval and RS gradient were predominantly enhanced with the increased presence of fibroblasts in cell cultures (NC, +20% FB, +40% FB cell cultures) either with or without the concurrent presence of functional modulations. On the contrary, there was greater reduction of time-frequency and frequency-domain features with the presence of functional modifications in MO cell cultures compared to the rest types of cell cultures.

Among other findings, it was interestingly found that gap junctional blockade using CBX on NC monolayers and NRVM/FB co-cultures affected the EGM morphology in a similar way to I_{to} blockade. It was observed in both cases of functional modifications that fractionation, RS interval, R-peak width, QS interval, EGM duration and S-peak width were dramatically increased, while EGM gradient, EGM amplitude and a variety of time-frequency and frequency-domain features, such as maximum modulus, variance of energy, Shannon entropy and the standard deviation of PSD estimates, decreased. Regarding the latter group of time-frequency and frequency-domain features, these were found to be significantly increased in a different cluster mainly consisting of I_{CaL} , I_{Na} , I_{Ks} and K_{ATP} blockade in various cellular experimental models with no clear focus on any of them.

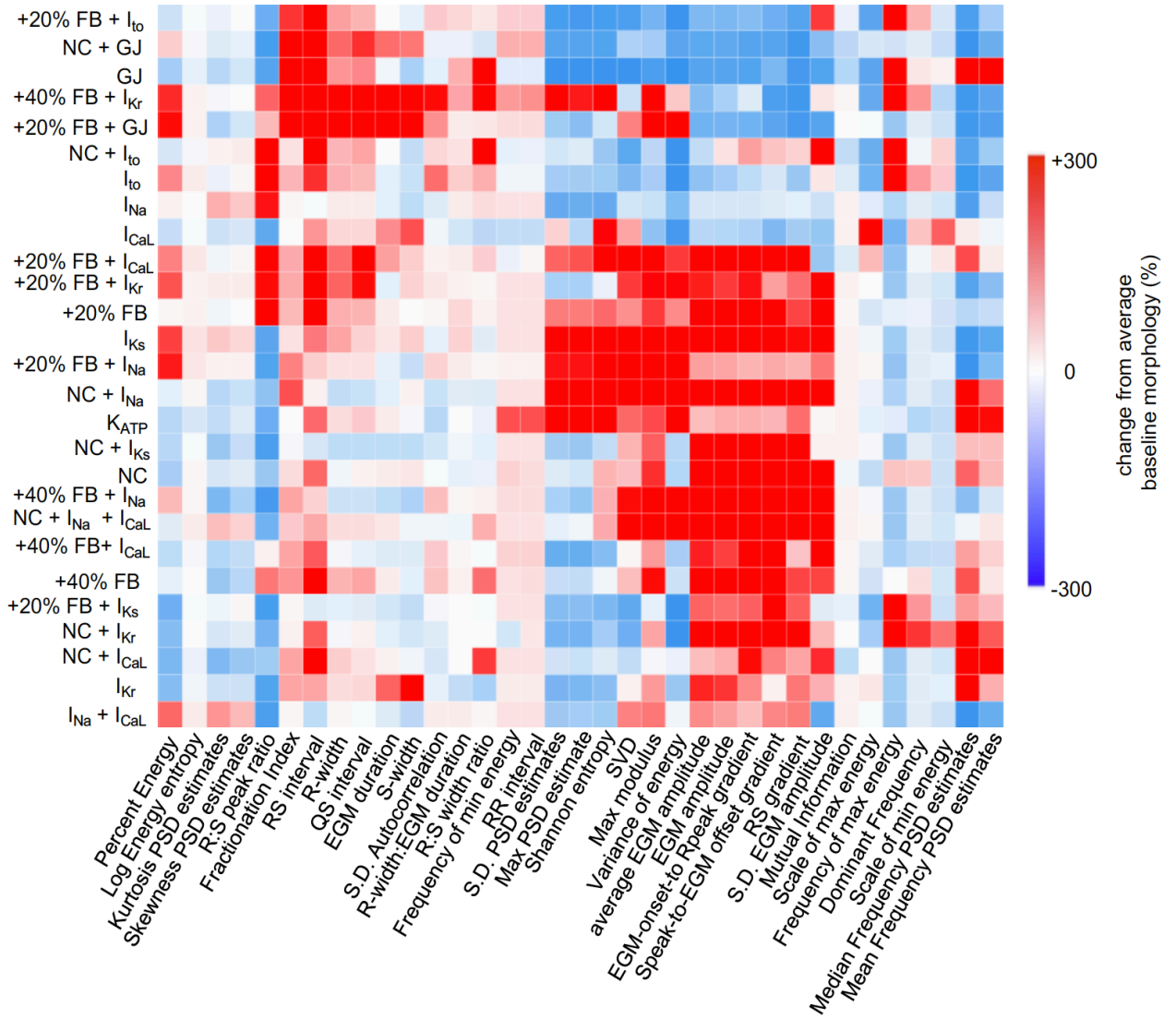


Figure 6.4: Heatmap presenting EGM feature changes per modulation. It was created using the Clustergrammer software. The colours range between red for minimum feature reduction (-99%) to green for maximum feature enhancement (+683%) of the average EGM feature measurement. The average control myocyte only EGM morphology was used as the base of measurements, and for this reason this data group is not presented on the heatmap.

6.2.3 Investigating EGM morphology discrepancies

Comparing the EGM data collected from different cell groups, it was counterintuitively observed that increased amounts of fibroblasts were correlated to increasing voltage of the signal. In order to investigate these inverse results in more depth, NRVMs were co-cultured with the electrophysiologically inactive HeLa cells, in order to investigate the EGM morphology effects of the electrophysiological activity of this cell culture. The total area of the cell culture covered by HeLa cells

was quantified by using the fluorescence excited by HeLa cells which were previously stained using the CellTracker Red CMTPX system. The CellTracker staining protocol is presented in detail in section 2.6.3. This coverage was similar to the one by fibroblasts (NRVM/HeLa: $50.9 \pm 1\%$; MO: $35.7 \pm 2\%$; NC: $63.9 \pm 4.5\%$; +20% FB: $72.7 \pm 3.5\%$; +40% FB: $76.2 \pm 3.4\%$; $p < 0.0001$; mean \pm SEM) (Figure 6.5A). EGM morphology analysis showed that the EGM amplitude was significantly lower in NRVM/HeLa co-cultures compared to any other cell group (NRVM/HeLa: $331.4 \pm 68.5\mu\text{V}$; MO: $546.2 \pm 70\mu\text{V}$; NC: $1312 \pm 199\mu\text{V}$; Low %FB: $1217 \pm 129.7\mu\text{V}$; High %FB: $1225 \pm 317\mu\text{V}$; $p = 0.001$; mean \pm SEM) (Figure 6.5B) and the fractionation was equally increased in the +40% FB cell group (NRVM/HeLa: 1.9 ± 0.2 ; MO: 1.04 ± 0.01 ; NC: 1.3 ± 0.14 ; +20% FB: 1.2 ± 0.07 ; +40% FB: 2.3 ± 0.7 ; $p = 0.002$; mean \pm SEM) (Figure 6.5C).

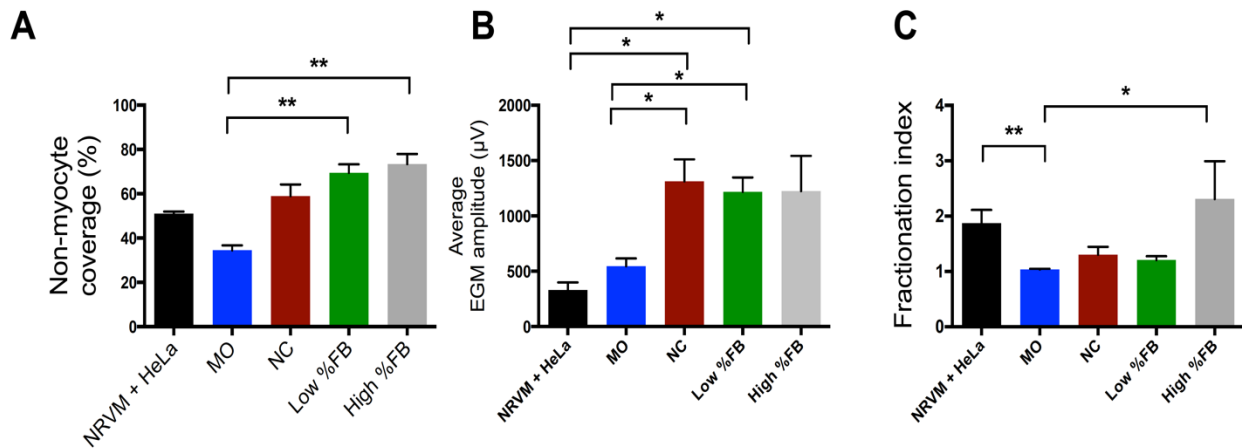


Figure 6.5: Evaluation of the EGM morphology effects with the absence of normal gap junctions. (A) Non-myocyte coverage in different cell groups. The size of area covered by HeLa cells was quantified in NRVM/HeLa co-cultures, while the size of area covered by vimentin, the fibroblast marker, was measured in the rest of cell groups. (B-C) Significant differences in voltage (B) and fractionation (C) were observed among cell groups. Kruskal-Wallis test analysis. All bar charts represent mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

6.2.4 Myofibroblasts influence EGM morphology

For the purpose of this study, a myocyte was defined as a cell with positive stain for α -actinin, fibroblasts and myofibroblasts were positive for vimentin, and α SMA was used as a myofibroblast marker. The MO group was used as control containing mainly NRVMs, the fibroblast (FB) group was the natural cell composition group which was hypothesised to include more fibroblasts than myofibroblasts, and the myofibroblast (MFB) group was co-cultures of NRVMs with low amount (+20%) of cultured fibroblasts, which were supposed to have differentiated into myofibroblasts.

The area coverage of vimentin was significantly lower in the myocyte group compared to the FB group ($p < 0.001$) and the MFB group ($p < 0.05$) (MO: $38 \pm 1.7\%$; FB: $65.6 \pm 4.4\%$; MFB: $61.3 \pm 4\%$; $p < 0.0001$) (Figure 6.6A). The percentage of total area covered by α SMA was also significantly decreased in the MO group compared to the FB ($p < 0.0001$) and MFB ($p < 0.0001$) groups (MO: $11.6 \pm 1.1\%$; FB: $54.3 \pm 2.9\%$; MFB: $82.6 \pm 2.6\%$; $p < 0.0001$; mean \pm SEM) (Figure 6.6B). EGM amplitude was significantly higher in the MFB group ($p < 0.05$) and there was also a trend towards an increase in the FB group (MO: $556.2 \pm 82\mu\text{V}$; FB: $1205 \pm 280\mu\text{V}$; MFB: $1333 \pm 170\mu\text{V}$; $p = 0.03$; mean \pm SEM) (Figure 6.6C). Regarding the rest of features which showed significant change, they were increased in the MFB group and there was a trend towards an enhancement of the same features in FB data (Figure 6.6D-G). This was observed for: RS gradient ($p = 0.18$, $p < 0.05$), EGM onset to R-peak gradient ($p = 0.051$, $p < 0.05$), S-peak to EGM offset gradient ($p = 0.059$, $p < 0.01$) and maximum modulus ($p = 0.059$, $p < 0.05$). The p -values correspond to the difference between the MO and FB groups and between the MO and MFB groups respectively. There was no significant change between FB and MFB groups for any of the characteristics.

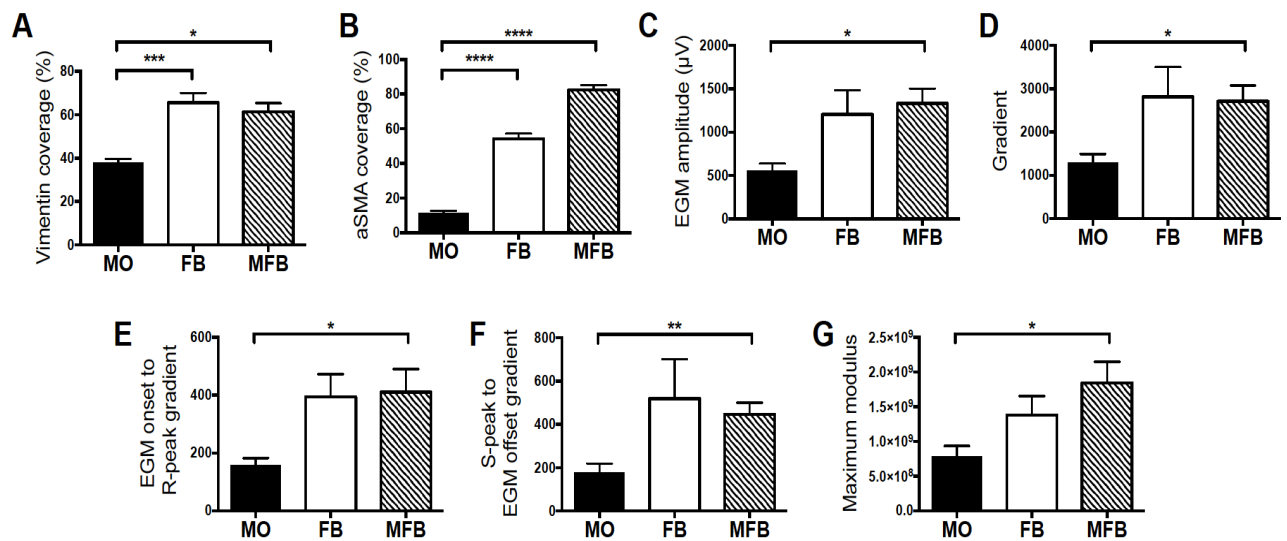


Figure 6.6: EGM morphology modifications showing a significant difference in the myocyte only group compared to either the fibroblast or the myofibroblast group or both. MO, Myocyte only group: $n = 13$ cell cultures; FB, fibroblast group (natural cell composition): $n = 12$ cell cultures; MFB, myofibroblast group (NRVM/20% cultured fibroblasts): $n = 10$ cell cultures. Kruskal-Wallis test analysis. All bar charts represent mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

6.2.5 *Ex vivo* effects of fibrosis on AP and EGM morphology

Fibrosis was quantified on human tissue slices by measuring the amount of slice area covered by collagen, which is a marker of fibrosis. SHG imaging data (Figure 6.7) presented that collagen coverage was between 6-24% of slice area in heart failure and donor samples used in dual modality performance experiments (Figure 6.9A). According to that information, and because there is no convention for describing tissue slice fibrosis, the samples were sorted into three groups based on their collagen coverage: “low” - <10% (n=7 slices), “moderate” – 10–20% (n=5 slices) and “high” - >20% (n=7 slices).

Examples of EGMs obtained from tissue slices with various collagen coverage are presented in Figure 6.8. APD₉₀ was significantly different among groups (low: 473.7±22.3ms; moderate: 376.7±46ms; high: 393.1±22ms; p = 0.035; mean ± SEM) (Figure 6.9B), but no differences were observed regarding FPD (low: 432.6±19ms; moderate: 365.1±41ms; high: 378±20ms; p = 0.36; mean ± SEM) (Figure 6.9C). Linear regression analysis showed a correlation between APD₉₀ and FPD ($r^2 = 0.69$, $Sy.x=59.3$, slope = 0.71 ± 0.07 , $p<0.0001$, $n=49$) and this relationship was maintained even within each group with no differences between groups (p=0.99) (Figure 6.9D). Conduction velocity was the same among groups (low: 11.5 ± 1.9 cm/s; moderate: 9.5 ± 2.2 cm/s; high: 13.5 ± 1.5 cm/s; p=0.43; mean ± SEM) (Figure 6.9E). No significant differences were observed in EGM duration (low: 41.5 ± 2.7 ms; moderate: 34.4 ± 2 ms; high: 37.6 ± 3.7 ms; p = 0.13; mean ± SEM) (Figure 6.8F). Moreover, no linear relationship was found between conduction velocity and EGM duration ($r^2=0.1$, $Sy.x=9.2$, slope = -0.77 ± 0.66 , p = 0.29, n=13) (Figure 6.9G).

EGM features were computed for data collected from slices with low (n = 9 slices), moderate (n = 7 slices) and high (n = 7 slices) collagen coverage. In total, 487 EGMs were analysed. Following the EGM morphology analysis, it was observed that 27 features significantly changed between groups (Figure 6.10). These were features belonging to all types of EGM morphology analysis. It is interesting that slices with moderate collagen coverage differed significantly from the rest of groups for a number of features, including EGM amplitude (Low: $2809 \pm 143\mu V$; Moderate: $1985 \pm 179\mu V$; High: $2669 \pm 182\mu V$; p<0.001 (Low vs. Moderate) / p<0.05 (Moderate vs. High)), RR interval (Low: 1026 ± 38 ms; Moderate: 1595 ± 79 ms; High: 1024 ± 14 ms; p<0.0001 (Low vs. Moderate) / p<0.0001 (Moderate vs. High)), dominant frequency (Low: 0.96 ± 0.02 Hz; Moderate: 0.81 ± 0.03 Hz; High: 0.97 ± 0.01 Hz; p<0.0001 (Low vs. Moderate) / p<0.0001 (Moderate vs. High); mean ± SEM) and all time-frequency features. There was no change between slices with low and high collagen coverage (Figure 6.10).

Surprisingly, fractionation did not change among groups (Low: 4.3 ± 0.3 , Moderate: 5.2 ± 0.4 , High: 4.4 ± 0.3 ; $p=0.13$; mean \pm SEM).

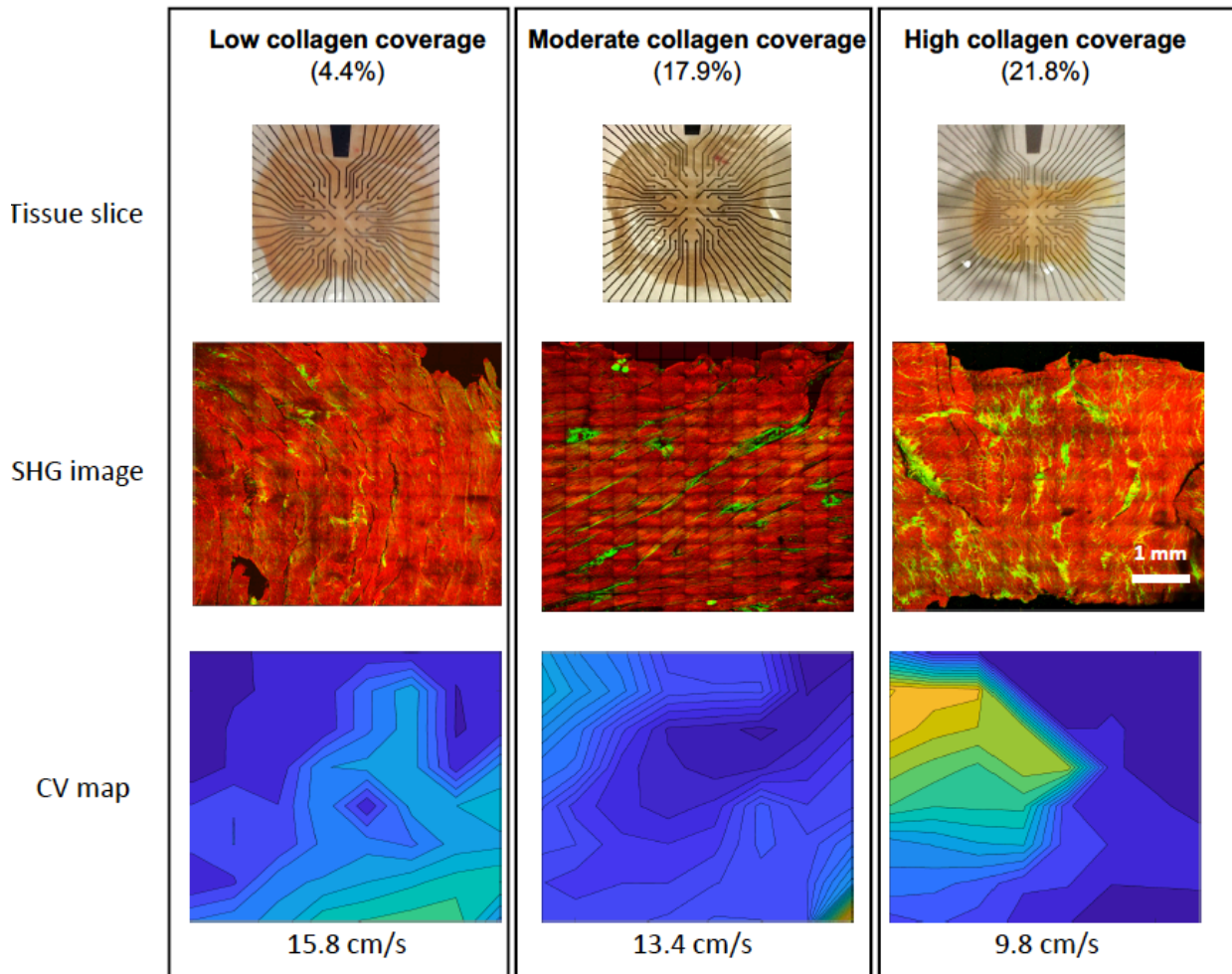


Figure 6.7: Human left ventricle tissue slices visualised in three different ways. (*Top*) Photos showing the slices placed on top of electrodes in a MEA plate. The reference electrode can be seen on the top side. (*Middle*) The same slices, and with the same orientation as on the top photos, scanned by SHG microscopy for quantification of collagen burden. Representative examples of slices with different sizes of collagen covered areas are presented here. Green represents collagen and red corresponds to the autofluorescence signal emitted from the rest of tissue areas. (*Bottom*) Isochronal activation maps for the wavefront propagation in each slice. Regarding the middle slice (moderate coverage), the conduction slowing observed in the top right area of the CV map corresponds to the fibrotic area captured during SHG imaging.

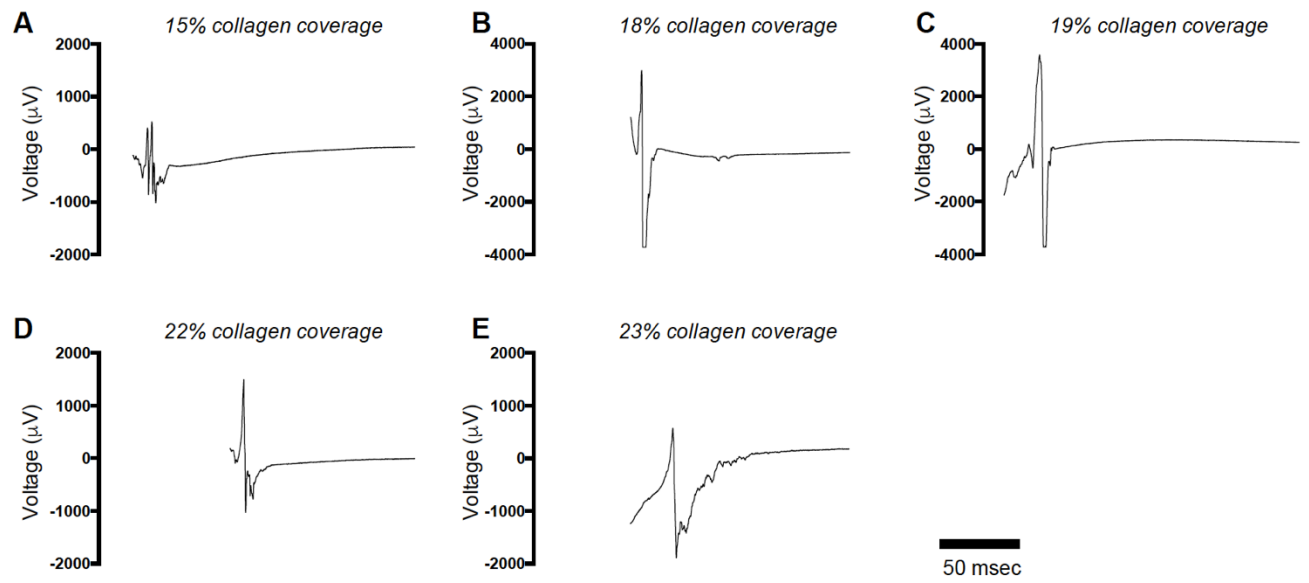


Figure 6.8: Representative examples of EGMs obtained from tissue slices with different collagen coverage. Here, 200msec of each recording are shown emphasising the variability of the QRS complex morphology.

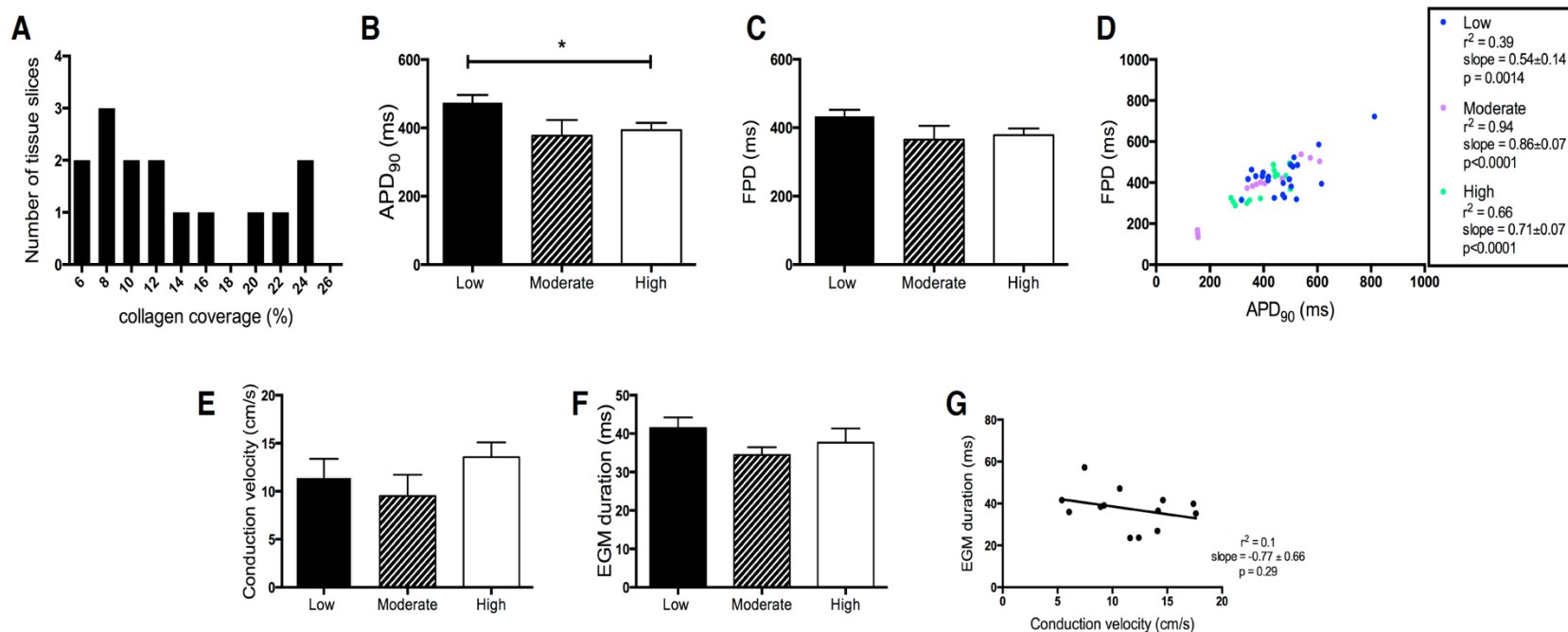


Figure 6.9: (A) Frequency distribution graph of the collagen coverage on human ventricular tissue slices used on dual modality experiments. These tissue slices were divided into three groups: “Low” – <10% collagen coverage (n=7 slices), “Moderate” – 10.1-20% collagen coverage (n=5 slices), “High” - >20% collagen coverage (n=7 slices). (B) Quantification of APD₉₀ in each experimental group. One-way ANOVA showed that APD₉₀ is significantly different among groups (p=0.041), even though Tukey’s post-hoc test did not show significant differences. (C) FPD in each group of slices. (D) Linear regression analysis showed significant correlation between APD₉₀ and FPD (r² = 0.69, Sy.x=59.3, slope = 0.71±0.07, p<0.0001). (E-F) Conduction velocity (E) and EGM duration (F) did not show any significant differences among collagen coverage groups. (G) No linear relationship between conduction velocity and EGM duration (r²=0.1, Sy.x=9.2, slope= -0.77±0.66, p=0.29). All bar charts represent mean ± SEM; *p<0.05. APD₉₀ and FPD plots were made using manually annotated data.

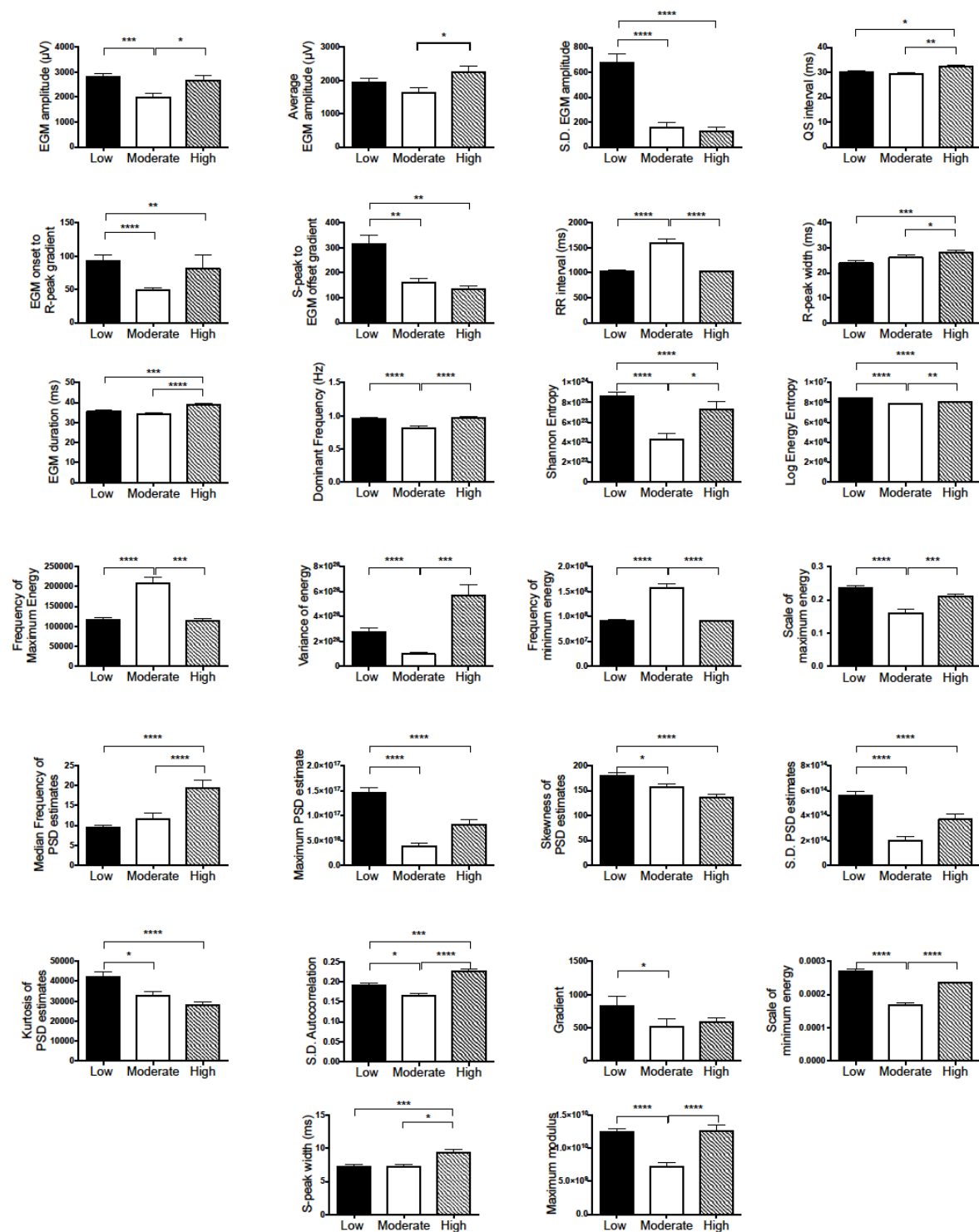


Figure 6.10: EGM properties modified due to differences in the size of collagen coverage, and subsequently the extent of fibrosis, in human left ventricle tissue slices. Low, <10% collagen coverage: n = 9 slices (165 EGMs in total); Moderate, 10-20% collagen coverage: n = 7 slices (102 EGMs in total); High, >20% collagen coverage: n = 8 slices (130 EGMs in total). Kruskal-Wallis test analysis. All bar charts represent mean±SEM; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

6.2.6 *In vivo* effects of scarred areas on EGM morphology

Unipolar EGMs, collected in patients with AF during electrical stimulation of regions of the LA endocardium at 3Hz (Figure 6.11), were analysed for their morphology. These data were obtained from 8 patients and consisted of 348 EGMs derived from control locations and 126 EGMs collected from scarred regions (≥ 2 S.D. blood pool mean). Scarred locations were defined as locations of the LA where the atrial intensity at each voxel, measured by LGE-MRI, was expressed as 2 S.D. above the blood pool mean (>2 S.D.). Control locations were defined as the locations where the raw atrial intensity was ≤ 0 S.D. below the blood pool mean. Data analysis showed no difference in voltage between control and scar locations (control: $2.17 \pm 0.1\text{mV}$; scar: $1.96 \pm 0.1\text{mV}$; $p=0.28$; mean \pm SEM) (Figure 6.12). However, there were 15 features that were significantly altered. These included RS interval (control: $23 \pm 1\text{ms}$; scar: $36.3 \pm 2.3\text{ms}$; $p<0.0001$; mean \pm SEM), EGM duration (control: $63.3 \pm 1.3\text{ms}$; scar: $85 \pm 3.4\text{ms}$; $p<0.0001$), fractionation (control: 2 ± 0.1 ; scar: 2.3 ± 0.2 ; $p=0.008$; mean \pm SEM) and logarithmic energy entropy (control: 2273 ± 145 ; scar: 2842 ± 176 ; $p = 0.0003$; mean \pm SEM) (Figure 6.13). Among the modified time-frequency characteristics were maximum modulus (control: 3.42 ± 0.1 ; scar: 3 ± 0.2 ; $p = 0.0008$; mean \pm SEM) and variance of energy (control: $225 \times 10^3 \pm 12 \times 10^3$; scar: $123 \times 10^3 \pm 14.2 \times 10^3$; $p<0.0001$; mean \pm SEM) (Figure 6.13).

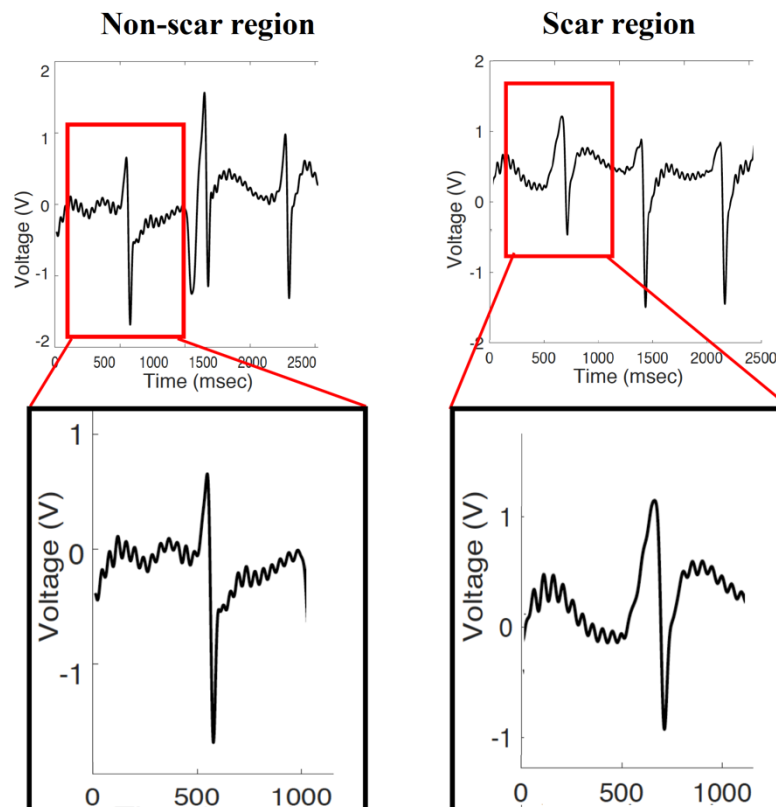


Figure 6.11: Representative examples of 2.5msec clinical recordings obtained at 3Hz from the non-scar and scar regions of the LA posterior wall. (*Top*) The original recording includes the noise. EGM amplitude alternans recorded in scar regions is also presented here. (*Bottom*) The first msec of each recording is zoomed in.

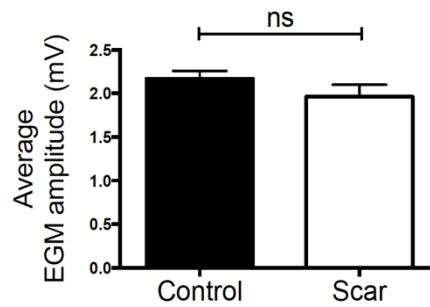


Figure 6.12: EGM amplitude was no different between control (≤ 0 S.D. blood pool mean) and scarred regions (≥ 2 S.D. blood pool mean). Unipolar paced EGM recordings were collected from 8 AF patients (control: $n = 348$ EGMs; scar: $n = 126$ EGMs in total). Mann-Whitney t-test. Bar charts represent mean \pm SEM; ns = non-significant.

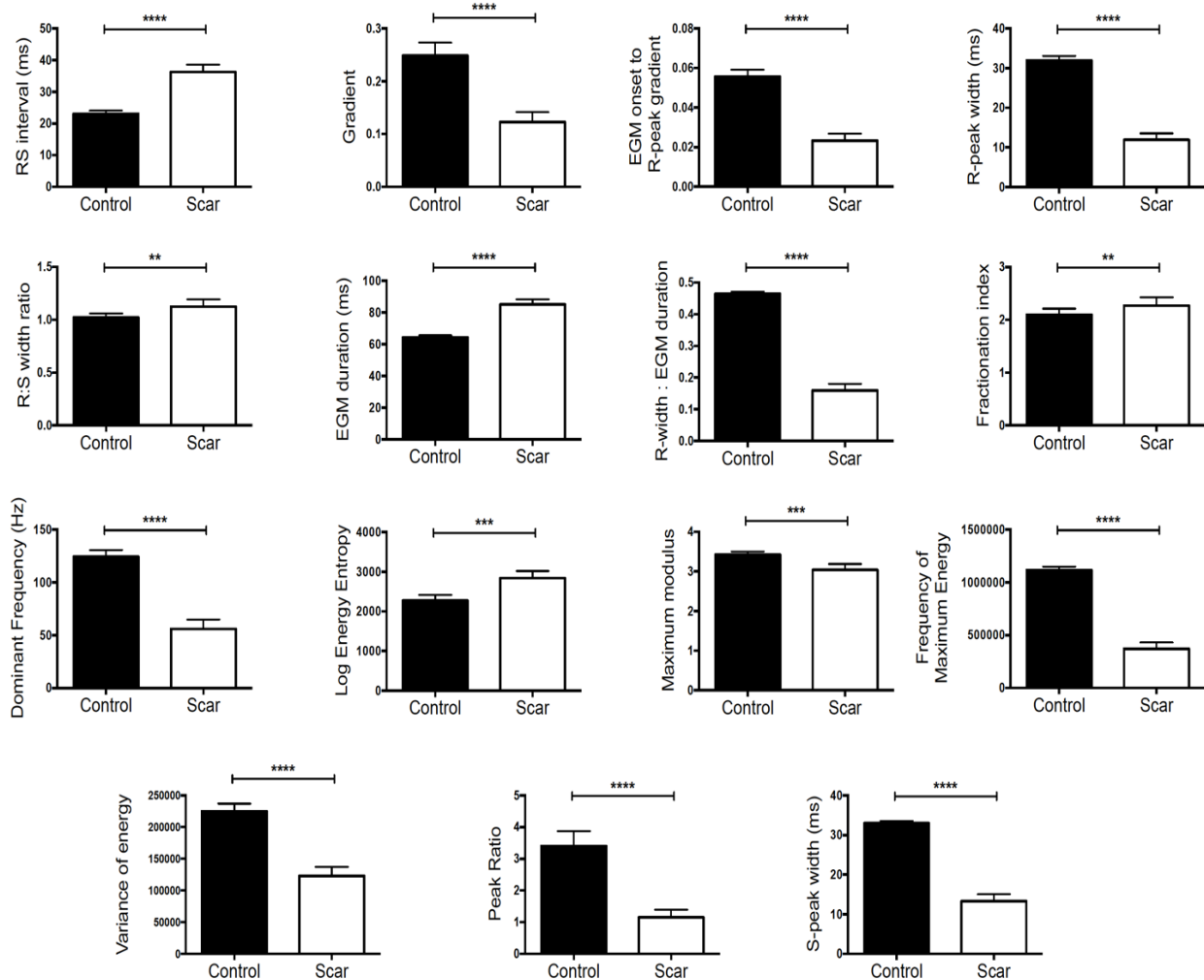


Figure 6.13: EGM morphology modifications due to the presence of scar, as measured on 3Hz data collected from control (≤ 0 S.D. blood pool mean) and scarred regions (≥ 2 S.D. blood pool mean) of the LA endocardium (control: $n = 348$ EGMs; scar: $n = 126$ EGMs in total). Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

6.3 Discussion

This chapter focused on the characterisation of EGM morphology as affected by *in vitro* modifications of structural parameters and any *ex vivo* or *in vivo* structural abnormalities. A general observation is that depending on the experimental model both the cellular component of scar and fibrosis can affect EGM morphology and lead to conduction slowing. It was also observed, mainly on *in vitro* and *in vivo* results, that feature modifications can be dependent on the amount of cellular scar or the degree of fibrosis respectively. An additional aim of this study was the investigation of the effect of myofibroblasts on cardiac electrophysiology by identifying alterations in EGMs. The data presented in this chapter were obtained from experimental models with increasing complexity. However, it is hard deriving conclusions for a model based on another, due to the different metrics for non-conductive tissue applied in each case and the specific electro-architectural characteristics, e.g. 2D structure in cell cultures and tissue slices vs 3D structure in the LA.

6.3.1 Fibroblast effects of EGM morphology

It was sought to be investigated how fibroblasts and conduction slowing affect EGM morphology in a 2D cellular model of scar. For this purpose, NC and NRVM/fibroblast co-cultures with +20% or +40% FB were compared to MO preparations, in order to assess the effects of fibroblast burden. All types of cell culture had significantly higher fibroblast coverage compared to MO monolayers, but there were no significant differences among NC, +20% FB and +40% FB cell cultures. However, it was shown that the degree of EGM fractionation did not agree with the fibroblast burden, as there was a significant increase of fractionation solely in the +40% FB group and it was maintained low for the rest of cell groups. Computational models of fibrosis reached similar conclusions revealing discontinuous conduction and EGM fractionation with high amounts of fibrosis implemented in the model (363). Similar to Jacquemet *et al.* (364), whose study presents the introduction of microfibrils via collagenous septa in a computational model of cardiac monolayer leading to increased fractionation, the results presented in this chapter suggest that fractionation can be also enhanced by significantly increasing the fibroblast burden. This is explained by the fact that high amounts of fibroblasts intermingled amongst myocytes create a heterogeneous substrate which is responsible for conduction slowing, discontinuous propagation and conduction delays across electrodes (4). The absence of increased fractionation in the NC and +20% FB cellular models, when compared to the MO monolayers, suggests that structural abnormalities in these cell cultures were insufficient to lead to discontinuous conduction. Therefore, we could propose the presence of a dose-dependent or a threshold effect of the fibroblast burden.

FPD and APD₉₀, which are equivalent characteristics (20,40), were prolonged in NC and +20% FB cell cultures compared to MO monolayers proposing that the coupling between myocytes and fibroblasts leads to slower repolarisation. The reason for this might be the heterogeneous coupling between myocytes and fibroblasts through gap junctions that can be formed not only by myocytes, but also *in vitro* by isolated fibroblasts (352,360). Moreover, due to the separation of myocytes by fibroblasts, there is reduced myocyte-to-myocyte coupling leading to enhanced APD in myocytes. The disruption of normal intercellular coupling is a substrate for discontinuous propagation promoting functional conduction blocks and even abnormal wave propagation (31).

Contrary to what was expected, increased fibroblast burden did not demonstrate amplitude reduction when compared to the MO group. This disagrees to what has been suggested by *in vivo* studies (6,106,365), which present that there is an inverse relationship between endocardial voltage of pacing or AF data and fibrosis, as determined by LGE imaging (365). However, the *in vitro* results presented here agree with the *ex vivo* findings next presented. In fact, the experimental model used in this thesis for extracellular recordings does not mimic the *in vivo* fibrosis, which is characterised by increased fibroblasts content and elevated amount of extracellular matrix (352). A reason for the higher EGM amplitude in NC and NRVM/FB co-cultures could be the number of myocytes on top of an electrode. The area under the EGM represents the total ion exchange (242), and therefore a large number of myocytes and higher cellular density may lead to higher ion exchange and increased EGM amplitude. This suggestion can be supported by clinical studies where it was presented that endocardial voltages were lower in scar LA regions. High collagen deposition and low cardiac myocyte population were found in this areas, as opposed to patchy fibrotic or healthy areas, where collagen deposition was lower, the myocyte population was elevated and the EGM amplitude was higher (7).

In order to investigate whether the higher EGM amplitude was a result of possibly enhanced conductivity due to changes in the electrophysiology of cultured fibroblasts, e.g. enhanced intracellular communication, cardiac myocytes were co-cultured with HeLa cells. HeLa cells do not express Cx43 and for this reason this is a commonly used cell line for transfections with the Cx43 gene (366). The HeLa cell coverage on cell cultures was similar to the fibroblast one in the NC cell group. The result was that the voltage of the signal recorded from NRVM/HeLa co-cultures was lower than the MO preparations and in fact it was the lowest voltage recorded *in vitro* under baseline conditions for the purposes of this thesis. EGM fractionation was also comparable to the one measured in EGMs obtained from +40% FB cultures. That confirmed that there are less gap junctions formed in the NRVM/HeLa co-cultures compared to the rest of cell preparations giving rise to EGMs with low amplitude. These experiments also proved that the absence of gap junctions can result to discontinuous action potential propagation, and therefore fractionated and low voltage signal. However, it appears that further

investigations need to be carried out in order to explain the increased EGM amplitude recorded in NRVM/FB co-cultures.

Moreover, K^+ -related pharmacological agents with opposite activity were applied on the NRVM/FB co-cultures. Following clustering, it was observed that time-frequency features changed towards the opposite direction after the enhancement of K_{ATP} current using pinacidil compared to the effects of K^+ channel blockade by 4-AP (I_{to} current blockade) or E-4031 (I_{Kr} current blockade). These findings agree with feature modifications presented in Chapter 4 of this thesis about individual ion channel modulations. In addition, specific time-frequency and frequency-domain features were predominantly increased after the blockade of I_{CaL} , I_{Na} , I_{Ks} and K_{ATP} . There is no previous experience on such EGM modifications, in order to explain any mechanistic insights, but these findings suggest that a common ion exchange mechanism may be influenced in myocyte/fibroblast co-cultures. However, each ion current plays a different role, and thus the blockade of each ion channel variably affects the degree of EGM modifications.

6.3.2 The role of myofibroblasts

Myofibroblasts are not considered to be part of healthy cardiac tissue and they usually appear following cardiac injury (358,367). They migrate to and are highly attracted by chemokines released at the site of injury. It has been found that myofibroblasts also produce cytokines, as a way to maintain the inflammatory response to injuries by themselves (358). Myofibroblasts have contractile ability, because a system of microfilaments connected to extracellular fibronectin has been found enabling cells to generate force to the extracellular matrix. Levels of α SMA, a myofibroblast marker (368), and therefore myofibroblast expression, are enhanced between 5 to 14 days following cardiac injury, and this is followed by a decrease in α SMA expression levels (358,369). This agrees with our findings about increased α SMA signal, and consequently enhanced amount of myofibroblasts, on NRVM/cultured fibroblast co-cultures. These fibroblasts had been cultured in low density, progressively leading to high cell density due to proliferation, for 7 days prior to be mixed with NRVMs. Recordings were obtained 2-4 days following seeding on MEA plates, as presented in section 2.3. These are experimental conditions that according to previous *in vitro* studies stimulate fibroblasts to differentiate into myofibroblasts (358,370), supporting the belief that the NRVM/fibroblast co-cultures in this study mainly consisted of myofibroblasts.

Fibroblasts usually express Cx43 and Cx45, which participate in gap junction formation (358). However, these connexins seem to be upregulated in myofibroblasts suggesting enhanced intercellular coupling (34,358). It has been supported that when myofibroblasts, derived from *in vitro* fibroblast

differentiation, are co-cultured with NRVMs, hetero-cellular gap junctions are formed between myofibroblasts and myocytes and homo-cellular gap junctions between cells of the same type (34). It is likely though that myofibroblasts create a barrier to normal conduction activity, because it is suggested that they are not as excitable as myocytes and they are intercalated among myocytes increasing the distance between neighbouring myocytes (358). This could explain the fact that a number of EGM features were influenced by the presence of myofibroblasts.

A limited number of features were modified due to the presence of myofibroblasts. These were EGM amplitude, EGM onset to R-peak gradient, S-peak to EGM offset gradient, RS gradient and the maximum modulus, which was the only affected time-frequency analysis feature. Apart from maximum modulus, the rest of EGM features correlate to the depolarisation phase of cardiac myocytes. These changes could be explained by the depolarising effect of myofibroblasts on myocyte electrophysiology, as demonstrated by Miragoli *et al.* (34,371). According to that study, myofibroblasts have a relatively depolarised resting phase of -50mV or less, which affects the more negative myocyte resting membrane potential of $\approx -70\text{mV}$ and it leads to its increase to -65 to -56mV (371). This effect was found to be more prominent when the amount of myofibroblasts exceeded 15% in the co-culture with myocytes, which was also the case in our experimental model. Thus, the depolarisation could be more spontaneous (371). Since EGM duration corresponds to the upstroke duration of depolarisation in an action potential (40), we could suspect that this effect of myofibroblasts leads to reduced EGM duration and higher EGM amplitude, and thus to increased gradients.

6.3.3 Characterisation of EGM feature modifications in adult human ventricular slices

The method used in this study for the investigation of fibrotic regions in adult human ventricular slices was SHG microscopy (372,373). This imaging method has been used multiple times for the visualisation of collagen distribution within a human myocardial slice (326,329,372). However, previous studies did not attempt to quantify fibrosis based on the amount of collagen, which is captured by SHG imaging. Thus, it is suggested that the size of the tissue area covered by collagen could be used both as a measure for quantifying fibrosis and as a marker of the potential of a tissue slice to be used for EGM recordings. The latter is a result of the observation that tissue slices, derived from heart failure myocardia, with a collagen coverage of over 24% could not be stimulated and no EGMs could be obtained from them. This is a confirmation that extended fibrotic areas obstruct the normal action potential propagation (32).

There is poor knowledge on EGM morphology recorded from human myocardial tissue slices. This has been mainly focused on the investigation of FPD modifications (20,40,249), due to the interest on assessing new anti-arrhythmic and non-cardiac drugs for effects on APD and any possible induction of QT prolongation (20,374). FPD recorded in slices in this present study was similar to the APD recorded simultaneously from them under the same experimental conditions using a novel apparatus for concurrent MEA and optical mapping recordings developed by Chowdhury, Tzortzis, *et al.* (40). These results were similar to data obtained from multicellular preparations, such as ventricular slices (20) or wedges (349) and papillary muscles (375) under comparable conditions. The FPD values were longer than those reported for human ventricular tissue collected from healthy donor hearts (349,376) and the APD measurements agree with previous findings that the APD is prolonged in heart failure (377). Slices benefit from inclusion of the several cell types that are found in the native tissue, such as myocytes, fibroblasts and endothelial cells (20).

Voltage and fractionation counter-intuitively changed among the groups of data, which were based on collagen coverage. The average EGM amplitude was enhanced due to the increased collagen coverage, while fractionation was not affected. Both observations do not agree with our knowledge about the effects of fibrosis, as it is known that voltage decreases and fractionation increases in scarred regions (6,106,127). However, a number of additional non-conventional features were influenced by the amount of collagen coverage, and therefore fibrosis, in tissue slices. These included RS gradient, S-peak width, Shannon entropy, S.D. EGM amplitude, variance of signal energy, EGM duration and autocorrelation. Regarding Shannon entropy, EGMs obtained from slices with >20% collagen coverage had higher Shannon entropy than the EGMs obtained from slices with 10-20% collagen coverage. This agrees with previous findings presenting Shannon entropy as a method for quantifying the complexity of EGMs obtained from scar regions, even though these referred to CFAE signals (127). Moreover, on the same way that increasing cellular scar in cell cultures leads to maximum modulus enhancement and a reduction of the skewness and kurtosis of PSD estimates, as earlier presented in this chapter, the same effects were observed on slices belonging to the moderate and high collagen coverage groups. This could suggest that abnormalities in conduction propagation comparably influence these time-frequency and frequency-domain features in multicellular preparations.

6.3.4 The *in vivo* relationship between EGM morphology and atrial fibrosis

This chapter also presented the analysis of paced EGM signal obtained from the LA of patients with persistent AF. Even though these data are not representative of arrhythmic events, they provide information about the ways that scar can manifest EGM morphology compared to non-scarred regions.

LA scar was quantified automatically using a method that collocated voltage mapping results with LGE imaging data used for the construction of a LA surface scar map (174). This map expressed the gadolinium signal intensity as multiples of SD above blood pool mean, as described in section 2.4. Scar has been originally defined as a region with >3 SD above blood pool mean (174), but for the purposes of this thesis and due to lack of adequate EGM data, regions with >2 SD blood pool mean were included as scarred regions. LA areas with <0 SD blood pool mean were defined as non-scar regions.

It is supported that the underlying atrial substrate, represented by low voltage areas, can predict the outcomes of catheter ablation of AF (98). Low-voltage areas may represent regions of slow conduction as a manifestation of underlying fibrosis or wave collisions representing more functional electrophysiologic change and these regions have been suggested to be crucial for the AF maintenance. Current catheter ablation strategies are mainly focused on targeting potential AF drivers involving targeting low voltage regions. Low voltage areas are more represented in persistent AF than in paroxysmal AF, as shown by studies utilising voltage in paced-rhythm (99) and AF (100) recordings. However, the results presented in this chapter do not agree with the literature, because there was no clear reduction of EGM amplitude for statistically significant conclusions. These findings are also opposite to the fact that the analysed data were obtained from the LA posterior wall, which is known from histological studies to be more prominent to low-voltage areas in patients with AF (101).

Another characteristic feature of the underlying substrate is the presence of fractionated EGMs, which are believed to coincide with areas of slow conduction or pivot points where the activation wavefront turns around at the end of a functional block (104). The study presented here showed that fractionation was indeed increased in scarred regions. This also agrees with previous studies by Nademanee *et al.*, where it was reported that fractionated signal obtained from persistent AF patients was believed to collocate with scar areas responsible for AF occurrence and regions targeted during catheter ablations (6,106).

The rest of features that were modified due to the presence of scar included EGM duration, RS interval, the width of R- and S-peaks, logarithmic energy entropy, maximum modulus, the variance of energy and dominant frequency. EGM duration and RS interval were both prolonged in scarred areas, which could be justified by the slower conduction. In general, these findings supported our initial hypothesis that scar is responsible for modulating EGM features and generating a morphology which is distinctive from the one recorded in healthy/non-scarred areas.

6.4 Limitations

Cultured fibroblasts were used during the preparation of the *in vitro* experimental models and they were co-cultured with myocytes. The purpose of these fibroblasts was to be used as a way to mimic the pathological conditions. However, fibroblasts isolated from rat diseased hearts should be ideally used, in order to reproduce the myocardial conditions.

An additional limitation of the above study is the lack of a reliable method for quantifying fibrosis in tissue slices. The method presented here uses the total size of the area covered by collagen I molecules, which are part of ECM. Even though SHG microscopy has been the standard method for non-invasive imaging of fibrotic areas in different types of tissue (329,372,378,379), this method has not been standardised and verified for quantification purposes, but only for qualitative observations. Consequently, apart from the types of fibrosis based on histology (20,32,380), there is currently no categorisation system based on quantification measurements. Thus, apart from the type of categorisation presented in this chapter, tissue slices could be classified in a variety of other possible ways, including the use of a collagen coverage percentage as a cut-off or dividing samples into two groups, based again on collagen coverage.

The *ex vivo* slice data presented in this chapter generally present that the degree of feature modifications is independent of the amount of fibrosis. A variety of features are significantly eliminated in the moderate collagen coverage group compared to the other ones and there may be no difference between the low and high coverage groups. Even though such trends seem impossible to be explained with electrophysiology terms, this may be a result of the categorisation method followed in this chapter.

6.5 Conclusion

In this chapter, it was shown how EGM morphology is modified due to the effect of a modified structure of the preparation. These changes involved the cellular component of scar in 2D cell preparations, variable amounts of fibrosis in tissue slices derived from heart failure myocardium and the presence of scarred regions in the posterior LA endocardium. It was found that such *in vitro* and *ex vivo* structural abnormalities modify not only the conventional features of voltage and fractionation, but also other morphological and complex EGM features. The amount of cellular scar is not always correlated to the degree of feature modifications. It can be also suggested that EGMs dramatically change due to *in vivo* scar presence, but voltage is not one of the features that change, as is conventionally considered. These findings will be useful when developing predictive models based on each experimental model. The *ex*

vivo and *in vivo* models especially have been investigated very little or not at all for the effects of functional abnormalities in this thesis, and there is a focus on structural abnormalities instead. Thus, the features used for predicting unlabelled EGMs, which are obtained under the conditions presented in this chapter, will be meaningful and the way that predictions occur will be clearer.

Chapter 7

Predicting functional and structural abnormalities
from cellular EGM morphology

7.1 Introduction

Primary cultures of NRVMs and FBs are widely used *in vitro* in cardiac research (381). Their advantages include the ability to study pathophysiological processes in an experimental setting that lacks the influence of hemodynamic factors existing *in vivo* and the feasibility to artificially control other concomitant factors using pharmacologic methods (381). In addition, the co-culture of NRVM and FBs provides the ability to robustly investigate a variety of parameters that affect cellular electrophysiology. This is allowed not only by the feasibility of growing both cell types, but also by the presence of gap junctional coupling between both cell types (279). Fibroblasts act as current sinks and impose an electrical load when electrically coupled to myocytes. Moreover, the resting membrane potential of FBs has been shown to be more positive compared to myocytes (382) and may become more hyperpolarized with activation (383). It has been suggested by experimental studies that the increased myocyte/FB coupling may lead to APD changes, electrotonic depression of myocytes, arrhythmogenic excitability gradients, altered conduction and unidirectional block (31,356,384–386). These are all factors that may influence the unipolar extracellular EGM morphology (242) and it worth investigating how each one of them affects specific EGM features. However, the aim of the work presented in this thesis is bi-directional. Thus, apart from the elucidation of functional and structural factors affecting the EGM morphology, it is aimed to be investigated how the same factors can be predicted using EGM features.

Functional abnormalities, such as gap junction uncoupling and ion channel blockade, are more difficult to detect, as opposed to the investigation of the cardiac tissue structure through LGE-MRI or histological techniques. For this reason, there are examples of drugs in late-stage development or with market approval, such as clobutinol, sibutramine, tegaserod, which are then withdrawn due to previously undetected drug-induced cardiotoxicity (387,388). Better safety has been recently facilitated through the adoption of US Food and Drug Administration guidelines that recommend screening new drugs with the hERG inhibition assay (387). Therefore, it becomes clear that there is a necessity for more accurate and faster pre-clinical detection methods leading to the emergence of a variety of screening platforms (387,389). The focus of this thesis is to predict the variety of factors that are known to cause disruptions in conduction velocity and action potential propagation using the EGM morphology. For this reason, there has been an increasing interest over the last decade for the development of predictive models using supervised machine learning methods.

Machine learning includes a variety of statistical methods which aim to build predictive models given a dataset. These techniques are an approach to optimize the utility of datasets generated from screening pharmacological agents and they enable handling multidimensional datasets in an automated fashion

(390). All these methods have proved their performances in many different scenarios of classification, when applied to biological data (389). In this chapter, it is hypothesized that multiclassification algorithms can be used to predict functional and structural abnormalities that take place *in vitro* using the unipolar extracellular EGM morphology. In order to address this, the aims were:

1. to create an EGM morphology database with data obtained
 - a. from a variety of *in vitro* experiments where cardiac disease-related functional modifications were caused by target-specific agents and
 - b. from a number of *in vitro* conditions mimicking structural abnormalities that can be proarrhythmic or related to heart failure and responsible for abnormal action potential propagation,
2. the development of a process for automated feature selection, classification training and validation of the final prediction model through unlabelled data,
3. to investigate a variety of classifiers, in order to find the most effective ones for analysing EGM morphology data related to targeted pharmacological agent effects,
4. to develop a model that has the ability to predict the heart-related drug-specific effects of various pharmacological agents and other structural *in vitro* conditions can be used to decode functional and structural abnormalities which are typically related to a number of cardiac diseases.

EGMs, which were obtained under the experimental conditions presented in Chapters 4-6, were described by values for each one of the time-domain, frequency-domain and time-frequency features that was introduced in Chapter 4. The process that I followed to elucidate the aims of this chapter included the methods presented in section 2.9. Firstly, training and test datasets were created consisting of EGMs obtained from all *in vitro* conditions presented in Chapters 4-6. The SFS method was applied on the training dataset for selecting a feature subset used during classification training. The model, which was extracted in the end of training, was applied on the test dataset of unlabelled EGMs for predicting the functional and structural abnormalities that occurred when these were obtained.

7.2 Methods

The experimental data used in this Chapter are the ones presented in Chapter 4 and derived under a variety of functional and structural modulations in cell monolayers. EGM morphology analysis was carried out using the methods presented in Chapter 3 and feature arrays were created in order to describe the morphology of individual EGMs. Based on these arrays, a dataset was created. EGMs were randomly selected using Matlab to create the training and test datasets. The training dataset consisted

of 59.7% of the initial pool of EGM data and the test dataset consisted of 40.3% of the same initial dataset. No EGMs were included in both datasets, because each EGM was selected only once.

Feature selection was carried out using the SFS method presented in section 2.9.2, in order to exclude the less informative feature and keep only those that can show differences between data and across all classes. The training dataset was used for this purpose.

The subset of data created after feature selection was used for classification training. The machine learning techniques used for this purpose were based on SVM, kNN, Decision trees and Discriminant Analysis, as presented and described in detail in section 2.9.1. Each classification training was assessed using specific metrics (average accuracy, error rate, precision, sensitivity, specificity, PPV, NPV) which are presented in section 2.9.4. The exported predictive model, following the application of the most efficient machine learning technique, was validated by providing the test dataset. The results of this prediction were used for assessing the performance of the model.

7.3 Results

7.3.1 Data used for supervised machine learning

Following the rejection of poor-quality signals collected from all MEA experiments presented so far, a total of 4,022 EGMs were analysed. These signals were obtained from a total of 24 functional and structural modulations either in isolation or in combination, as presented in Chapters 4-6 (Table 7.1). Regarding the functional abnormalities that were introduced to +20% FB and +40% FB cell cultures, the +40% FB data were combined with the corresponding +20% FB data (High % FB). The reason for this is that the size of data obtained from the +40% FB cell cultures were not adequate to be used both during training and evaluation of the prediction model. In addition, the morphology changes were not significantly different between these data groups at the baseline state or following the introduction of specific types of ion channel blockade, as presented in Chapter 4.

The EGM features extracted from the automated analysis, which was presented in Chapter 3, formed a matrix of 36 features. These EGMs were obtained both during the concurrent MEA/optical mapping experiments (IC₅₀ drug concentration) and during titration experiments for each of the pharmacological agents used in this thesis – data collected after the addition of the drug concentration closest to the IC₅₀ were used here. The data were partitioned into a training dataset, used for classification training, and a

test dataset, used for evaluation of predictive models. The training dataset consisted of 2,400 randomly selected EGMs (59.7% of total EGMs) and the test dataset consisted of the remaining 1,622 EGMs (40.3% of the total) (Table 7.2). The training dataset consisted of 100 EGMs per class, in order to have an equal amount of data representing each condition during classification training and avoid the biased analysis.

Table 7.1. Total number of EGMs collected under a variety of experimental conditions from cell cultures, which were analysed and further used for the development of training and test datasets.

	# of EGMs analysed		# of EGMs analysed		# of EGMs analysed		# of EGMs analysed	
Control (MO)	347	I _{CaL} blockade	113	NC + I _{Kr} blockade	126	High % FB + GJ uncoupling	113	
I _{to} blockade	118	K _{ATP} opening	125	NC + I _{Ks} blockade	111	High % FB + I _{to} blockade	110	
I _{Kr} blockade	112	GJ uncoupling	179	NC + I _{Na} blockade	139	High % FB + I _{Kr} blockade	124	
I _{Ks} blockade	109	NC	493	NC + I _{Na} + I _{CaL}	107	High % FB + I _{Ks} blockade	111	
I _{Na} blockade	149	NC + GJ uncoupling	103	NC + I _{CaL} blockade	104	High % FB + I _{Na} blockade	184	
I _{Na} + I _{CaL}	117	NC + I _{to} blockade	138	High % FB	489	High % FB + I _{CaL} blockade	201	
TOTAL	4,022							

Table 7.2. Number of EGMs collected per experimental condition used for classification training and evaluation of predictive models.

# of EGMs analysed			# of EGMs analysed			# of EGMs analysed			# of EGMs analysed		
Training		Test	Training		Test	Training		Test	Training		Test
Control (MO)	100	247	I _{CaL} blockade	100	26	NC + I _{Kr} blockade	100	26	High % FB + GJ uncoupling	100	13
I _{to} blockade	100	18	K _{ATP} opening	100	11	NC + I _{Ks} blockade	100	11	High %FB + I _{to} blockade	100	10
I _{Kr} blockade	100	12	GJ uncoupling	100	79	NC + I _{Na} blockade	100	39	High %FB + I _{Kr} blockade	100	24
I _{Ks} blockade	100	9	NC	100	393	NC + I _{Na} + I _{CaL}	100	7	High %FB + I _{Ks} blockade	100	11
I _{Na} blockade	100	49	NC + GJ uncoupling	100	3	NC + I _{CaL} blockade	100	4	High %FB + I _{Na} blockade	100	84
I _{Na} + I _{CaL}	100	17	NC + I _{to} blockade	100	38	High % FB	100	389	High %FB + I _{CaL} blockade	100	101
Total # of EGMs for Training dataset			2,400								
Total # of EGMs for Test dataset			1,622								
TOTAL			4,022								

7.3.2 Feature selection for predicting functional and structural modulations

The SFS method for automated feature selection was only applied to the training dataset. The number of features used by the classifier was reduced to a subset with a variable number of features depending on the method used during the SFS process. Features derived from any type of morphology analysis were included in the final feature subsets. Twenty-two features were selected using the LDA method, 12 by applying kNN, 11 using SVM and 12 using the Bagging Decision Trees (TreeBagger) algorithm (Appendix B-I).

7.3.3 Identification of the optimal classifier

Initially, a comparison of 20 machine learning algorithms was undertaken, in order to identify those which are most suitable for the specific classification problem. A training dataset consisting of the 24 classes presented in Table 7.2 (2,400 EGMs) was used for this purpose. The performance of each classifier was assessed by the average classification accuracy (Table 7.4). The highest accuracy was achieved by the Bagging Ensemble method (85.8%) and the lowest classification accuracy corresponded to simple decision trees with coarse distinctions between classes, where the maximum number of splits was 4 (18.2%). Comparing the four families of supervised machine learning methods, it was observed that there was significant difference in average classification accuracy between them. However, the SVM, kNN and Ensemble methods tend to have better performance on the same training dataset (Table 7.3).

Since the Bagging Ensemble method had the best performance, it was chosen for use in subsequent analysis and further optimized. Table 7.4 lists the order that the features were selected during the SFS process. Assessing how each of these features affects classification accuracy, it was found that using the frequency of maximum energy in isolation gave an accuracy of 44.3%, which increased with the step-wise addition of the next 5 features (logarithmic energy entropy, QS interval, mutual information, maximum modulus, frequency of minimum signal energy) reaching the 82.2% of accuracy (Figure 7.1). However, there was not significant change after the addition of the rest of features and the classification accuracy was only increased to 85.8% in the end. That shows that the first 6 features selected during the SFS process were the most discriminative.

Table 7.3. Classification performance for each of the classification methods considered. This is the average classification accuracy which is used as an index of the performance on the overall training dataset.

Prediction method	Accuracy (%)	Prediction method	Accuracy (%)
Ensemble (bagging)	85.8	KNN (cubic)	60.5
SVM (Cubic)	75.2	Ensemble (boosted)	58.2
SVM (Quadratic)	74.6	Ensemble (subspace KNN)	56.8
KNN (fine)	71.1	SVM (Fine Gaussian)	54.1
KNN (weighted)	69.9	Ensemble (subspace discriminant)	53.1
SVM (Linear)	69.9	Decision Tree (medium)	44.5
Decision Tree (fine)	67.7	Linear Discriminant Analysis	44.3
KNN (medium)	64.6	Quadratic Discriminant Analysis	42.6
KNN (cosine)	63.3	KNN (coarse)	39.7
Ensemble (RUSBoosted Trees)	61.4	Decision Tree (coarse)	18.2

Table 7.4. Feature subset extracted by forward SFS using the Ensemble Bagging method during feature selection process listed in order of significance from the most to the least important.

EGM features selected during Sequential Forward Selection	
1. Frequency of maximum energy	7. Variance of signal energy
2. Logarithmic energy entropy	8. Kurtosis of PSD estimates
3. QS interval	9. Scale at minimum signal energy
4. Mutual Information	10. Maximum PSD estimate
5. Maximum modulus	11. Median frequency of PSD estimates
6. Frequency of minimum signal energy	12. S.D. EGM amplitude

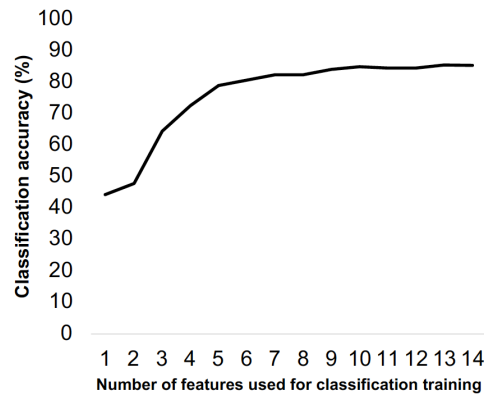


Figure 7.1: Classification performance achieved using in vitro data. Starting from the most discriminative feature, additional features were included during training in the order of selection. The initial increase phase in the curve is followed by a plateau, as the addition of more features does not make significant difference in classification performance.

The parameters of the Bagging ensemble method were optimised. The out-of-bag error, which is the average predictive accuracy for each training sample when not included in the training, was estimated for trees with 5 different leaf sizes (1, 5, 10, 15, 20) and different ensemble sizes (1–500 trees), in order to identify the optimal parameters for the minimization of prediction error (Figure 7.2). Classification training using the minimum leaf size of one resulted to the minimum out-of-bag classification error and it remained stable above 50 trees. The out-of-bag classification error was significantly different across leaf sizes (leaf size 1: 0.16 ± 0.002 , leaf size 5: 0.18 ± 0.002 , leaf size 10: 0.2 ± 0.003 , leaf size 15: 0.22 ± 0.002 , leaf size 20: 0.25 ± 0.002 ; $p < 0.0001$). Based on these findings, a leaf size of 1 and 50 trees were chosen as the optimal conditions for classification training using the Bagging Ensemble method.

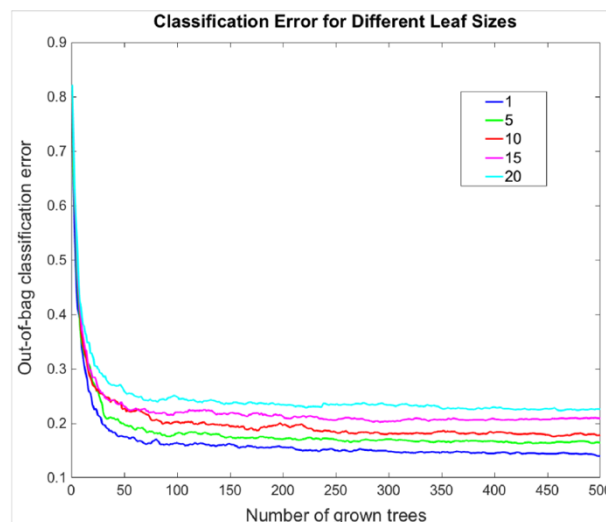


Figure 7.2: Optimisation of the Ensemble Bagging method. Out-of-bag classification errors measured for five leaf sizes (1, 5, 10, 15, 20) as a function of number of grown decision trees. A variety of ensemble sizes with up to 500 decision trees were investigated using the same training dataset.

7.3.4 Predicting functional and structural modulations using cellular EGMs

Classification training was carried out using the optimal conditions for the performance of Bagging ensemble method. This algorithm was applied on the training dataset using 14 features and EGMs were classified against 28 classes of functional and structural modifications. The result was a classification accuracy of 85.8% with error rate of 4.4% and precision 85% (Table 7.5). Figure 7.3 presents the per-class predictions. It is shown that data correlated to 20 modulations were correctly predicted at a rate of >75%. Table 7.5 presents the details of the assessment of classification training performance. It was observed that the sensitivity of the control group for each type of cell culture (MO, NC, +20%FB, +40%FB) was much lower than the rest of modulations within each cell group. In correlation with the confusion matrix from Figure 7.3, the control groups resulted to a higher number of false negative classifications compared to the rest of classes, which could be the reason for low sensitivity. There is high confidence in the results presented in this confusion matrix, since all classes of the training dataset had an equal size ($n = 100$ EGMs per class), in order to avoid any biased predictions.

The receiver operation characteristic (ROC) was used as a more specific measure of accuracy as a function of the false positive rate. It is another way to visualise the classification performance. Figure 7.4 presents the ROC curves for the modifications with the highest true positive/false positive rate, which is gap junction uncoupling in MO monolayers, and the lowest true positive/false positive rate at the +20% FB data. The area under the curve (AUC) is indicative of the overall accuracy of the model with AUC=1 indicating zero errors in classification. AUC for gap junction uncoupling was 0.99 and for +20% FB was 0.89, indicating high prediction accuracy for both conditions. The rest of classes had in-between AUC results.

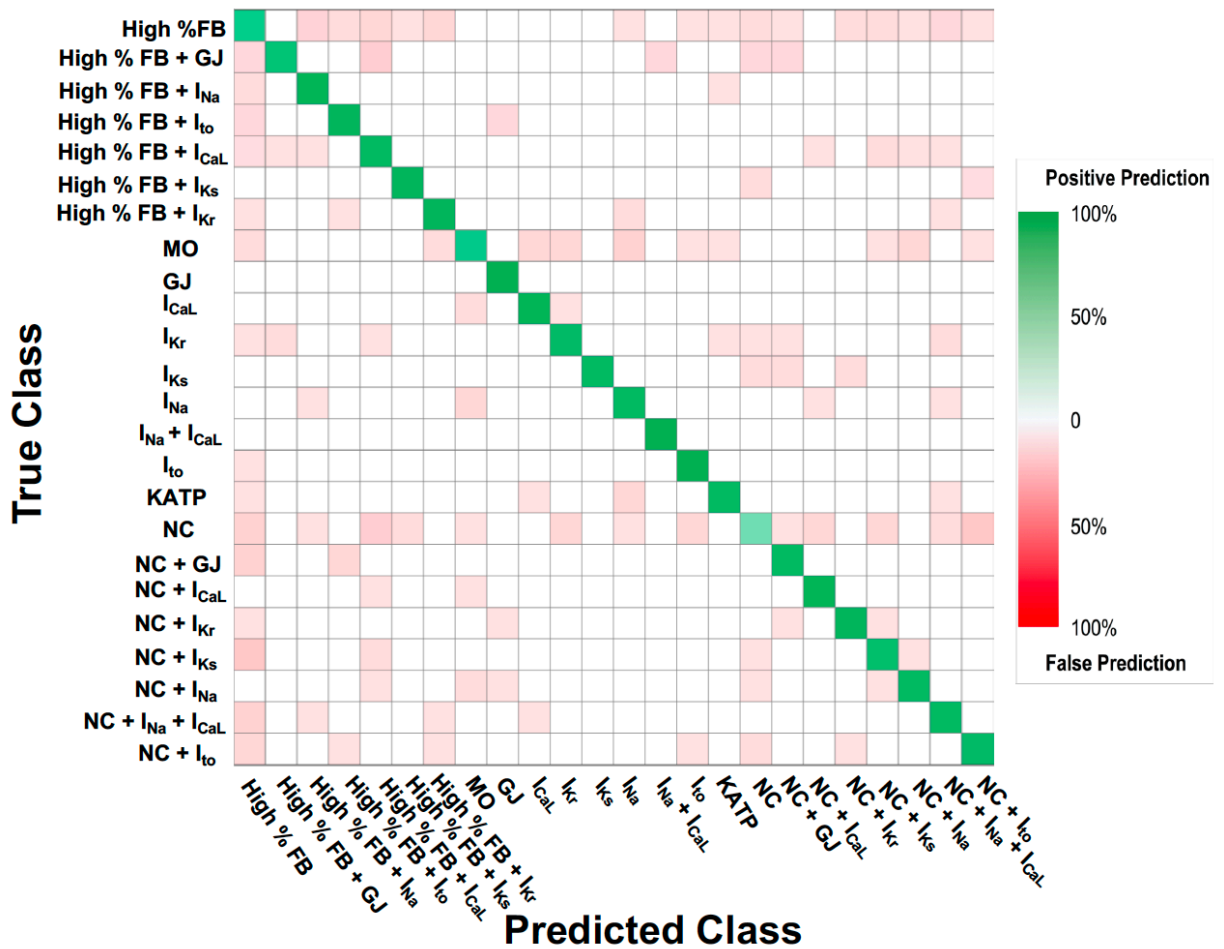


Figure 7.3: Classification performance achieved on the EGM signal dataset by Bagging ensemble classification. The dataset consisted of all functional and structural modifications applied on cell cultures. The confusion matrix presented here was determined by the known and predicted classes using the training dataset. The rows correspond to the true class, the columns show the predictions and the diagonal cells show for what percentage of the observations the trained network correctly estimated the classes. Green cells correspond to correct predictions.

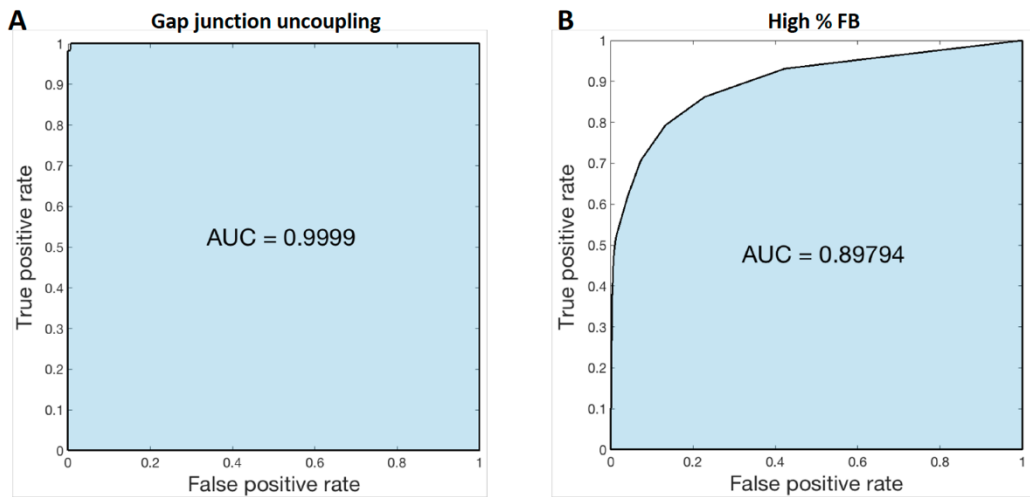


Figure 7.4: ROC curves obtained after classification training using the Ensemble Bagging method. The curves correspond to data obtained from MO culture with gap junction uncoupling (A) and +20% FB monolayers at the baseline state (B), which are the modifications with the most and the least correctly classified EGMs respectively.

The lowest rates of correct predictions were achieved at the baseline class of each type of cell culture. These poor predictions were due to the fact that the EGMs obtained from baseline classes had hardly different SFS features from a group of other classes. This resulted to high false negative rates and consequently low sensitivity. This became obvious on the confusion matrix presented on Figure 7.3, across the row corresponding to each baseline class (MO, NC, High % FB), where red cells indicate horizontally the classes that were false negatively detected. In specific, the baseline of MO cultures was mainly confused with the I_{Na} blockade at MO cell cultures and functional abnormalities at the NC and High % FB cultures. The baseline of NC was confused with various functional abnormalities at NC and High % FB cell cultures. The baseline class of High % FB cultures was poorly predicted due to the misclassification of data mainly for High % FB + I_{Kr} blockade data and EGMs obtained from a variety of functional abnormalities in NC cell cultures. Figures 7.5-7.7 confirm the absence of statistically significant differences between the baseline and the rest of classes for each cell group (MO, NC, High% FB), apart from a few exemptions indicated on each figure. This might be the reason for high false negative rates, the relatively higher number of false predictions for the baseline state of each cell group, as shown in the confusion matrix of Fig.7.3, and subsequently the reduced precision and classification accuracy.

	Control (MO)	K_{ATP} opening	I_{CaL} blockade	I_{Na} blockade	I_{Kr} blockade	Gap junction uncoupling	I_{Na} + I_{CaL}	I_{to} blockade	
Sensitivity (%)	56.7	93	97.7	95.2	100	100	100	92.8	
Specificity (%)	97.1	98.1	97	95.3	96.1	100	100	95.9	
PPV (%)	82.2	55.4	83.3	57.8	66	100	100	65	
NPV (%)	92.9	99.2	99.5	99.6	100	100	100	99.9	
	NC	NC + I_{CaL} blockade	NC + I_{Na} + I_{CaL}	NC + I_{Na} blockade	NC + I_{to} blockade	NC + I_{Ks} blockade	NC + I_{Kr} blockade	NC + Gap junction uncoupling	
Sensitivity (%)	48.4	70.5	88	98	81.7	89.5	93.7	94.2	
Specificity (%)	96.1	94.9	95.2	92	97.3	97.4	98.1	97.9	
PPV (%)	92.3	65.4	53.2	69.5	49.5	61.1	77.2	98	
NPV (%)	78.5	97.8	99.2	99.9	98	99	99.3	97.9	
	I_{Ks} blockade	High %FB	High %FB + I_{CaL} blockade	High %FB + I_{Ks}	High %FB + I_{Na} blockade	High %FB + I_{to} blockade	High %FB + I_{Kr} blockade	High %FB + Gap junction uncoupling	
Sensitivity (%)	98.3	52.4	93.3	90.1	98.2	92.7	87.7	95.1	
Specificity (%)	97.5	98.5	92.2	94.7	92.5	91.3	89.9	99.2	
PPV (%)	85.7	78.2	67.9	60.2	69.5	80	81.4	95.3	
NPV (%)	99.3	89.6	99.6	99.6	99.9	98.3	93.1	99.5	
		Overall performance							
		Precision (%)			85				
		Error rate (%)			4.4				
		Classification accuracy (%)			85.8				

Table 7.5. Assessment of classification training using training dataset. The classification training was carried out against 24 classes. Statistical indices (sensitivity, specificity, PPV, NPV) were calculated separately for each class and average indices (precision, error rate, classification accuracy) were measured for the overall performance.

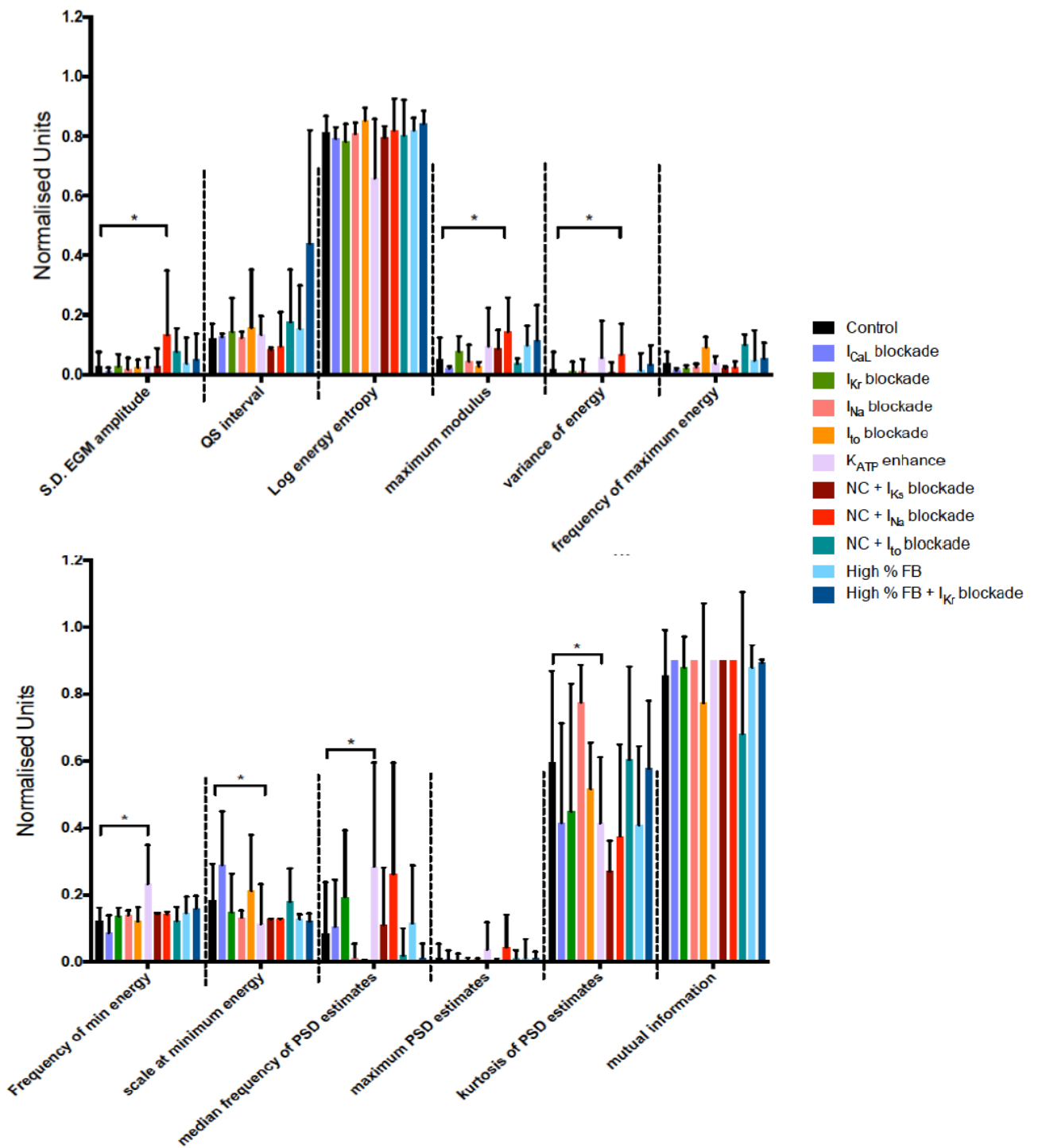


Figure 7.5: Bar graphs presenting how the SFS features used during classification training changed across a range of classes which resulted to high false negative rates, and consequently low sensitivity, in relation to the control (MO baseline) class. Two-way ANOVA and multiple t-test analysis within each group of classes showed no significant changes between control and the rest of classes for any feature, apart from a few exceptions. Classes are grouped together per feature and dashed lines distinguish the analysis per feature. Normalised units correspond to the division of each feature measurement by the highest measurement for that particular feature. Mean \pm SD; * $p < 0.05$.

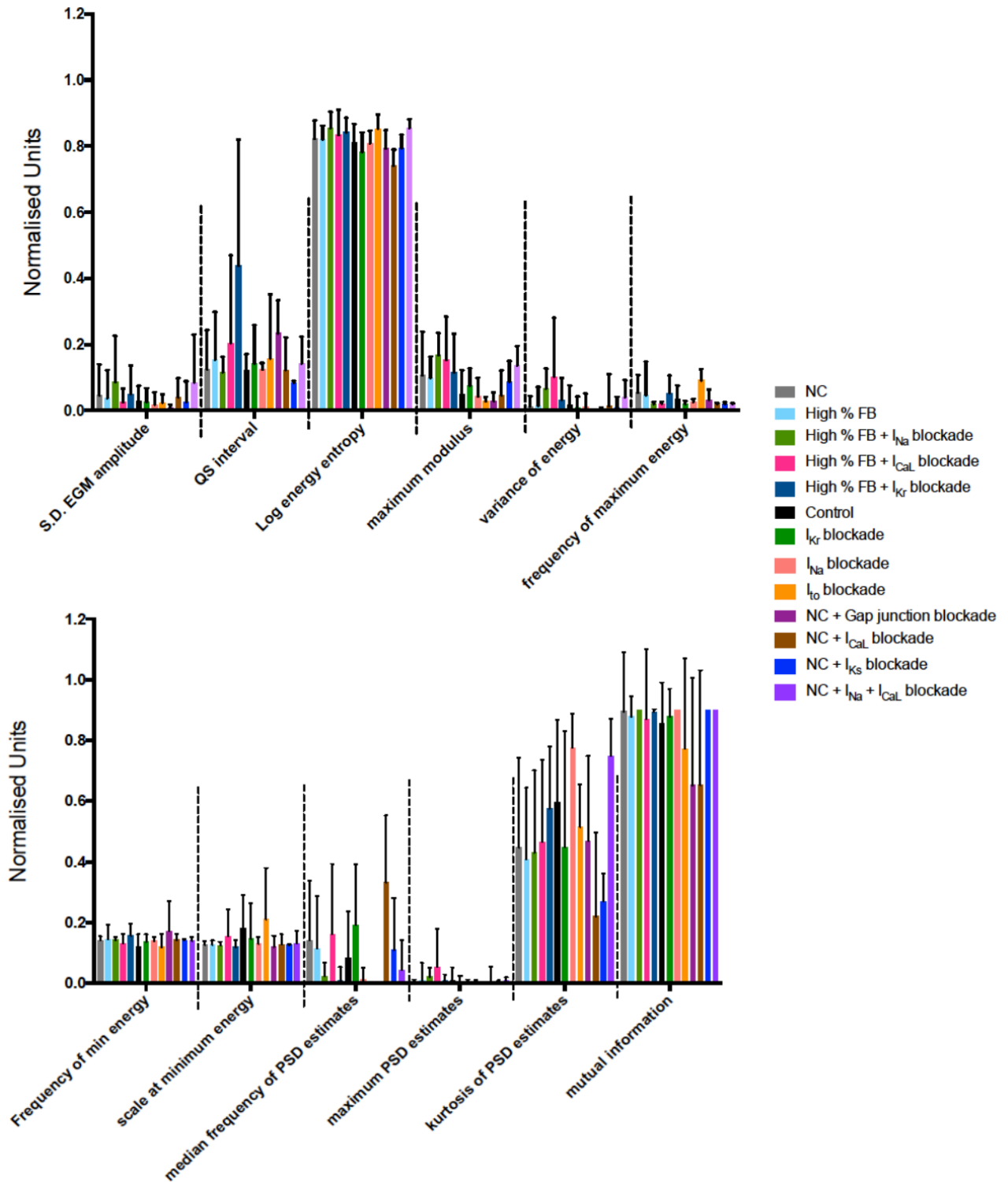


Figure 7.6: Bar graph analysis presenting no significant differences between the baseline state of NC cell cultures and the rest of classes with which a high false negative rate was found comparing the SFS feature changes. Two-way ANOVA and multiple t-test analysis within each group of classes were carried out. Classes are grouped together per feature and dashed lines distinguish the analysis per feature. Normalised units correspond to the division of each feature measurement by the highest measurement for that particular feature. Mean \pm SD.

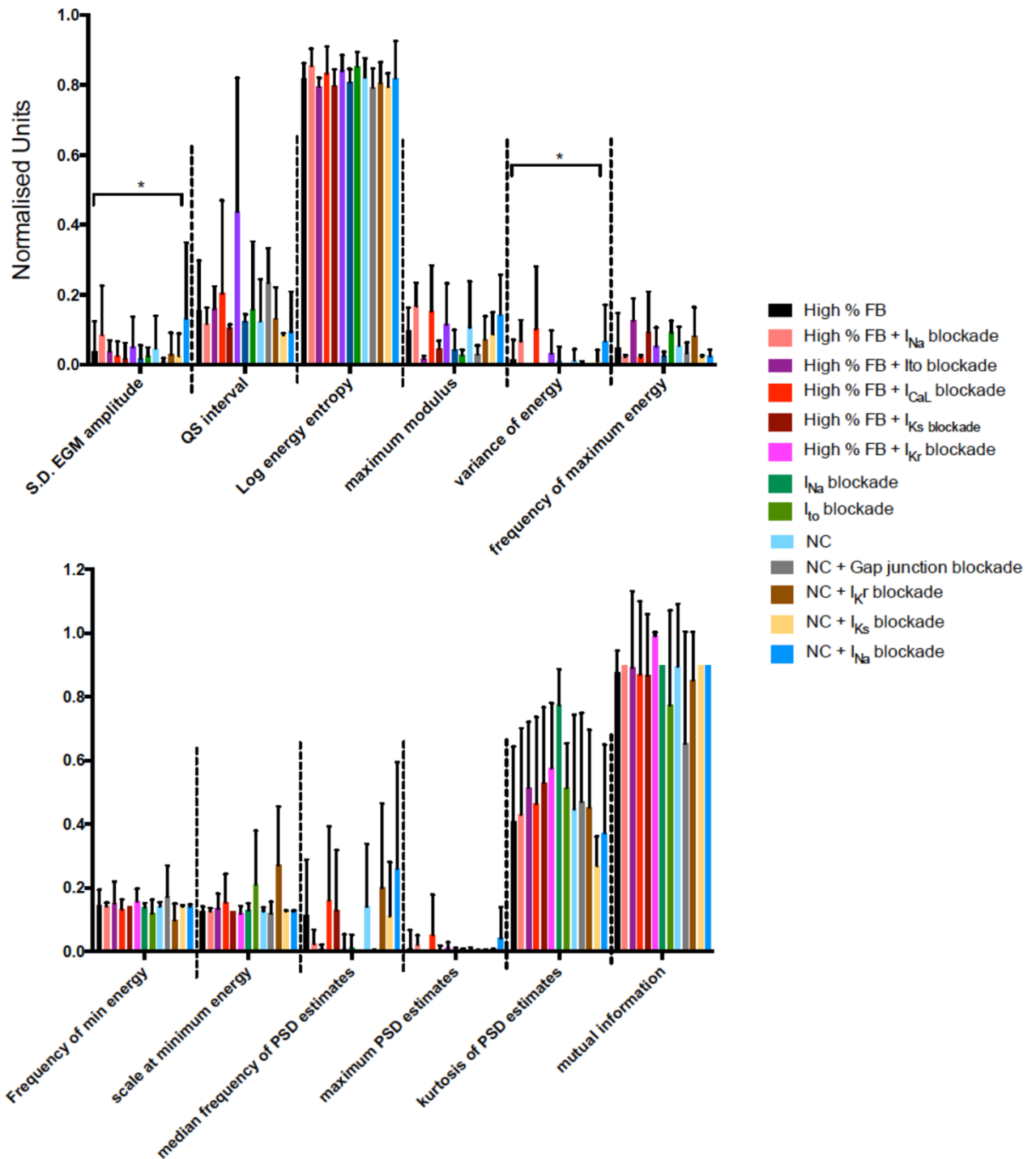


Figure 7.7: High false negative rates for the baseline class of High % FB cell cultures can be explained by the absence of significantly different SFS features between that class and a group of others. There is an exception on this regarding the S.D. EGM amplitude and variance of energy between the High % FB and NC + I_{Na} blockade classes. Classes are grouped together per feature and dashed lines distinguish the analysis per feature. Normalised units correspond to the division of each feature measurement by the highest measurement for that particular feature. Mean \pm SD; * $p < 0.05$.

7.3.5 Validation of prediction model

The prediction model produced at the end of the process described so far was evaluated using a test dataset of 1,622 EGMs. The number of tested EGMs corresponding to each class is presented in Table 7.2. These data were obtained under the same experimental conditions as those that were used for the training dataset and they were not included in the training dataset. The performance assessment demonstrated high classification accuracy in overall (86.6%) (Table 7.6). The data analysis following the prediction model validation agreed with the training results regarding the low predictability of EGMs obtained from MO, NC and High % FB cell preparations at the baseline state, because the sensitivity was lower than the rest of classes (MO: 41.2%, NC: 53.4%, High % FB: 66%).

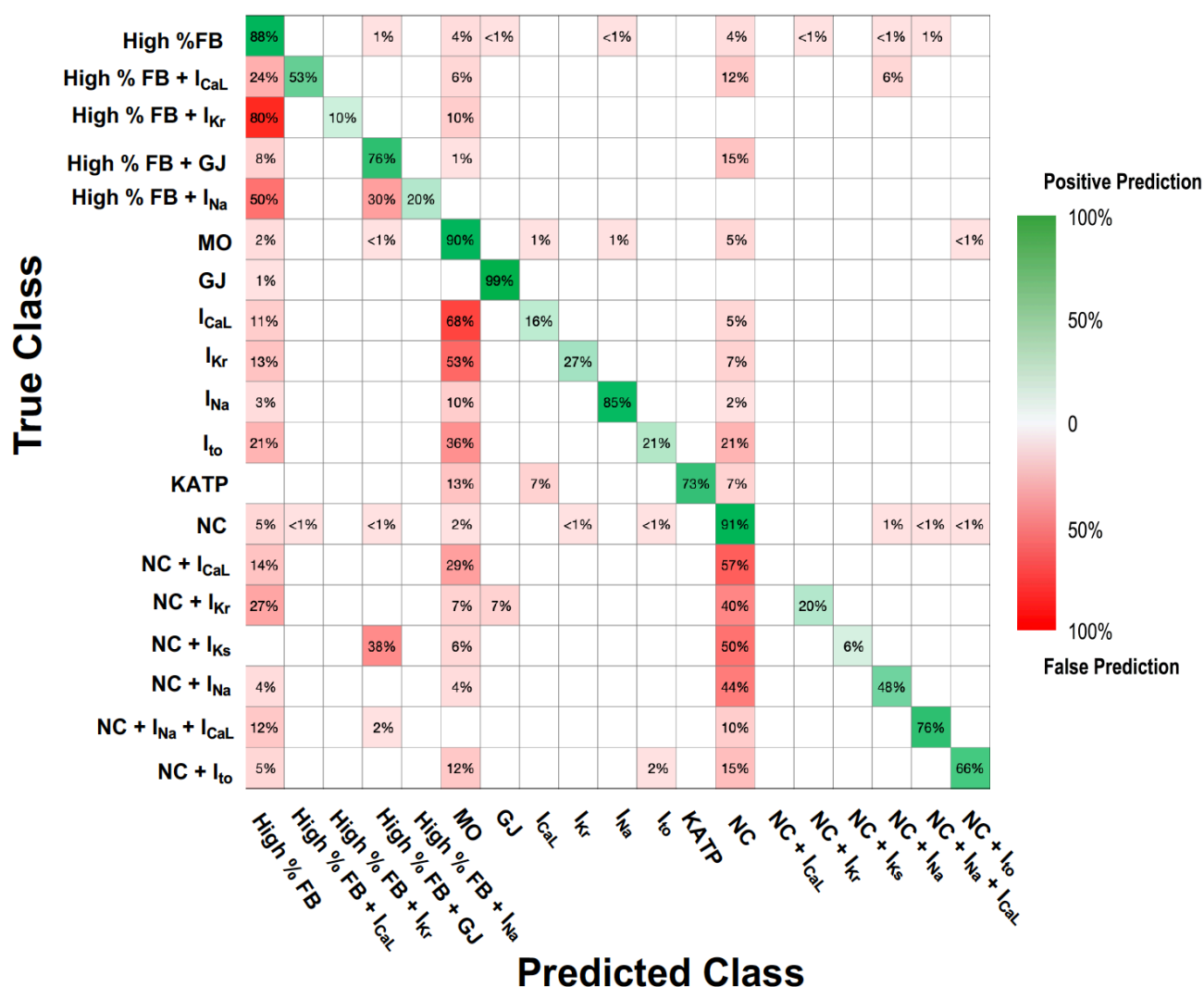


Figure 7.8: Confusion matrix presenting how the predictive model presented in section 7.3.4 performed in each class of the test dataset. The test dataset consisted of 1,622 EGMs derived from functional and structural modifications applied on cell cultures. The confusion matrix presented here was determined by the known (true) and predicted classes using the test dataset. The rows show the true class, the columns show the predictions and the diagonal cells show for what percentage of the observations the model correctly estimated the classes. Green cells correspond to correct predictions. The true positive and false negative rates (correct or incorrect classifications respectively) are presented in each square of the matrix.

Table 7.6. Evaluation performance of prediction model using a test dataset. The model was developed for predicting 24 classes. Statistical indices (sensitivity, specificity, PPV, NPV) calculated separately for each class and average indices (precision, error rate, accuracy) measured for the overall performance.

	Control (MO)	K_{ATP} opening	I_{CaL} blockade	I_{Na} blockade	I_{Kr} blockade	Gap junction uncoupling	I_{Na} + I_{CaL}	I_{to} blockade	
Sensitivity (%)	41.2	93.3	74	95.2	99.8	99.1	95.2	92.8	
Specificity (%)	98.6	98.8	97.1	95.3	96.3	100	96.3	95.9	
PPV (%)	90.3	56	48	57.9	70.1	100	90.9	65	
NPV (%)	84	99.8	99.6	99.7	99.8	99.9	95.3	99.9	
	NC	NC + I_{CaL} blockade	NC + I_{Na} + I_{CaL}	NC + I_{Na} blockade	NC + I_{to} blockade	NC + I_{Ks} blockade	NC + I_{Kr} blockade	NC + Gap junction uncoupling	
Sensitivity (%)	53.4	71.9	88	98	81.7	89.5	93.7	94.2	
Specificity (%)	97	95.1	95.2	92	97.3	97.4	98.1	97.9	
PPV (%)	90	68.2	53.2	69.5	49.5	61.1	77.2	98	
NPV (%)	78.6	97.8	99.2	99.9	98	99	99.3	97.9	
	I_{Ks} blockade	High %FB	High %FB + I_{CaL} blockade	High %FB + I_{Ks}	High %FB + I_{Na} blockade	High %FB + I_{to} blockade	High %FB + I_{Kr} blockade	High %FB + Gap junction uncoupling	
Sensitivity (%)	95	66.0	94.1	95.1	77.2	84.9	80	96.7	
Specificity (%)	96.1	96.4	95.5	95.4	99	97.5	96.6	97.9	
PPV (%)	77.9	85.8	27.1	80.2	80.2	68.4	20	98.2	
NPV (%)	95.9	90.6	99.9	95.6	99.6	98.8	99.8	96.4	
		Overall evaluation performance							
		Precision (%)		79.2					
		Error rate (%)		5.2					
		Classification accuracy (%)		86.6					

The confusion matrix presented in Fig. 7.6 confirms the low predictability of these three classes: MO, NC and High % FB, also presented in Table 7.6. It becomes clear from the confusion matrix that a high percentage of data are incorrectly classified and a result a characteristic column of false negative rates is formed for each one of these classes. The sensitivity and specificity were >80% for the majority of the rest of classes and the overall error rate was 5.9%. This is also visualised on the confusion matrix. Even though the size of test data are unequal across classes, this does not affect the validation results. Each EGM is examined individually against all classes of a predictive model, in order to make a prediction. The equal representation of all classes is crucial in training datasets, which is indeed equal in this case, as already presented in Table 7.2, and not in test datasets.

7.3.6 Predictive modelling differences between functional and structural abnormalities

The predictability of functional modulations in isolation, consisting of ion channel blockade and gap junction uncoupling in MO cell cultures, was also investigated and compared to the predictability of structural modifications. The structural abnormalities consisted of the three different baseline types of cell culture with various amounts of fibroblasts (MO, NC, High % FB). This analysis was carried out in order to search whether such a large number of classes and the high degree of problem complexity, as the one investigated in previous section, are factors that prohibit the performance of classification training.

The same process as described in previous sections was followed for the development of a predictive model focused on functional modulations and 7 features were selected by SFS using the TreeBagger algorithm. The training dataset used for this purpose consisted of 900 EGMs in total obtained from different types of functional abnormalities at MO cultures. These randomly selected EGMs were chosen anew, and the datasets used for classification and validation were different from those used in previous section. Features indicated during the SFS process were used for classification training (Appendix B-II). The average accuracy following the classification training was 96.7% and as it can be observed from the confusion matrix (Figure 7.9A) some classes had a better predictability than others. The results of the training were different for the same classes compared to the ones presented in the previous section. This is due to the use of a different training dataset and the decreased number of classes involved in classification. However, better classification results were achieved (Table 7.7) compared to the use of both functional and structural abnormalities, as presented in previous section. Figure 7.9B presents the ROC curves of an example of a class with high percentage of positive predictions ($I_{Na}+I_{CaL}$, AUC = 0.99) and the class with the lowest percentage of positive predictions, which is the baseline MO cell monolayer (AUC = 0.85). These ROC curves visualise the predictability of each class by showing how the ratio of true positive rate by false positive rate change.

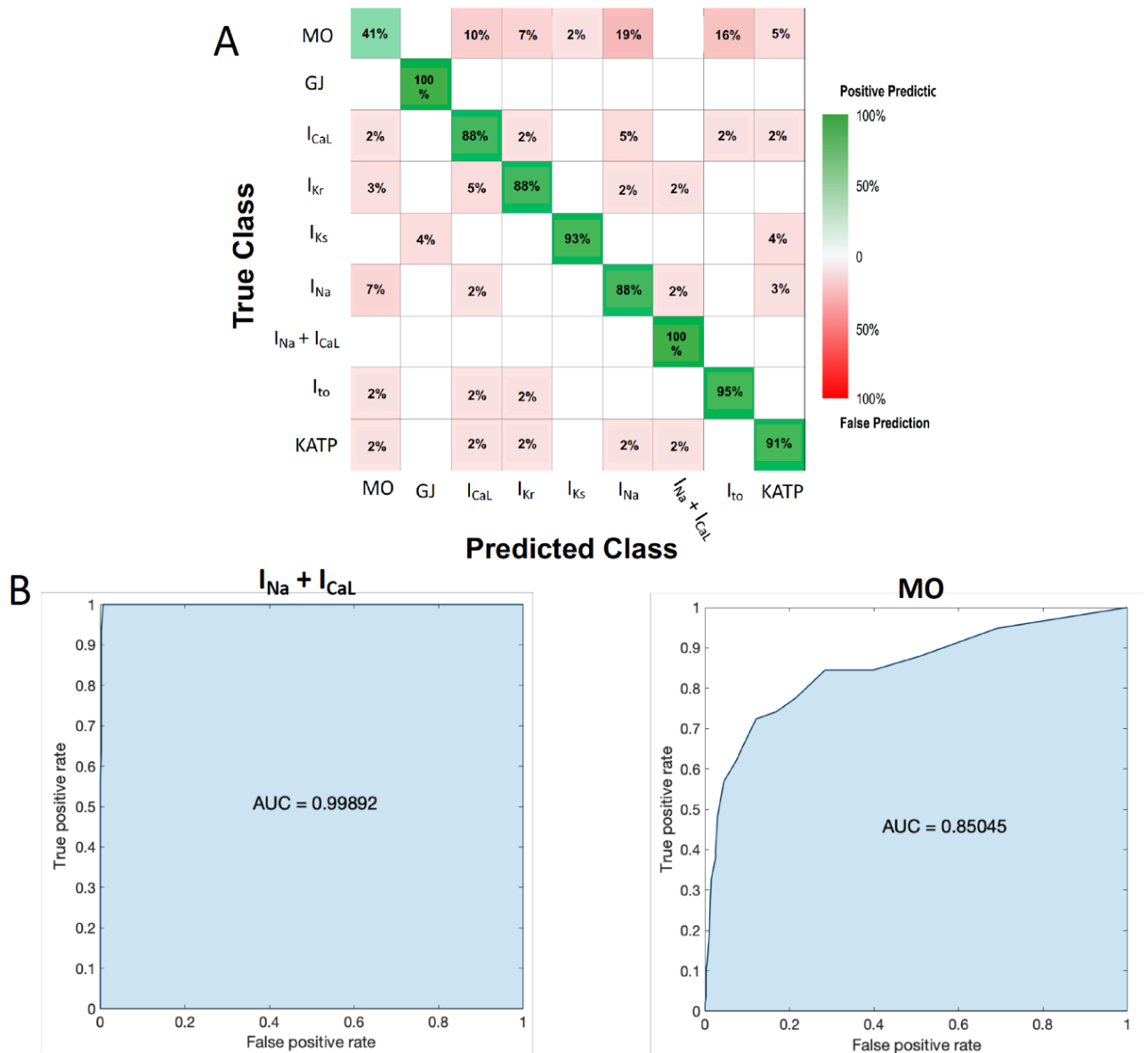


Figure 7.9: Classification performance following training for the functional modifications. (A) Confusion matrix presenting the classification performance using a training dataset consisting of EGMs obtained with or without the presence of functional modulations in MO preparations. The panel was determined by the known and the predicted class data during classification training. The data were described by time-domain, time-frequency, frequency-domain and information theory-based features. Green cells correspond to correct predictions. (B) ROC curves corresponding to the class with the highest ($I_{Na}+I_{CaL}$) and the lowest (MO) percentage of positive predictions.

Table 7.7. The 10-fold cross-validation classification training performance was assessed. Classification performance indices were calculated for the performance per class (Se, Sp, PPV and NPV) and the overall performance (precision, error rate, accuracy).

	Control (MO)	Gap junction uncoupling	I_{CaL} blockade	I_{Kr} blockade	I_{Na} blockade	I_{Na}+I_{CaL} blockade	I_{to} blockade	K_{ATP} blockade	I_{Ks} blockade
Sensitivity (%)	41.4	100	87.9	87.9	86.1	100	94.8	91.4	88.2
Specificity (%)	97.6	99.7	96.8	98.1	95.8	99.3	97.4	98.2	98.3
PPV (%)	72.7	98.3	80.9	87.9	75.8	90.9	84.6	88.3	83.4
NPV (%)	91.6	100	98.1	98.1	97.9	100	99.2	98.7	98.4
Precision (%)	98.1								
Error rate (%)	3.3								
Classification accuracy (%)	96.7								

Table 7.8. Validation performance of the model developed for predicting functional abnormalities using a test dataset with 9 classes. Statistical indices (sensitivity, specificity, PPV, NPV) calculated separately for each class and average indices (precision, error rate, accuracy) measured for the overall performance.

	Control (MO)	K_{ATP} opening	I_{CaL} blockade	I_{Na} blockade	I_{Kr} blockade	Gap junction uncoupling	I_{to} blockade	I_{Na}+I_{CaL} blockade	I_{Ks} blockade
Sensitivity (%)	49.7	66.7	89.4	86.2	86.7	97.5	100	82.2	80.1
Specificity (%)	95.7	91.7	86.7	86.1	91.6	100	88.5	89.5	90.3
PPV (%)	91.4	50.4	52.2	62.6	48.9	100	50.2	85.3	82.9
NPV (%)	68.2	98.4	99.4	97.2	99.4	98.6	100	90.1	92
Precision (%)	60.4								
Error rate (%)	10.5								
Average accuracy (%)	86.8								

The prediction model was validated using a test dataset of 468 EGMs. The performance assessment showed an average classification accuracy of 86.8% and error rate of 10.5%. Apart from the control and K_{ATP} classes, which showed the lowest sensitivity (49.7% and 66.7% respectively), this was >80% for the rest of classes confirming the classification training results (Table 7.8).

In order to investigate the predictability of control groups, where no functional abnormalities were included, the above process was following as well. Again, the data used for representing these classes and for constructing the training and test datasets were different from the ones previously used as they were selected anew. The training dataset consisted of 400 EGMs. The performance analysis following classification training showed that comparable correct predictions were made among classes (Table 7.9). The correct predictions ranged between 71-81% (Figure 7.10). The prediction model was evaluated using a test dataset of 1,029 EGMs. The classification accuracy was 88.1% and the error rate 9.8%. The rest of performance indices (sensitivity, specificity, PPV, NPV) were also similar among classes (Table 7.10). The average accuracy following the classification training for the baseline of different types of cell culture (MO, NC, High % FB) was 84.9%.

True Class	MO	80%	10%	10%
	High % FB	8%	81%	11%
	NC	11%	19%	71%
		MO	High % FB	NC
		Predicted Class		

Figure 7.10: Classification performance following training and using an EGM dataset consisting of data obtained from the four different types of cell monolayer used in this thesis at the baseline state. (A) Confusion matrix determined by the known and predicted groups as developed during classification. (B) ROC curves corresponding to each in vitro experimental model and obtained after classification training.

Table 7.9: Assessment of the classification training performance using EGMs obtained from cell cultures with different amounts of cellular scar. Classification performance indices were calculated for the performance per class (Se, Sp, PPV and NPV) and the overall performance (precision, error rate, accuracy).

	Control (MO)	NC	High % FB
Sensitivity (%)	90.7	84.3	90.8
Specificity (%)	94.1	93.9	97.3
PPV (%)	86.5	84.9	92.4
NPV (%)	96	93.7	96.2
Precision (%)		89.1	
Error rate (%)		6.2	
Classification accuracy (%)		93.8	

Table 7.10: Evaluation of the prediction model based on the different types of cell culture used in this thesis. Statistical indices (sensitivity, specificity, PPV, NPV) calculated separately for each class and average indices (precision, error rate, accuracy) measured for the overall performance.

	Control (MO)	NC	High % FB
Sensitivity (%)	89.6	82.5	87.2
Specificity (%)	89.2	93.6	90.9
PPV (%)	89.3	89.6	66.5
NPV (%)	89.5	89.6	97.2
Precision (%)		82.2	
Error rate (%)		9.3	
Classification accuracy (%)		89.4	

7.4 Discussion

In this chapter, a framework for classification of functional and structural abnormalities that occurred *in vitro* using unipolar extracellular EGMs was presented. Different types of cell culture were used where functional abnormalities were pharmacologically introduced. High-dimensional EGMs were projected to a low-dimensional feature space, and supervised machine learning tools were used to train a model and make predictions. The abnormalities could be predicted with high specificity and accuracy. EGMs were classified into 24 classes using 4 time-domain, 5 time-frequency and 4 information theory characteristics. Twenty feature-based supervised learning methods were compared. The outcome of this investigation is that the EGM morphology can be used to accurately predict a variety of localized cellular abnormalities, but the complexity of a dataset combined with the size of data can be critical factors for the efficiency of a predictive model.

7.4.1 Features describing the datasets and data reduction

EGM correlates of AP changes have been previously characterized (247), but more extensively presented in Chapters 4-6 of this thesis. As a result, the effects of a range of functional and structural modulations can be theoretically predicted. However, the inverse problem of predicting abnormalities from EGM morphology has been studied to a much lesser extent. Previous machine learning studies using unipolar EGM data obtained from AF patients have been published, but these mostly focused on the classification of AF phenotypes based on the presence of scarred regions and they completely ignored the presence of electrical remodelling and any cellular level changes (154,391). As is the case

with the work presented in this chapter, the nature of the outputs becomes multidimensional when multiple experimental conditions are present or a multiplex system is used (392). This chapter reported the use of unipolar EGM morphology for predicting numerous cellular abnormalities, which are known to be linked with a number of heart failure and arrhythmogenic phenotypes. While distinctive from clinical EGMs, the data used in this chapter enabled the capabilities of supervised learning algorithms to be assessed in a controlled context and ensure signals could be labelled accurately. Therefore, an additional advantage of the work presented in this chapter compared to previous studies, which also used *in vitro* data, is that modifications were studied in a step-wise manner. EGMs were obtained not only from isolated modifications, but also from combinations of them.

Feature extraction was carried out based on time-domain morphology analysis (393,394), time-frequency, frequency-domain and non-linear features (125,395). Their combination has been reported to achieve better performance (213). Obtained results show that the mixture of features can more efficiently predict the variety of abnormalities that were investigated. The training results clearly support this statement: selected time-domain features (53.2%), selected time-frequency features (65.3%), combined time-domain and time-frequency features (70.2%). The addition of frequency-domain and information theory-based features achieved the optimum classification performance. These findings agree with clinical studies that had been conducted for simulation modelling (125) or human (396) and animal (397) models, where it was suggested that the combination of EGM features is a promising method for accurate discrimination of arrhythmogenic substrates. This implies that the particular combination of features better describes EGMs derived from specific conditions.

The model reduction, through the automated feature selection, was determined by the TreeBagger algorithm. However, it cannot be concluded whether the SFS features agree with the *a priori* knowledge about effects of specific cellular abnormalities on EGM morphology, due to the large number of abnormalities investigated during classification and the complexity of the dataset. SFS-selected features may be the result of different abnormalities. It should be mentioned though that, regarding the classification based on structural modifications, the R-peak width was included in the final feature subset. Even though this feature was not increased significantly, as shown in the heatmap of section 6.2.2, its prolongation was responsible for the increased EGM onset-to-Rpeak gradient in all NRVM/FB co-cultures presented in section 6.2.1. This feature was also included in the feature subset for classifying tissue slices according to the amount of fibrosis in Chapter 8. Thus, it could be suggested that R-peak width is a strong predictor of the amount of cellular scar through its prolongation under these circumstances. The R-peak prolongation is justified by the conduction slowing that occurs in a region with increased amount of fibroblasts, due to the increased time that it takes for the action potential to pass across an electrode (242). Apart from that feature, there were no similarities between the subsets selected for predicting functional or structural abnormalities in isolation. This could suggest that each

group of modifications is responsible for different effects in electrophysiology and subsequently for changing a different set of EGM features.

However, a number of other features were also identified by feature selection as providing useful information to distinguish abnormalities, even though their relationship with electrical remodelling is not currently understood, as discussed in chapter 4. This suggests that time-frequency and time-domain features both capture valuable information about underlying electro-architectural modifications and should be taken into consideration alongside traditional time-domain characteristics when quantifying abnormalities. Predictions based on a range of these features were characterized by higher sensitivity and specificity.

The correlation of EGM features with the various interventions presented in this chapter, as well as chapters 4-6, and ECG features is not possible. Even though a number of the EGM features used in this thesis were initially studied in ECGs, there have not been successful publications presenting the effect of specific electro-architectural factors on ECG features in a similar way as presented in this thesis. The use of these ECG features is limited in the use of classification and predictive modelling for cardiac diseases, such as AF, without interest on the disease-specific electrophysiology factors responsible for modifying these particular features under these circumstances. For this reason, the EGM approach presented in this thesis using *in vitro* and *ex vivo* data could be considered as superior, because the majority of features used for machine learning can be traced back to cellular level electrophysiology abnormalities, also responsible for a variety of cardiac diseases, that took place when the EGM data were collected. Therefore, the SFS feature subsets presented in this chapter are meaningful and supported by experimental data, like the ones presented in previous chapters.

7.4.2 Classification training for *in vitro* cellular level abnormalities

The Bagging Ensemble method was chosen for all classification cases in this chapter. This algorithm provided a relatively high accuracy of classification and ease of interpretation when compared to other supervised machine learning approaches. A variety of conditions were firstly examined for the effects on predictive performance. The conditions that yielded the best performance in predicting both functional and structural abnormalities achieved a classification accuracy of 85.8%. For some of the modifications, such as gap junction uncoupling, cellular changes were correctly identified with 100% specificity, because of the significant change of the majority of morphology characteristics making those classes easily distinguished from the rest.

The preferred leaf size for classification training was one, as the highest accuracy was achieved under these conditions. This leaf size results in a deep tree with good generalization ability and is correlated to high accuracy, even though it requires more training time compared to larger leaf sizes (193). A large leaf size results in shallow trees, it does not guarantee better classification performance and it is not recommended for complex datasets. Despite the advantage of using small leaf size, as it was presented in this chapter, this strategy has a significant weakness which must not be overlooked. Using a small leaf size and training data with noise or a small training dataset, which struggles to produce a representative sample of the true target function, result to deep trees that overfit the training data. Overfitting can occur in a hypothesis space H , when a hypothesis $h \in H$ gives smaller error than hypothesis $h' \in H$ over the training examples (132). That practically means that an EGM could be misclassified, because the generalisation error is smaller when this classification takes place instead of the right one. This phenomenon is linked to the relatively small size of the training dataset or the fact that each class may consist of a small number of observations. It could be considered that this was the case with the classification training that is presented in this chapter. However, the use of separate training and validation datasets, which was preferred here, is a common approach to avoid overfitting. Even though the learner may be misled by random errors and coincidental regularities within the training set, the validation set provides a safety check against overfitting and is unlikely to exhibit the same random fluctuations of the training set (132).

The sensitivity was much lower for the baseline status of each type of cell culture using the Bagging Ensemble method, as it ranged between 40-70%. The reason for this is the high false negative rate. SFS features were chosen based on the fact that they can maximize the differences among all classes. However, it was proved that it does not mean that the same features are the optimal ones for distinguishing individual features, such as the control cell group, from the rest. The absence of significant differences between the morphology of baseline classes and multiple other classes, focusing on the SFS features, resulted in the inability to efficiently distinguish the classes during training. In addition, the validation of classification training results using the functional modifications in isolation showed that the model is characterised by high accuracy, but low precision (60.4%) and high error rate (10.5%). These additional parameters show that the predictive model based on functional modifications is weak. On the contrary, the model based on the different types of cell culture showed a performance during validation which was as high as during the classification training. A reason for this difference could be the small number of observations per class in the first case compared to the high number of EGMs per class in the second case of structural modifications-based predictive model. The high performance of both training and validation using data derived from all classes could be a compromise for the low predictability of a high number of functional modifications.

The most common misclassifications were the false positive classification of control cells as I_{Na} blockade, the gap junction uncoupling of high %FB cell cultures for the gap junction uncoupling of NC cultures and the false positive classification of NC cultures as variable modifications. Despite using lidocaine and carbenoxolone at concentrations close to their IC_{50} , their cellular effects were smaller compared to the rest of pharmacological agents. This problem could be solved either by computing additional EGM features, training on more data or by applying higher concentrations of these drugs, in order to make the feature changes more prominent. Heylman *et al.* (155) also proposed the same solutions as a way to reduce any misclassifications when applying isoproterenol on hiPS cardiac myocytes. Overall, while the classification performance shows that the prediction model had the capability to predict the abnormalities that take place on the substrate, there are opportunities to further improve model performance and obtain higher assessment scores and reduction in errors.

In general, supervised learning allows data-driven predictions and classification of unknown drug treatments (140), and subsequently functional modifications. Unknown drugs, and subsequently their effects, could be also classified by simple comparison of their effect on specific morphology parameters, but only one-dimensionally on a parameter by parameter basis (155). Supervised learning augments then these individual parameter comparisons by synthesizing all parameter data from multiple drug treatments and determining the relative contribution of each parameter to the classification of each drug. For this reason, multiple pharmacological agents could be used for a specific ion channel blockade and this is something that could be carried out in the future for the improvement of the predictive model presented in this chapter (155). This is important because many drugs may affect multiple EGM parameters and to varying degrees, due to their low level of specificity for an ion channel, such as 4-aminopyridine (398).

7.5 Limitations

A potential limitation of the work presented in this chapter is the limited number of modifications that were investigated compared to the complexity and variety of arrhythmogenic conditions that take place in the heart. Due to the fact that many events of electrophysiological remodelling coexist, it is difficult to distinguish the contribution of each abnormality (399). Thus, prediction models should be developed in future which incorporate the information about EGM modifications when more combinations of ion channel blockade simultaneously occur or functional abnormalities combined with structural modulations, such as cellular anisotropy or patterning. In addition, the present predictive models presented in this chapter could be further optimised by including a larger number of EGMs per class, in

order to improve the predictability of such a large number of modifications and avoid problems, such as overfitting and low precision of the model.

Classification training was carried out using the algorithms provided on MATLAB. However, there are supervised learning algorithms which were not investigated in this thesis. Such a classifier is the *optimum path forest* (OPF), which is considered to be a fast, simple, multi-class and parameter independent method (400). Another method that was not tested was *adaptive boosting* (AdaBoost), which is an ensemble learning algorithm that can be used for multi-class training and is more resistant to overfitting than many other algorithms (401). Such methods can be impressively efficient and significantly reduce the computational time when large datasets are used.

7.6 Conclusions

This chapter presented the successful prediction of a variety of functional and structural modifications, either in isolation or combinations of them, applying supervised machine learning algorithms on unipolar extracellular EGMs obtained *in vitro*. The prediction of these abnormal conditions using unlabelled data was carried out with high sensitivity and specificity. A combination of selected features which were extracted using different types of signal analysis were necessary, in order to achieve high classification performance. A biological relationship was found between SFS-selected features and specific predicted abnormalities, which was explained by findings presented in early chapters of this thesis. The overall goal of this approach and application of machine learning is to render contact EGMs more useful for more effective diagnosis of cardiac diseases and personalized pharmacological and interventional antiarrhythmic strategies.

Chapter 8

Predicting fibrosis from *ex vivo* and *in vivo*
obtained EGM morphology

8.1 Introduction

8.1.1 Applications of supervised machine learning in cardiac electrophysiology

Even though EGM morphology has been poorly correlated to specific electro-architecture determinants, it has been used for correlations to arrhythmic conditions and mainly for predicting AF phenotypes and AF-related structural abnormalities. The mechanism of AF remains unclear, but it has been suggested that EGM morphology during AF may be correlated with different types of conduction, such as conduction slowing, a collision of activation waves or reentrant circuits and conduction blocks (93). High-frequency EGM recordings or signals with chaotic patterns are also associated with AF (70,402). Therefore, catheter ablation strategies which are EGM-guided are developed as an alternative to ablations targeting CFAE or localized reentrant sources driving AF (403). The lack of a strict CFAE definition leaves the decision of selecting the target sites for ablation on the expertise of the electrophysiologist and leading to inefficient ablations (153,404).

Fractionation has been proposed as an alternative method for identifying possible target areas (405). Since the simplistic characterization of the different levels of fractionation has been inadequate though, a recent study by Orozco-Duque *et al.* (153) suggested different types of analysis of CFAE recordings, including time-domain, frequency-domain and time-frequency analysis. This aimed to distinguish each level of fractionation by building a single map encoding waveform differences of CFAE upon the anatomical atrial structure. Since the EGM signal can be labelled for different types of fractionation, supervised learning has been considered the most appropriate machine learning method (153).

Despite the fact that contact EGMs are useful for performing ablation procedures, the analysis of the ECG is more prevalent, due to the wide availability of this type of data. The majority of machine learning applications in cardiac electrophysiology are based on ECG (156). ECG morphology has been used for distinguishing VF and VT arrhythmic events (171) as well as AF (406).

8.1.2 Clinical relevance of classifying cardiac arrhythmias

There is a large number of publications, especially over the last decade, which present applications of machine learning methods in cardiology. There is a flexibility in the type of data to be used, and thus imaging data, i.e. LGE-MRI and echocardiography, and electrophysiology data, i.e. EGM, ECG and optical imaging, have been used for predictive modelling. Machine learning offers the ability to make

predictions for the improvement of diagnosis and the optimisation of treatment methods for cardiac arrhythmias.

The most routine ablation strategy for treating AF is the pulmonary vein isolation. The combination of CFAE targeting with pulmonary vein isolation though is increasingly used, as it is suggested to reduce the long-term AF recurrence (38). However, this approach is debatable due to the uncertainty on interpreting CFAE morphologies (36). For this reason, there are published methods such as the semi-supervised mapping (153) and the application of supervised learning algorithms that use features derived from combined extraction approaches (154), which can locate the distribution of different fractionated EGM patterns on the atria and improve the outcome of ablating drivers that maintain AF in persistent AF patients. Another example of machine learning application in cardiac electrophysiology is a recent study by Hajimolahoseini *et al.* (169). They presented an algorithm to extract the active intervals of intracardiac bipolar EGMs during AF. The analysis included the characteristics of the signal waveform at its inflection points (IPs), which are the points at which the concavity of intracardiac EGMs changes from being convex to concave or vice versa. Signal features included the distance between the consecutive IPs, the slope of signal waveform at the IPs and the energy concentrated between them, which were all inspired by the way the electrophysiologists annotate the EGMs visually. The data were then used for an Expectation Maximization algorithm for Gaussian mixtures for automatic clustering. The important outcome of that study was the significant reduction of the mean computational time to 31ms per 1s recordings obtained from 10 channel EGMs which makes the algorithm able to be applied in real-time clinic applications.

In addition to AF, VF and VT are studied for improvement of their diagnosis by applying machine learning techniques. Both VF and VT are dangerous arrhythmic events leading to sudden death if no defibrillation shock is applied to the subject within a few minutes (170). A number of algorithms have been proposed for the efficient VF/VT detection based on processing the ECG signal, mainly in the time-domain (171). The reason for this is that correct detection and classification of these arrhythmic events is of extreme importance for an automatic external defibrillator and patient monitoring.

8.1.3 Experimental models for the development of prediction models

There are numerous studies presenting ECG or EGM data obtained from patients with different types of cardiac arrhythmia or heart failure. It is understood that these data are directly linked to pathophysiological or electro-architectural abnormalities that may occur during an arrhythmic event. Even though extensive research into the cardiac arrhythmia and heart failure mechanisms has been carried out using animal models as well, it is rare conducting these studies in the human myocardium.

This is due to the limited availability of human samples and the lack of representative experimental models that can be prepared from small biopsies (20). The use of intact sections of the human myocardium is an alternative solution to this problem, which has been increasingly preferred over the last decade. There have already been studies using samples derived from the trabeculae (376), papillary muscles (407) and ventricular wedges (349).

Another approach is using thin slices of viable cardiac tissue. These slices retain the structural and function properties of native myocardium and remain viable for several hours *in vitro* (323,408). There is a number of studies presenting electrophysiological data obtained from this type of experimental model. It has been showed that tissue slices obtained from adult human heart failure samples maintain electrophysiological, structural and biochemical characteristics representative of native myocardium (20). Regarding EGM morphology, Camelliti *et al.* (20) observed that the FPD recorded in slices was similar to APD measured under comparable experimental conditions using intracellular microelectrodes or optical mapping in multicellular preparations, such as papillary muscles and ventricular wedges. In addition, compared with single cells or cell culture, slices benefit from inclusion of the several cell types that are included in the native tissue, such as myocytes, fibroblasts and endothelial cells (20). These are advantages of slices over single cell experimental models and cell culture, as their electrophysiological characteristics reflect a more *in vivo* profile.

Machine learning processes need large datasets, in order to generalise sufficiently to make accurate predictions on new observations. The collection of data from tissue slices exposes a number of technical challenges. This is not only due to the limited availability of cardiac samples, that can be used for tissue slices, but also due to the fact that consistent vital cardiac slices with uniform thickness and minimal tissue damage are necessary for electrophysiological experiments and consistent signal recordings. Moreover, sufficient oxygen diffusion to all cells inside the section has to be maintained throughout recordings (324). These conditions make slices a less attractive experimental model compared to tissue wedges, cell culture and isolated cells, and result to small datasets. It becomes clear that the collection of slice data for machine learning applications is a demanding task and for this reason, there are very few published studies using this model. The collection of signal recordings from patients may be an easier task, as intracardiac EGMs can be obtained during catheter ablation procedures (153). The disadvantage of this data collection process though is the variability of the tissue substrate, such as wall thickness, the clinicopathological characteristics of the subject and any functional abnormalities, i.e. ion channel, gap junction abnormalities, which are virtually impossible to be measured *in vivo*. Therefore, *in vivo* EGM datasets are small, as they include only a few hundreds of EGMs, collected from a relatively small number of patients (153,391), when compared to the ECG datasets which are larger and contain hundreds to thousands of data recordings (409–411).

In this chapter, I sought to investigate the prediction of structural determinants of EGM morphology using supervised machine learning methods. For this purpose, the aims of this chapter were to:

1. group the *ex vivo* data using a structure-based system,
2. search the supervised learning algorithms, in order to find the most suitable for the *ex vivo* or *in vivo* data,
3. optimize the learning process for the development of a predictive model, and
4. assess the classification performance and validate the extracted model.

The methods used in this chapter for the elucidation of the above aims have already been described in detail in Chapter 2 of this dissertation. These include the collection of LV slices from end-stage heart failure human hearts and one rejected donor heart and SHG microscopy for obtaining data regarding the collagen disposition in slices. Even though, the *ex vivo* data presented in Chapter 6 were organised into three groups, based on the percentage of collagen coverage, it was preferred to use different groups and perform binary classifications for this chapter. This was forced by the necessity for an adequate number of training data and in order to investigate the role of collagen coverage in more detail. Furthermore, clinical recordings were used for the purposes of this chapter and the application of supervised machine learning algorithms for classification training and predictive modelling was carried out.

8.2 Results

8.2.1 Collagen coverage: the fibrosis measure on tissue slices

In order to quantify fibrosis, the area of tissue slice covered by collagen type I molecules was measured using the data collected during the performance of SHG microscopy, as described in section 2.7. These data were imported to FIJI for the quantification of the size of collagen-covered area in the same way as presented in section 2.6.4, where a threshold was set on the image and that was adjusted for the signal collection solely from collagen. These measurements were derived using SHG images obtained from heart failure (n=29) and donor heart (n=2) slices. Most of the slices were used for EGM recordings (n=24), but there was a minority which could not be stimulated (n=7). It was found that EGMs were obtained from slices with a collagen coverage of 4-24% of the slice area, while slices with >25% coverage could not be stimulated.

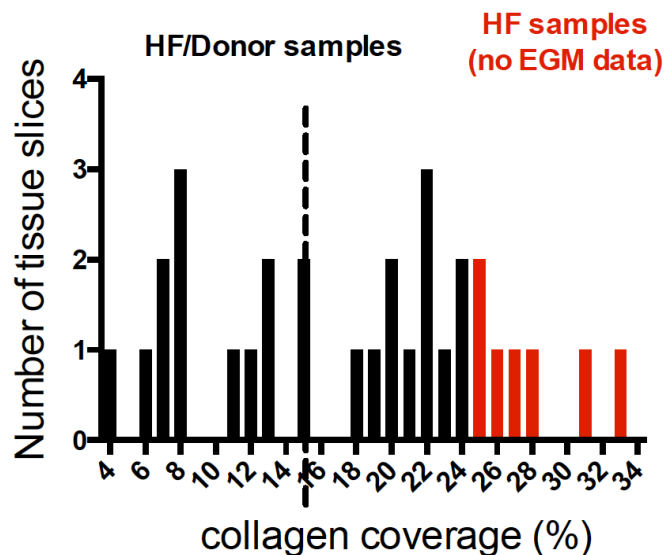


Figure 8.1: Frequency distribution graph of tissue slices used in this thesis based on the percentage of the area of each slice covered by collagen. The majority of slices (black; n=24) could be used for EGM recordings. The dashed line represents the middle point of the group of these samples corresponding to 15% collagen coverage. There were slices (red; n=7) that could not be stimulated, and no EGM recordings were obtained from them.

8.2.2 Feature selection using *ex vivo* EGMs

Based on the collagen coverage data, a similar number of slices were stimulated on either side of 15% coverage (<15%: n=11 slices; >15%: n=13 slices) providing EGM datasets of a similar size. Therefore, it was attempted to automatically differentiate EGMs from tissue with less than 15% collagen coverage (153 EGMs) from those with more than 15% collagen coverage (187 EGMs). The training dataset consisted of 340 EGMs in total, which was 57% of the total amount of EGMs obtained from tissue slices. The rest 43% (147 EGMs) were used for the test dataset. The selection of data for the training and test datasets was randomly performed from the pool of EGMs.

The training dataset was used for feature selection, in which case the SFS method was applied. Five different classification methods were used during feature selection (Bagging Decision Trees, SVM, Decision Trees, kNN and LDA) with each one giving a different feature subset (Appendix B-III). Features extracted from all types of morphology analysis were chosen.

8.2.3 Optimising the classification training process

An assessment of 20 supervised machine learning algorithms, available on MATLAB, was made and these were compared using the average classification accuracy (Table 8.2). Even though the followed process was the same as previously presented in Chapter 7, data obtained from a different experimental model were used for this piece of work and different feature subsets, consisting of less features, were selected during SFS. Thus, it was necessary to repeat the modelling optimization procedures. The default classification conditions with 10-fold cross-validation were used during this initial assessment. It was observed that the highest classification accuracy was achieved using the Bagging Ensemble method (88.6%) and the lowest accuracy was achieved by the coarse kNN method (64.1%).

Based on its performance, the Bagging Ensemble method was chosen for further optimization and classification training. It was observed that only three of the features, chosen by applying the TreeBagger algorithm for SFS, were significantly different between classes: RR interval, logarithmic energy entropy and R-peak width (Figure 8.2).

Table 8.1. Classification performance presenting the accuracy after applying 20 classification training methods which are available on MATLAB. A 10-fold cross-validation was applied in each case. The average classification accuracy is presented here which is used as an index of the performance on the overall training dataset. The highest accuracy was achieved by the Bagging ensemble method.

Prediction method	Accuracy (%)	Prediction method	Accuracy (%)
Decision Tree (complex)	83.7	kNN (fine)	84.9
Decision Tree (medium)	83.7	kNN (medium)	80.1
Decision Tree (simple)	76.5	kNN (coarse)	64.1
Linear Discriminant Analysis	67.7	kNN (cosine)	78.5
Quadratic Discriminant Analysis	72.5	kNN (cubic)	76.1
SVM (Linear)	74.1	kNN (weighted)	84.9
SVM (Quadratic)	85.7	Ensemble (boosted)	86.1
SVM (Cubic)	86.1	Ensemble (bagging)	88.6
SVM (Fine Gaussian)	79.3	Ensemble (subspace discriminant)	70.9
Ensemble (subspace kNN)	84.9	Ensemble (RUSBoosted Trees)	82.9

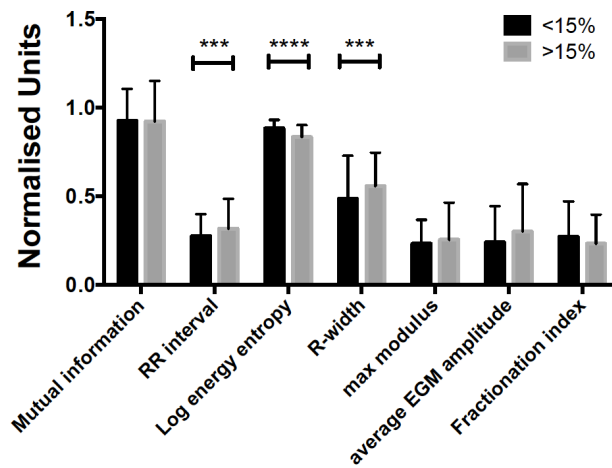


Figure 8.2: EGM properties modified due to differences in the size of collagen coverage, as measured on data obtained from human left ventricle slices (n=24). The data presented on the bar graph belong to the training dataset and are categorized as obtained from slices with collagen coverage below or over 15% of the total sample area (n=187 and n=153 EGMs respectively). Feature measurements are normalized. Mean \pm S.D. ***p<0.001; ****p<0.0001.

In order to optimize the algorithm, the out-of-bag error was estimated for trees with leaf sizes of 1, 5, 10, 15 and 20 and a tree ensemble ranging between 0-500 trees (Figure 8.3A). It was observed that the leaf size of 1 is the best for achieving the least out-of-bag classification error and that remained relatively stable above 160 trees. The out-of-bag classification error was significantly different between leaf sizes (leaf size 1: 0.09 ± 0.02 , leaf size 5: 0.1 ± 0.03 , leaf size 10: 0.13 ± 0.03 , leaf size 15: 0.14 ± 0.03 , leaf size 20: 0.14 ± 0.02 ; p<0.0001) (Figure 8.3B).

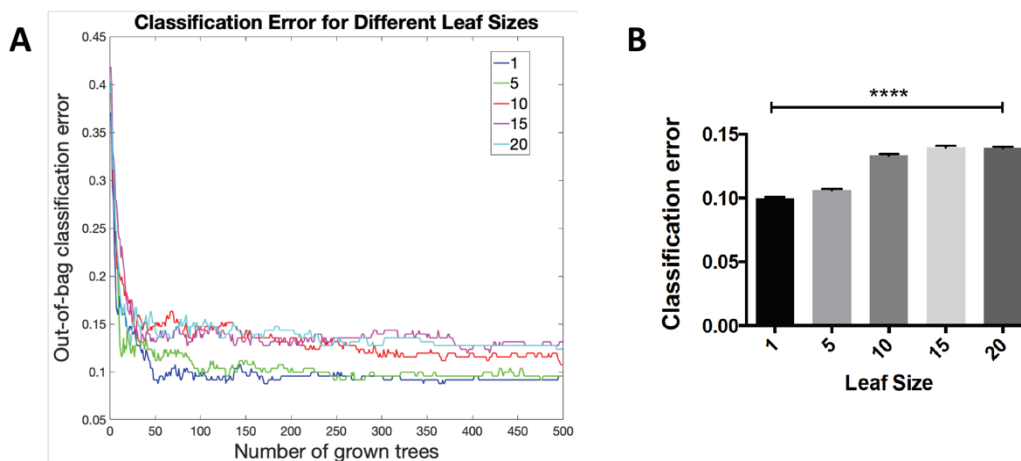


Figure 8.3: Optimisation of the Ensemble Bagging method. (A) Out-of-bag classification error measured for five leaf sizes (1, 5, 10, 15, 20) as a function of number of grown decision trees. A variety of ensemble sizes with up to 500 decision trees were investigated using the same training dataset. (B) Kruskal-Wallis analysis comparing the classification error values for each leaf size.

8.2.4 Predicting fibrosis levels using *ex vivo* data

Classification training was carried out using the training dataset and the optimized classification conditions. The Bagging ensemble method was applied on the dataset which was characterised by 6 EGM features (Table 8.2). The classes were determined by 15% collagen coverage (<15% / >15%). The classification training results showed that a classification accuracy of 90% was achieved with 9.9% error rate and 90.9% precision. Correct classifications were carried out for 95% of <15% data and 84% of >15% data (Figure 8.4A). The high classification performance was also confirmed by ROC curves created for each class (Figure 8.4). AUC for <15% was 0.89 and for >15% was 0.9 meaning the high predictability for both classes.

Table 8.2. Feature subset extracted by forward SFS using the Ensemble Bagging method during feature selection process listed in order of significance.

EGM features selected during Forward Feature Selection	
1.	Mutual information
2.	RR interval
3.	Logarithmic energy entropy
4.	Maximum modulus
5.	Average EGM amplitude
6.	R-peak width

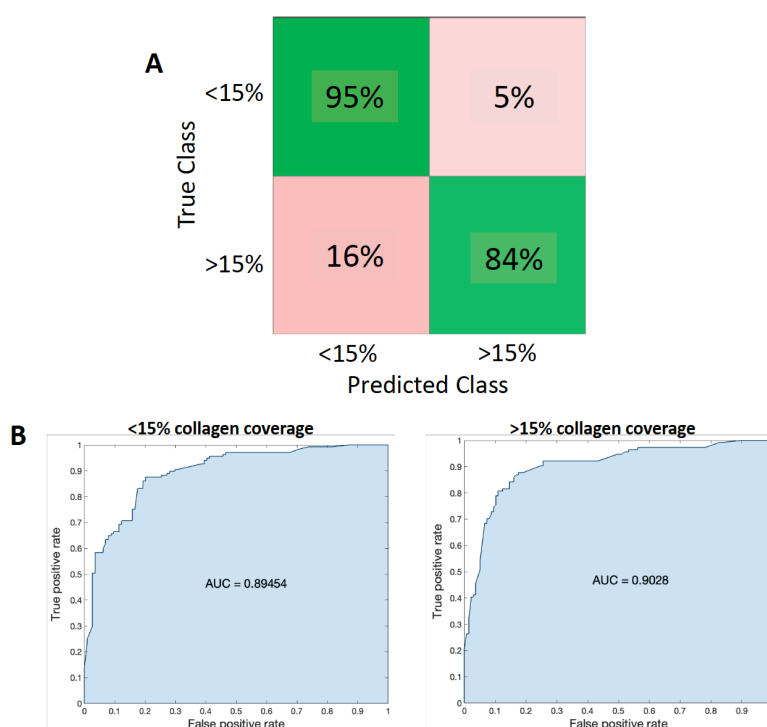


Figure 8.4: Classification training performance using instances classified as obtained from tissue slices with collagen coverage below or above 15% of the total slice area. (A) Confusion matrix presenting the classification performance. The panel was determined by the known and the predicted class data during classification training. Green cells correspond to correct predictions. (B) ROC curves corresponding to each class, <15% (left) and >15% (right) collagen coverage.

The deterioration of classification training was also assessed by modifying the number of EGM features. Table 8.2 presents the order that features were selected during SFS. In order to investigate this deterioration and find the extent of classification accuracy modification due to each feature, multiple iterations of classification training were carried out by removing each time one of the features starting from the most significant one and followed by the less important ones. In the end, it was observed that there was a linear increase trend for classification accuracy. Mutual information, when used in isolation, was responsible for an accuracy of 51.8%. This accuracy increased to 90% with the step-wise addition of individual features (Figure 8.5).

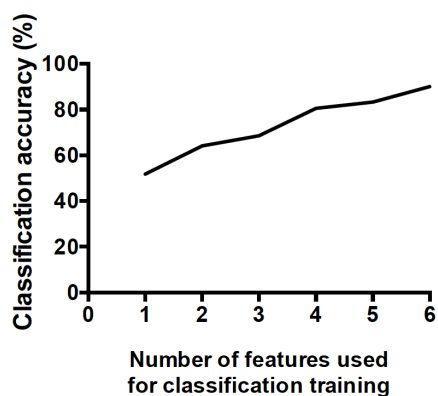


Figure 8.5: Classification performance achieved using *ex vivo* data classified as obtained from tissue slices with <15% or >15% collagen coverage. Starting from the least important feature which was selected last, more features were included on training keeping the order of importance at the same time. This resulted to an increase of classification training accuracy.

The prediction model that was extracted in the end was assessed by measuring a variety of classification performance indices (Table 8.3). Due to the binary classification that was carried out, the performance quantifications were similar for both classes. This is explained by the way that sensitivity and specificity are measured, according to the equations presented in Table 2.3. The sensitivity for the <15% data was 96.4% and the specificity was 82.5%, while the sensitivity for the >15% data was 82.5% and the specificity was 96.4%. The overall average classification accuracy was 90% with an error rate of 9.9%.

The assessment of classification training was followed by the validation of the predictive model. For this purpose, a test dataset of 146 EGMs was used. The performance assessment showed an average classification accuracy of 86.2% with an error rate of 13% (Table 8.4). The positive predictive and negative predictive values were 92.8% and 79.4% respectively for the <15% data and 79.4%. The same indices were 79.4% and 92.8% for the >15% data.

Table 8.3: Assessment of the 13-fold cross-validation classification training performance. Sensitivity (Se), specificity (Sp), Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were calculated separately for each class and average indices (precision, error rate, classification accuracy) were measured for the overall performance.

(%)	<15%	>15%
Se	96.4	82.5
Sp	82.5	96.4
PPV	86.8	95
NPV	95	86.8
Precision	90.9	
Error rate	9.9	
Classification accuracy	90	

Table 8.4: Evaluation performance of the model developed for predicting the collagen coverage on tissue slices. The assessment was carried out either separately for each class or for the overall performance depending on the statistical index.

(%)	<15%	>15%
Sensitivity	85.6	89.3
Specificity	89.3	85.6
PPV	92.8	79.4
NPV	79.4	92.8
Precision	86.1	
Error rate	13	
Average accuracy	86.2	

8.2.5 Including conventional scar markers at prediction model

It was investigated whether the scar markers of voltage and fractionation play a role for the prediction of collagen coverage using *ex vivo* data. For this reason, the same training dataset was used, but the data were characterized by features in three different ways: a) using the SFS features, b) the EGM amplitude and fractionation only or c) a combination of the SFS features alongside fractionation, as the average EGM amplitude had been already chosen during feature selection.

The results of this three-way classification training are presented in Figure 8.6 and Table 8.5. The classification training performance was worse when the traditional scar markers only were used, because the average accuracy was 55.3% and the error rate was 42.5%. The low performance was also depicted on confusion matrix which shows that 40% of the <15% EGMs and 52% of the >15% EGMs

were misclassified. However, the classification training performance was similar with or without fractionation (average accuracy: 86.2% and 85.4% respectively; error rate: 13% and 13.3% respectively).

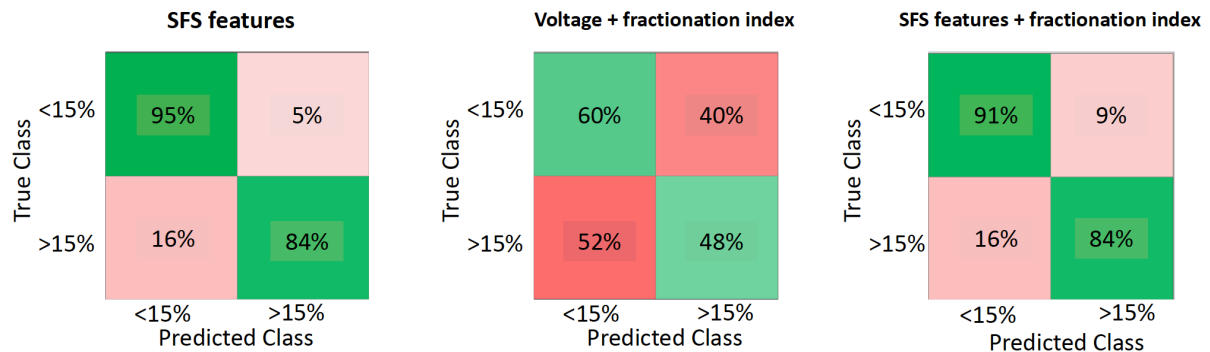


Figure 8.6: Confusion matrices extracted from classification training with the same training dataset but using different features each time. The EGM parameters used for training were the SFS features including EGM amplitude (left), voltage and fractionation (middle) or the fractionation in addition to the SFS features (right). The best performance was achieved using the SFS features and the worst using the voltage and fractionation only.

Table 8.5: Assessment of classification training and evaluation of subsequent prediction models using the training dataset and different feature subsets on each case. The features used were the SFS features, voltage and fractionation in isolation or the SFS features alongside fractionation.

Classification training			
(%)	SFS features	Voltage + Fractionation	SFS features + Fractionation
Classification accuracy	90	54.6	87.6
Precision	90.9	54.1	87.7
Error rate	9.9	45.4	12.4
Evaluation of prediction models			
Average accuracy	86.2	55.3	85.4
Precision	86.1	55.1	84.8
Error rate	13	42.5	13.3

8.2.6 Predicting different degrees of collagen coverage

Apart from the classification of *ex vivo* EGMs based on 15% collagen coverage, an investigation of the ability to efficiently predict the collagen coverage using different cut-offs was attempted. For this purpose, the same EGM data were used for the creation of new training and test datasets, but in this case within the range of 16-23% collagen coverage. The datasets consisting of EGMs derived from tissue slices with 16% or 17% collagen coverage were the same, so they were grouped. The same occurred for slices with 18% and 19% coverage. Therefore, 6 new training datasets and their corresponding test datasets were created, as presented in Table 8.6.

Feature selection was carried out for each training dataset separately. The selection process indicated 4-7 features, depending on the dataset, and features extracted from all types of EGM analysis were chosen in general (Appendix B-IV). The algorithms that were preferred during feature selection were the Ensemble Decision Trees and SVM, because these are the ones also applied during classification training.

Table 8.6: This table presents the number of tissue slices used per class (above or below each % collagen coverage), the number of EGMs which were recorded and analysed in each case and the size of training and test datasets.

	N° of slices with <X% (# EGMs)	N° of slices with >X% (# EGMs)	# of EGMs	
			Training dataset	Test dataset
15%	13 (170)	13 (170)	340	160
16/17%	14 (169)	12 (167)	336	164
18/19%	15(157)	11 (157)	314	186
20%	16 (154)	10 (154)	308	192
21%	19 (124)	7 (124)	248	252
22%	21 (95)	5(95)	190	310
23%	23 (71)	3 (71)	142	358

Based on the classification performance results for each percentage of collagen coverage, as also presented in Figure 8.7 and Table 8.7, it can be suggested that the classification based on 21% gave the best results. In this case the average accuracy was 84.5% with an error rate 14.5%. However, the performance of these datasets was worse than the one previously presented which was based on 15% of collagen coverage. The classification accuracy for the 16-23% collagen coverage datasets ranged between 75.6%-84.5%, while it was 90% for the 15% collagen coverage. The precision of the classification training also ranged between 75.7%-84.6%, but it was 90.9% for the classification training based on 15% coverage.

The relatively low predictability of collagen coverage using this classification was confirmed by the evaluation of each prediction model. The average accuracy ranged between 66.1% - 81.8% and the error rate ranged between 14.3% - 41.3% (Table 8.7). The performance of the model based on 15% collagen coverage was still better than the rest, but more data were included in the training dataset, also meaning more data points per class. For this reason, the rest of training datasets could be considered as weaker, as there are not as enough data to per class.

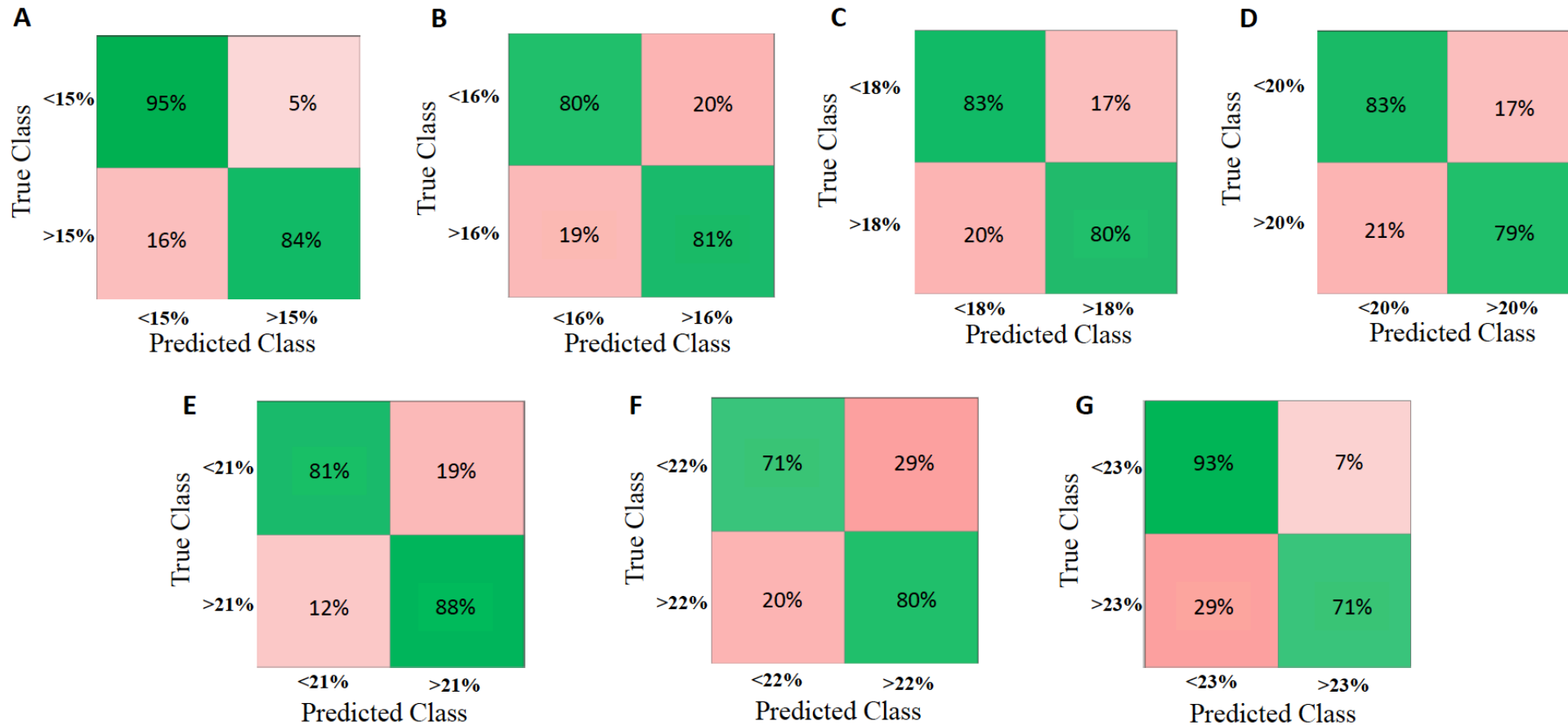


Figure 8.7: Panel of confusion matrices presenting the classification training performance for each dataset. EGMs were classified based on a certain percentage of the slice area covered by collagen and that ranged between 15-23%. Each confusion matrix was determined by the known and predicted groups. Green cells correspond to correct predictions.

Table 8.7: (Top) Assessment of classification training using datasets based on a range of collagen coverage percentages, in order to investigate the impact of different classification systems on the same dataset. Statistical indices were either calculated separately for each class or as an average for the overall performance. The classification method used in each case was the optimum for each dataset. (Bottom) The prediction model extracted in each case was evaluated using unknown data.

(%)	<16%	>16%	<18%	>18%	<20%	>20%	<21%	>21%	<22%	>22%	<23%	>23%
Se	79.8	80.7	83.2	80.4	82.7	78.8	81.1	87.8	71.1	80	92.8	71.4
Sp	80.7	79.8	80.4	83.2	78.8	82.7	87.8	81.1	80	71.1	71.4	92.8
PPV	80.5	80	80.9	82.6	79.6	82	86.9	82.3	78	73.5	76.5	90.9
NPV	80	80.5	82.6	80.9	82	79.6	82.3	86.9	73.5	78	90.9	76.5
Precision	80.3		81.8		80.8		84.6		75.7		83.6	
Error rate	19.7		18.2		19.2		15.5		24.4		17.9	
Classification accuracy	80.3		81.8		80.7		84.5		75.6		82.1	

(%)	<16%	>16%	<18%	>18%	<20%	>20%	<21%	>21%	<22%	>22%	<23%	>23%
Se	77.5	83.3	83.3	77.8	85.3	88.5	71.9	77.8	58.4	63.6	82.6	28.6
Sp	83.3	77.5	77.8	83.3	88.5	85.3	77.8	71.9	63.6	58.4	28.6	82.6
PPV	95.2	46.3	95.6	44.7	97.9	48.9	97.6	17.7	97.7	5.4	98.2	3.3
NPV	46.3	95.2	44.7	95.6	48.9	97.9	17.7	97.6	5.4	97.7	3.3	98.2
Precision	70.8		70.1		73.4		57.7		51.6		50.8	
Error rate	21.4		17.5		14.3		27.7		41.3		18.5	
Average accuracy	80.3		81.8		80.8		77.7		69.4		66.1	

8.2.7 Feature selection using *in vivo* EGMs

In order to predict the presence or not of scar regions, unipolar EGMs obtained from the endocardial LA surface of 8 patients with clinical history of persistent AF were used. The analysed data were 300ms recordings obtained from the LA posterior wall. In total, 452 EGMs could be analysed for their morphology (55 EGMs per patient on average), rejecting the signals which were characterized by high levels of noise. EGMs were labelled as derived from scarred/non-scarred areas, based on the categorization system presented in section 2.4. Based on LGE-MRI data obtained from the LA of these patients, followed by a correlation between EGMs and the LGE-MRI data, scar and non-scar regions were determined. The non-scar regions were defined by <0 S.D. above blood pool mean and the scar regions were defined by >2 S.D. above blood pool mean. The features selected by SFS using the training dataset (379 EGMs) were different depending on the classification algorithm used each time and they ranged between 2-5 features (Appendix B-V). Dominant frequency and frequency at signal of maximum energy were selected both when using the Decision Trees and kNN methods. The frequency of minimum or maximum energy was always picked up by SFS.

8.2.8 Classification training using *in vivo* EGMs

The 20 supervised machine learning algorithms provided through MATLAB were investigated, in order to detect the most suitable one for the training dataset. The classification training by applying the Bagging method gave 99.7% classification accuracy (Table 8.8). This means that all EGMs belonging to a class were correctly classified. Since such a high accuracy was achieved, it was not necessary to perform an optimization. The Ensemble Bagging method was used for the extraction of a prediction model.

Table 8.8. Classification performance presenting the accuracy after applying 20 classification training methods which are available on MATLAB. A 10-fold cross-validation was applied in each case. The average classification accuracy is presented here which is used as an index of the performance on the overall training dataset. The highest accuracy was achieved by the Bagging ensemble method.

Prediction method	Accuracy (%)	Prediction method	Accuracy (%)
Decision Tree (complex)	97.4	kNN (fine)	97.6
Decision Tree (medium)	97.4	kNN (medium)	92.9
Decision Tree (simple)	97.4	kNN (coarse)	90.8
Linear Discriminant Analysis	93.1	kNN (cosine)	91.3
Quadratic Discriminant Analysis	90	kNN (cubic)	92.3
SVM (Linear)	93.4	kNN (weighted)	95.8
SVM (Quadratic)	95.5	Ensemble (boosted)	73.4
SVM (Cubic)	97.9	Ensemble (bagging)	99.7
SVM (Fine Gaussian)	95.5	Ensemble (subspace discriminant)	92.3
Ensemble (subspace kNN)	98.2	Ensemble (RUSBoosted Trees)	97.9

A 10-fold cross-validation classification training was performed which is comparable to the cross-validation fold used for the *in vitro* and *ex vivo* data classification. Frequency of maximum energy and S-peak width were the features describing the training dataset. As also shown in Figure 8.8A-C, there were no misclassified EGMs, obtained from non-scar areas, during training and all classification performance indices were >99% with insignificant error rate (Table 8.9).

In addition, the use of voltage and fractionation as predictors was investigated. Applying the same training conditions, the data were characterized by features in three different ways: (a) the SFS features and voltage, (b) a combination of the SFS features with voltage and fractionation, and (c) the voltage and fractionation only. The evaluation of the classification performance showed that the training was efficient in overall, but still less efficient than using only the SFS features. The accuracy was 99.6% when the SFS features were used alongside voltage, 99.2% when voltage and fractionation were combined with the SFS features and 98.2% when using the voltage and fractionation only (Figure 8.10).

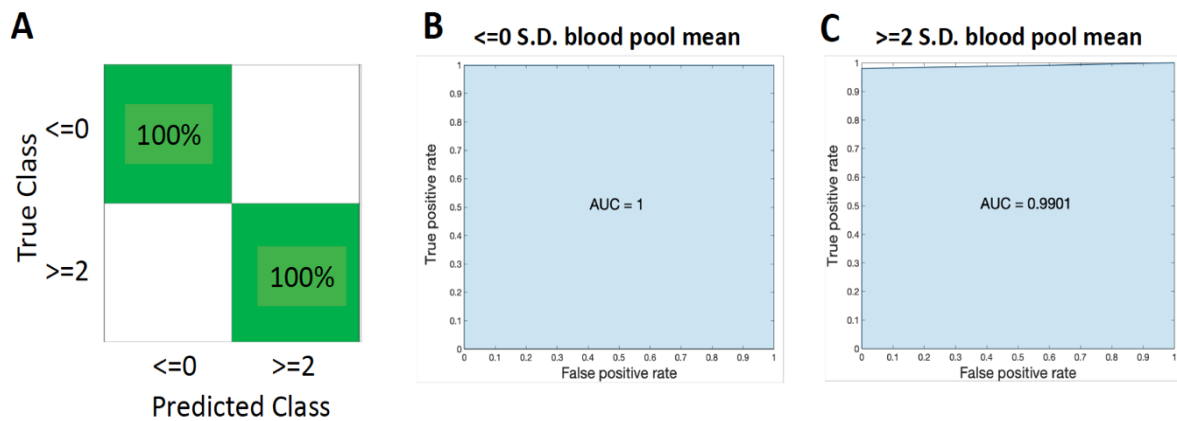


Figure 8.8: (A) Assessment of classification training using the Ensemble Bagging method by presenting the results on a confusion matrix. This matrix is determined by the known and predicted groups as developed during classification training. It is presented here that all EGMs recorded from non-scar regions were correctly predicted, as there are only green cells on the top row of the confusion matrix. (B-C) ROC curves obtained after classification training. Each curve corresponds to a class, the presence (B) or the absence of a scar (C).

Table 8.9: Assessment of the 10-fold cross-validation classification training. Calculation of indices for the overall and per class performance.

(%)	≤ 0 S.D.	≥ 2 S.D.
Se	100	99
Sp	99	100
PPV	99.6	100
NPV	100	99.6
Precision	99.8	
Error rate	0.3	
Classification accuracy	99.7	

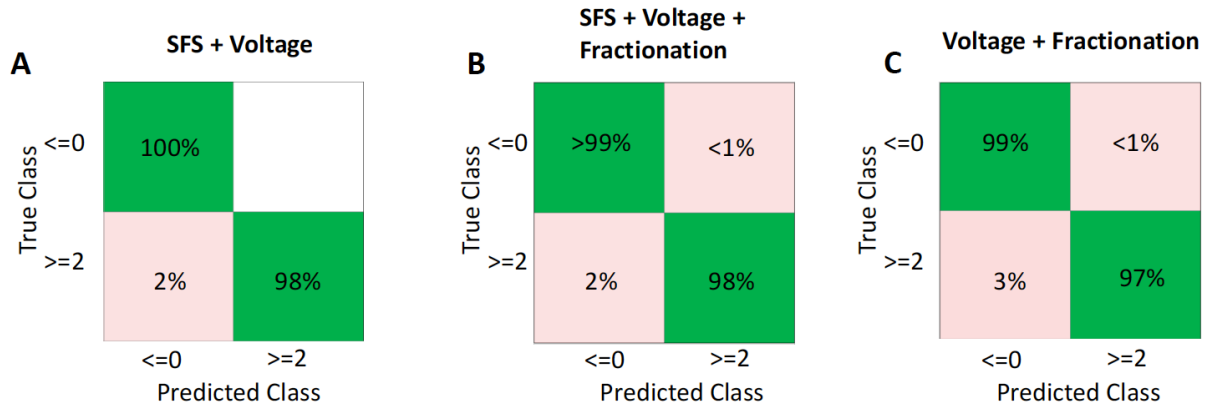


Figure 8.9: Confusion matrices presenting the classification performance when the predictors were the SFS features in addition to voltage (A), the SFS features alongside voltage and fractionation (B) or the voltage and fractionation in isolation (C).

Table 8.10: Assessment of the classification training using different EGM features as predictors. Classification indices computed for the overall and per class performance.

	SFS + voltage		SFS + voltage + fractionation		Voltage + fractionation	
	<=0 S.D.	>=2 S.D.	<=0 S.D.	>=2 S.D.	<=0 S.D.	>=2 S.D.
(%)						
Se	100	99	99.3	99	98.6	97
Sp	99	100	99	99.3	97	98.6
PPV	99.6	100	99.6	98	98.9	96.1
NPV	100	99.6	98	99.6	96.1	98.9
Precision	99.8		98.8		97.5	
Error rate	0.3		0.8		1.8	
Classification accuracy	99.6		99.2		98.2	

8.2.9 Validation of predictions based on *in vivo* EGM data

The prediction model, that was produced based on clinical data, was evaluated using a test dataset of 102 EGMs. These data were obtained under the same clinical conditions as the ones that were used for the training dataset. The test dataset contained different EGMs to those belonging to the training dataset. It was observed through the performance assessment that an average accuracy of 85.1% was achieved with 64.2% precision and 13.4% error rate (Table 8.11). Due to the binary classification, the rest of performance indices were similar between the two classes. Thus, the non-scarred areas could be predicted with 87.9% sensitivity and 82.5% specificity, while the scarred regions could be predicted

with 82.5% sensitivity and 87.9% specificity. These results are significantly lower than the classification training performance.

Table 8.11: Evaluation performance of the model developed for predicting the presence of scar on the substrate where the EGMs were collected from.

(%)	≤ 0 S.D.	≥ 2 S.D.
Se	87.9	82.5
Sp	82.5	87.9
PPV	78	83.3
NPV	83.3	78
Precision	64.2	
Error rate	13.4	
Average accuracy	85.1	

8.3 Discussion

In this chapter, a new method of establishing fibrosis measurement in *ex vivo* and *in vivo* models is proposed. This method is explicitly useful for classifying EGMs obtained *ex vivo* and *in vivo* from substrates with cardiac disease-related structural conditions. A variety of supervised learning methods were investigated and the most appropriate one, based on its performance, was further optimized and used for the extraction of a prediction model. Feature subsets were used in each case. Voltage, but not fractionation, was included in the feature subset for the classification of EGMs obtained from tissue slices. Despite the fact that these two features have been used in the clinic as markers for ablating locations supposedly responsible for maintaining AF, the work presented here showed that they are weak predictors. The clinical data used in this chapter also showed that a combination of features extracted from different types of analysis achieved better performance.

8.3.1 *Ex vivo* structural determinants as a limitation to EGM signal recording

The cardiac tissue has high cellular density which makes it difficult to image. The penetration of current imaging methods is relatively weak when compared to the tissue thickness and it is about 20-30 μ m and 50-80 μ m for confocal and two-photon microscopy respectively (412). Based on the excessive

disposition of collagen type I, which is a characteristic of cardiac diseases with interstitial fibrosis (413,414), SHG microscopy was used in this thesis for the visualization of fibrosis in cardiac tissue slices. This technique has been used for the visualization of structural alterations on cardiac tissue (329,412). Despite the fact that the collagen disposition on cardiac tissue has been visualized not only on 2D using classical histological techniques (355,415), but also on 3D using SHG microscopy (329), there are no suggested systems for the categorization of the myocardial tissue based on the amount of collagen.

In this chapter, the system that was suggested is based on the size of the area covered by collagen type I, which is measured during the post-processing of the microscopy images, and as described in detail in Chapter 2. It was found that no EGM recordings could be obtained from adult human left ventricle tissue slices with more than 25% collagen coverage. This may be explained by the presence of a threshold above which the distance between myocytes is too long to sustain the normal action potential propagation, due to the increased amount of intermingled fibroblasts (4). In addition, fibrosis could be so extended that not even the “zig-zag” propagation (4) could take place.

8.3.2 Are voltage and fractionation useful predictors of the structural fibrotic substrate?

Voltage, but not fractionation, was included in the feature subset that was selected to describe the training dataset during classification training. A classification accuracy of 90% was achieved when using the Bagging method and excluding fractionation from the feature subset, in order to classify data derived from tissue slices with less or more than 15% collagen coverage. In addition, it was observed that better classification accuracy is achieved when fractionation is excluded from the feature subset, as the accuracy drops in that case to 87.6%. Voltage and fractionation cannot be the sole features used to discriminate the two classes of the training dataset, as the classification training is not efficient and the accuracy is as low as 54.6%. Furthermore, both voltage and fractionation were not selected as informative features of the *in vivo* EGMs during classification training and they were excluded from the final feature subset. The classification accuracy was not dramatically reduced in that case, as it happened with the *ex vivo* data. However, it was 98.2% when voltage and fractionation were the only used features, compared to 99.7% when the feature subset indicated by the SFS process was used for classification training.

These findings suggest that low voltage and high fractionation cannot be necessarily associated with fibrosis, despite the common notion for the opposite (416). Low amplitude EGMs are commonly used as markers of areas of fibrosis, which produce a reduced electric field and consequently a lower potential difference between the measuring electrode and the reference point. Voltage was similar between scar and non-scar areas, as recorded *in vivo*, and the absence of significant difference between classes could explain why this feature was not selected to the final feature subset. It has been suggested that the decrease in the unipolar voltage in diseased myocardium may be a manifestation of less depolarized myocytes and more interstitial fibrosis (417), which contradicts our findings.

In contrast to voltage, fractionation was significantly higher in scar regions, as presented in Chapter 6. Despite this fact, fractionation was not chosen for classification, but it has been reported that different levels of fractionation can be used for distinguishing different incidents of chronic AF and for identifying critical sites that could be targeted for AF ablation. A number of studies implemented supervised learning methods for this purpose. Faes *et al.* (418) created a 3-level classification of fractionation using a similarity index. Schilling *et al.* (213) used decision trees to classify fractionated intracardiac bipolar EGMs and obtained a mean correct rate of 80.65%, which is lower than what was presented in this chapter using unipolar data collected from the tissue surface. Possible explanations for the better performance that was presented here can be the signal morphology with unipolar EGMs having more prominent features compared to the bipolar ones, as well as the combination of features used to characterize the signal. Moreover, Duque *et al.* (391) used a k-Nearest Neighbours method to classify 4 levels of fractionated EGMs, correlated with the electrophysiological activity during AF, such as the presence of rotor in persistent AF. However, these studies were mainly focused on fractionation, measured through different types of analysis, while the work presented in this chapter suggests the use of features computed by different types of analysis, such as time-frequency and frequency-domain. This combination seemed to be more informative compared to the mainstream time-domain characteristics.

8.3.3 Important features for identifying fibrotic tissue

Current commercial mapping systems used in the clinic, such as the EnSite NavX system, are solely programmed for the extraction of time-domain contact EGM features (419). Thus, the identification of EGM morphology changes is limited to simple characteristics, such as the presence or not of fractionation (420). This type of information would be more helpful in understanding the nature of the underlying disease (58,80). For this reason, there is a strong interest in using the EGM recordings in the diagnosis and treatment of pathophysiological abnormalities of the cardiac tissue (154,157). However, apart from a relatively small number of studies focused either on the classification of clinical EGM data

according to different states of AF (157) or on the prediction of the effects of pharmacological agents on hiPSCs (389), our knowledge about the information that we could obtain from EGM morphology and how this is correlated to all types of cardiac arrhythmia is still in a preliminary level.

This chapter, as well as chapter 6 which supports it by presenting specific feature changes due to the differences in slice structure, shows that features used for classification of *ex vivo* EGMs based on 15% collagen coverage included R-peak width, EGM amplitude, logarithmic energy entropy, RR interval, maximum modulus and mutual information. The importance of R-peak width can be justified by the fact that R-peak is part of the QRS complex which can be prolonged in a scar area due to the conduction slowing that occurs in that area and the extended time that it takes for the action potential to pass across an electrode (4). The next feature was EGM amplitude which was found to be lower in tissue slices with >15% collagen coverage. It has been suggested that voltage is lower in highly fibrotic areas, both in the atria through CFAE (6) and the ventricles, especially in cases of ventricular tachycardia and dilated cardiomyopathy (421). Logarithmic energy entropy is another way to evaluate the distribution of a dataset and compute the information uncertainty of the signal. Entropy measures have been extensively used for the automatic analysis of the electroencephalogram, because decreased complexity has been shown to correlate with anaesthesia effect (422). In addition, Shannon entropy has been used to differentiate CFAE from non-CFAE signals (127), as defined by Nademane *et al.* (37) and it is a feature increasingly being used for the identification of CFAE areas in the atria (125). Logarithmic energy entropy has been used as a possible feature to distinguish different types of AF using *in vivo* data (153), but the findings presented in this thesis suggest that it could be also used as a marker of highly fibrotic areas in tissue slices.

RR interval is a feature that has been adopted for the detection of AF in ECG signals (423), because RR intervals during AF have a larger standard deviation and a shorter correlation length than those during sinus rhythm (424). However, in the work presented in this chapter RR intervals were quantified using sinus rhythm EGM signals from ventricular tissue slices. As presented in Chapter 6, there is variability in RR interval between data groups, which could be justified by the differences in collagen coverage, and subsequently the degree of fibrosis, between slices. Maximum modulus is a time-frequency characteristic and its biological meaning is not clarified, as there is poor experience in correlating this with any determinants of EGM morphology. Its use has been mostly limited to the identification of maxima and minima across the different scales of a wavelet transform. The maximum modulus lines, which are normally measured, enable the identification of the onset of signal activation through the relationship between the wavelet transform and the derivative of the signal (16). Mutual information, an information theory-based feature (425) that has been used for the identification of EEG signal abnormalities (426), has been poorly studied on ECG and EGM signals. Even though mutual

information scoring did not show any changes among datasets with different collagen coverage in Chapter 6, this feature was selected by SFS as informative enough for classification. This fact is enhanced by the dramatic decrease of classification accuracy when mutual information is excluded from the training process, as shown in this chapter. As is the case with all information theory features used for EGM morphology analysis, more investigations need to be conducted in order to clarify how these measures correlate to the functional and structural determinants of the signal.

The predictors of the *in vivo* EGM morphology were two features derived from the time-domain and time-frequency EGM morphology analysis. In the same way as with tissue slices, a combination of EGM features is a promising way to achieve better classification performance. One of the features was the S-peak width. This feature is significant for the classification for the same reason previously explained for R-peak width regarding its prolongation due to conduction slowing in fibrotic-rich regions (4). The frequency of the maximum signal energy was the second selected feature. As shown in Chapter 6 of this thesis, this feature is correlated to the amount of fibrosis on the substrate. It was found to be reduced in highly fibrotic substrate both in tissue slices and in patients, even though those data were collected from different areas of the heart, the LV and LA respectively. It is a poorly investigated feature and thus, no further conclusions can be made.

However, it is questioning why only two features were selected for the classification of *in vivo* EGMs given the complexity of these data and the presence of noise associated with the recording environment. In addition, the classification training accuracy was unexpectedly high even when voltage and fractionation were only used to describe the data. The machine learning results using the tissue slices showed that these two features are not sufficient for predicting structural abnormalities, as discussed earlier in this section. This contradiction could be explained by the small training dataset, and more specifically by the limited number of EGMs belonging to each class. Even though there was an attempt for equal number of EGMs in each class, the small dataset was sensitive to overfitting. As already discussed in chapter 7, overfitting may lead to high classification accuracy, but the prediction model is unable to generalise to future unknown data. Therefore, the validation of the resulting prediction model showed significantly lower performance due to this fact. It can be then suggested that further optimisation of the model for predicting the scarred and non-scarred regions using clinical EGMs is needed by using a larger training dataset.

8.4 Limitations

The types of fibrosis, such as interstitial, patchy or scar, have been defined using traditional histological techniques, but they have not been verified using the SHG microscopy. Therefore, it was difficult carrying out this type of analysis in this thesis. Moreover, the anisotropy is another structural determinant of EGM morphology, which was difficult to be precisely controlled in the experiments.

It is still not entirely clear whether LGE-MRI, which was used in this case for the quantification of S.D. above blood pool mean, truly reflects myocardial fibrosis. There has been a study correlating collagen content with structural remodelling on LGE-CMRI using 14 biopsy samples from 10 patients with a history of AF (131), but the majority of studies have been conducted in animal models (427). It has been suggested that LA areas with >3 S.D. above blood pool mean are scar areas (174), but the availability of data from these areas was limited and a training dataset for use in classification training was not able to be created. Thus, it was proposed that scar areas are those with >2 S.D. above blood pool mean for this thesis. Therefore, these issues have to be addressed through further investigations.

8.5 Conclusions

Feature subsets extracted from EGMs, which were obtained from left ventricle slices, can be used for accurate predictions of the amount of fibrosis on the myocardial tissue. Voltage, but not fractionation, was included in the feature subset for the classification that showed the best performance using the *ex vivo* EGMs. The extracted prediction model based on biologically meaningful EGM features showed high predictability. The analysis of *in vivo* data though showed that neither voltage nor fractionation was informative for efficient classification of the data based on the presence of scar regions. Even though these two features are linked to CFAE morphology and are used as markers for ablating locations responsible for maintaining AF, they do not seem to be informative enough. A different combination of features is suggested to be necessary for accurate predictions of structural abnormalities, but these results need to be reconsidered given the limited number of *in vivo* data used for the training dataset.

Chapter 9

Predicting heart failure phenotypes from *ex vivo* electrograms using supervised machine learning

9.1 Introduction

Apart from investigating the ability of performing supervised machine learning for predicting functional and structural abnormalities related to cardiac arrhythmias and diseases, as presented so far, I sought to study whether it is possible to predict heart failure phenotypes using EGM morphology. This Chapter therefore presents a systematic characterisation of all feature changes that occur in each case, followed by the predictive modelling process for developing a model that can efficiently predict these phenotypes.

9.1.1 Heart failure

Heart failure is generally considered as the cardiac syndrome in which the heart cannot adequately provide blood and oxygen to the organs, due to reduced pumping action (428). Heart failure is a common syndrome affecting about 2% of the adult population worldwide. Studies have shown that its prevalence is age-dependent, as less than 2% of people younger than 60 years and more than 10% of those older than 75 years are affected by chronic heart failure (429). Heart failure is the final pathway of multiple pathologies, including myocardial infarction, hypertrophic cardiomyopathy, dilated cardiomyopathy and arrhythmogenic right ventricular cardiomyopathy (428). Heart failure patients usually have a history of hypertension, cardiomyopathies, coronary artery disease, valve disease or a combination of these (430). Common symptoms of heart failure are breathlessness, ankle swelling and fatigue, and its signs are peripheral oedema, pulmonary crackles and raised jugular venous pressure. These are caused by structural and functional cardiac abnormalities leading to elevated intra-cardiac pressures or reduced cardiac output at rest or during stress (429). Heart failure cases are most commonly classified as acute and chronic heart failure or as heart failure with reduced left ventricular ejection fraction and heart failure with preserved left ventricular ejection fraction (429).

The pathophysiology of heart failure with reduced ejection fraction involves a progressive condition starting from risk factors (i.e. hypercholesterolaemia, hypertension, diabetes, obesity, familial history of heart failure) leading to cardiac injury and the development of myocardial dysfunction. In the end, worsening symptoms are observed in patients until end-stage heart failure is diagnosed. Cardiac injury consists of the loss of myocytes and eccentric hypertrophy of remaining myocytes, leading to fibrosis, progressive left ventricular dilatation and a change in the shape of left ventricle from elliptical to spherical (429,431). These changes are called left ventricular remodelling. The result is increased myocardial oxygen consumption and reduced efficiency of myocardial contraction (431).

However, the pathophysiology of heart failure with preserved ejection fraction is not completely clear and incompletely studied. It is hypothesised that symptoms, such as chronic obstructive pulmonary disease (COPD), obesity, hypertension and diabetes cause a pro-inflammatory state followed by the production of reactive oxygen species by coronary microvascular endothelial cells and reduced nitric oxide (429,432). This condition leads to myocyte hypertrophy, increased collagen deposition and titin hypophosphorylation (429,433).

9.1.2 Ischaemic Heart Disease

One of the well-studied fundamental causes of heart failure, which is also taken into account during diagnosis, is ischaemic heart disease (IHD). IHD, also called coronary heart disease, is responsible for about 70% of heart failure patients (434). Clinical presentation of IHD is significantly different between women and men. Sex differences are extended to the symptoms recognised by patients and physicians, the outcome and the response to treatment (435). Unfortunately, the determinants of such differences are unclear and there is insufficient knowledge about the biological mechanisms (435).

In general, there are a number of mechanisms through which IHD leads to heart failure. Ischaemic episodes can be responsible for episodes of pulmonary oedema in patients who have well preserved left ventricular systolic function. Ischaemia-induced abnormalities of diastolic dysfunction may persist even when ischaemia-induced systolic dysfunction had resolved (434). Patients with heart failure can be more prone to IHD without symptoms, also known as silent ischaemia. These cases are associated with defective autonomic function, and especially myocardial denervation reducing sensory inputs caused by previous myocardial infarction episodes (434). Myocardial ischaemia and infarction are also responsible for arrhythmias. AF, affecting about 20% of heart failure patients, and ventricular arrhythmias are observed in the majority of arrhythmia cases (434).

Myocardial infarction is another important factor taking part in the progression of ventricular dysfunction. Myocardial infarction is usually followed by ventricular remodelling, which occurs very rapidly in the immediate post-infarction period and then more slowly thereafter (436). During that process an expanding myocardial scar is formed, in addition to this the non-infarcted myocardium adapts in terms of size and shape to the infarcted tissue. Myocardial infarction can also have chronic haemodynamic effects including the induction of mitral regurgitation (434).

9.1.3 Dilated cardiomyopathy

Dilated cardiomyopathy (DCM) is another heart failure phenotype studied in this Chapter. It is suggested that DCM is an infrequent cause of heart failure (434,437), even though there is high variability in DCM-induced heart failure incidents among clinical trials (434). Patients aged less than 50 years are a minority of those with heart failure, but this group is more prone to DCM (434). Non-ischaemic DCM is associated with significant morbidity and mortality. The implantation of an implantable cardioverter-defibrillator (ICD) has been found to reduce the risk for cardiac death in non-ischaemic DCM patients, even though it is an expensive solution. Although fibrosis is implicated after infarction, the significance of fibrosis in DCM is unclear (438). Midwall fibrosis is correlated to approximately 30% of DCM patients and it has been suggested that it can be used for accurate prognosis in patients with DCM (438). Both interstitial and replacement patterns of fibrosis are seen in DCM and microvascular ischemia has been suggested as a leading factor (438). A number of defective genes implicated in familial DCM have also been found to code for cytoskeletal proteins setting up a chronic injury-repair scenario which leads to fibrosis (439). There is poor knowledge though regarding the outcome of identifying myocardial fibrosis *in vivo* in DCM patients (438).

9.1.4 Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy (HCM) is a common inherited cardiovascular disease and recognised as an important cause for heart failure, cardiac sudden death and AF (440). It is recognised as a genetic heart disease with a suggested prevalence of at least 1 in 500 in the general population. Even though, HCM occurs so often, it is much less frequently recognised in clinical practice and as a result many individuals remain undiagnosed (440).

Genetic studies show that dominant mutations in 11 or more genes are responsible for HCM. These genes encode thick and thin contractile myofilament protein components of the sarcomere or the adjacent Z-disc (440). However, 70% of genotyped HCM patients had mutations in two genes only, β -myosin heavy chain (MYH7) and myosin-binding protein C (MYBPC3) (440). There is phenotypic heterogeneity between and within families, due to the fact that sarcomeres cannot be the sole determinant of the HCM phenotype (441).

The HCM phenotype includes ventricular remodelling with increasing left ventricular wall thickness during various phases of life (440). The conventional clinical diagnosis of HCM is carried out by

imaging with two-dimensional echocardiography, MRI or both. Imaging data show an increase in left ventricular wall thickness (440). Other common findings are mitral valve systolic anterior motion or hyperdynamic left ventricle, but these are not obligatory for diagnosis (442). Absolute left ventricular wall thickness ranges from mild (13-15mm) to massive (>50mm) (440). Another observation proving the heterogeneity of this disease is the asymmetrical pattern of left ventricular hypertrophy that may be observed. Although about 50% of patients have a diffuse wall thickening, an important minority (10-20%) have segmental hypertrophy limited to small portions of the chamber (443). ECGs are abnormal at 75-90% of HCM patients and there is a strong correlation with T-wave inversion on ECG, due to the sarcomere mutations (444).

9.1.5 Lymphocytic myocarditis

Myocarditis commonly results from viral infections, i.e. coxsackievirus B, parvovirus B19, even though there are reported cases as a result of other pathogens, toxic or hypersensitivity drug reactions or sarcoidosis. Lymphocytic myocarditis is distinguished into fulminant and acute myocarditis. Fulminant lymphocytic myocarditis has a distinct onset with a viral prodrome within 2 weeks before the onset of symptoms and it has a good prognosis. On the contrary, acute lymphocytic myocarditis more does not have a distinct onset and hemodynamic compromise, and it often results in the need for cardiac transplantation or in death. Acute myocarditis is frequently diagnosed as non-ischemic dilated cardiomyopathy and its symptoms vary among patients (445). Cases of lymphocytic myocarditis leading to clinical evidence of heart failure are very rare and due to this reason there are no published clinical trials of therapy for heart failure in patients with myocarditis (446). It also remains unclear why patients lymphocytic myocarditis are more prevalent in fatal or nonfatal sudden arrhythmia (446).

9.1.6 Electrophysiological changes in heart failure

There is a variety of electrophysiological changes during heart failure, which depend on the underlying disease and they include ion channel remodelling, extracellular matrix remodelling, scar presence, the activation of the sympathetic nervous system, Ca²⁺ handling alterations, dilatation and stretch. There is not a standard set of electrophysiological changes. This fact, in combination with the existence of multiple arrhythmogenesis mechanisms, explain the complications on the efficiency of heart failure treatments and why it is difficult to find the most appropriate therapeutic approach (447,448).

The majority of heart failure animal models and heart failure patients have action potential prolongation, due to changes in repolarisation. In fact, T-wave alterations and QT prolongation are clinically detected (428). The reason for this may be myocardial strain, even though there is no direct clinical evidence for this. T-wave alternans, which is linked to sudden cardiac death, is another result of changes in repolarisation (449). It has been suggested that ventricular action potential prolongation, due to amplitude increase or enhancement of the Ca^{2+} -transient, compensates for the loss in contraction. The reduction of I_{to} and I_{Ks} currents, either in isolation or in combination with the late inward Na^+ current which is absent from the healthy myocardium, lead to AP prolongation (428). QT prolongation, which is associated with dispersed repolarisation, has been found in heart failure patients. It is responsible for EAD and the myocardium excised from heart failure patients has demonstrated EADs, but with non-sustained arrhythmia (450). Moreover, QT prolongation, when combined with a certain degree of conduction slowing, is responsible for the initiation of re-entries, as known from studies with heart failure patients (451,452).

The decreased expression of Cx43 is a feature of heart failure, but large reductions of Cx43 are needed in order to have conduction changes (453). These low levels are probably not reached in heart failure, but when they occur in combination to Na^+ current reduction, they could lead to conduction slowing (454). QT prolongation was observed *in vivo* in a rabbit model of moderate to severe heart failure, but this was not observed in cardiac myocytes isolated from the same model. It was also shown that Cx43 expression levels were reduced in transmural tissue samples, but the conduction velocity was not affected, suggesting the effect of the modified intercellular coupling on transmural conduction velocity to be compensated by increased cell dimensions facilitating the conduction (455).

As it has become obvious, a variety of heart failure phenotypes are associated with increased fibrosis, and therefore with increased amount of extracellular matrix. The increased extracellular matrix separating the myocardial fibres appears to obstruct propagation, and especially transversely to the fibre direction, by disrupting the intercellular coupling. This leads not only to anisotropy, but also arrhythmogenesis (428,456) and contributes to source-sink mismatch conditions and conduction block (457). The extracellular matrix is excess in fibrotic regions and fibrosis has been associated with EGM fractionation. It has been also suggested that the fibrotic muscle is responsible for a “zig-zag” action potential propagation along the electrically isolated myocardial fibres leading to activation delay, but without reduction in conduction velocity to be observed (35). Thus, it has been proposed the activation delay to be caused by the increased length of the activation path and not by functional abnormalities. Considering all heart failure conditions described so far, it becomes clear that it is crucial to develop more efficient prognostic and diagnostic tools, in order to improve heart failure management and treat

each heart failure patient at the early stages of the disease. The majority of heart failure phenotypes are linked to structural remodelling of the ventricles, which is one of the main factors for abnormal EGM morphology and linked to specific EGM modifications.

The EGM morphology derived from various heart failure phenotypes was investigated in this Chapter and it was hypothesised that each heart failure condition can be accurately predicted using unipolar extracellular EGM morphology. The hypothesis to be investigated was that tissue slices derived from end-stage heart failure phenotypes induce propagation discontinuities, leading to consequent EGM modifications, i.e. increased EGM fractionation and EGM duration, comparing to tissue slices derived from rejected donor heart transplants. The hypothesis that each heart failure phenotype is responsible for a distinct EGM morphology was also studied.

The experimental data used in this Chapter are the same as the *ex vivo* ones presented in Chapters 5 and 6, where they were used in order to answer different questions. The patient characteristics from whom the samples were derived are presented in Tables 2.1 and 2.2. EGM morphology analysis was carried out using the methods presented in Chapter 3. Chapter 2 includes the feature selection analysis and classification training algorithms which were applied on training and test datasets for the extraction of a predictive model.

9.2 Results

9.2.1 Quantification of collagen coverage

Quantitative comparisons were made between the heart failure phenotypes investigated in this thesis with the amount of scar measured as a percentage of the total slice area covered by collagen (Figure 9.1). The mean \pm SEM collagen coverage was $9.8 \pm 5.4\%$ ($n = 2$) for control biopsies, $13.96 \pm 2.7\%$ ($n = 8$) for IHD, $17.85 \pm 1.7\%$ ($n = 11$) for DCM, $19.86 \pm 2\%$ ($n = 2$) for HCM and $9.9 \pm 3.5\%$ ($n = 2$) for lymphocyte myocarditis (LM). One-way ANOVA showed that there were no significant differences among groups, possibly due to the low amount of slices for some categories ($p = 0.29$). By qualitative assessment, it can be observed that the highest collagen coverage was found in HCM samples, IHD and DCM had moderate amount of collagen coverage and LM slices had the lowest collagen coverage among HF phenotypes, which was similar to the control.

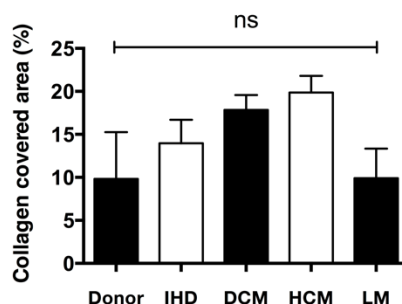


Figure 9.1: Collagen coverage of tissue slices derived from different heart failure phenotypes. Quantification of collagen coverage in each phenotype as a percentage of total size of slice. Donor (n=2), IHD (n=8), DCM (n=11), HCM (n=2), LM (n=2). Mean \pm SEM.

9.2.2 Measurement of conduction velocity

Conduction velocity was measured on the same direction as stimulation and it was found to be significantly different between groups (donor: 8.3 ± 2.3 cm/s, IHD: 11.57 ± 1.3 cm/s, DCM: 12.34 ± 1 cm/s, HCM: 21.09 ± 3.6 cm/s, LM: 7.3 ± 2.7 cm/s; $p = 0.006$) (Fig. 9.2A). The post-hoc analysis revealed that all phenotypes are significantly different from HCM, but there is no difference in conduction velocity among the rest (donor, IHD, DCM, LM). There was no correlation between conduction velocity and collagen coverage ($r^2 = 0.04$, slope = 0.13 ± 0.15 , $p = 0.4$) (Figure 9.2C).

The fibre orientation (longitudinal/transverse) on the side which was paced each time was unclear. However, high conduction velocity is known to normally take place alongside the longitudinal fibre orientation and the action potential propagation is slower on the transverse side of fibres (330). Therefore, the anisotropic ratio of the highest conduction velocity measured from one side of a slice by the lower conduction velocity measured from an orthogonal side was computed. (Fig. 9.2B). The results showed that the CV_{fast}/CV_{slow} ratio tended to be lower for the heart failure phenotypes rather than the donor data (Donor: 2.01 ± 0.05 ; IHD: 1.64 ± 0.3 ; DCM: 1.46 ± 0.4 ; HCM: 1.62 ± 0.3 ; LM: 1.46 ± 0.3 ; $p=0.32$), but these data were not statistically significant and there were no difference between heart failure phenotypes.

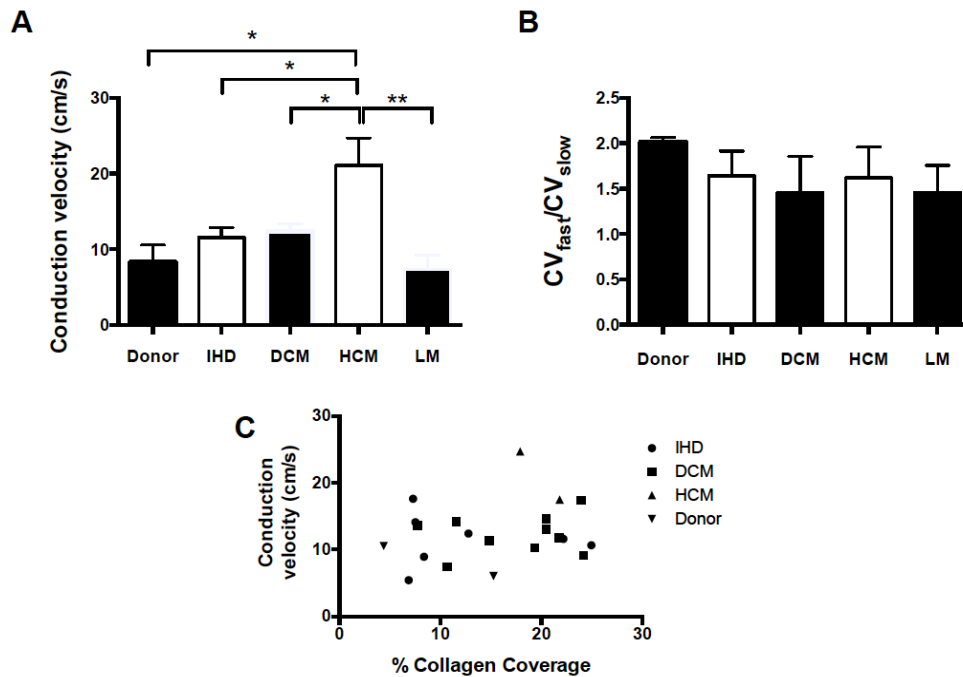


Figure 9.2: Average conduction velocity and conduction velocity anisotropy across tissue slices. (A) Conduction velocity measured as the average of the stimulation from two orthogonal directions in each slice – donor (n=2), IHD (n=8), DCM (n=11), HCM (n=2), LM (n=2). One-way ANOVA with Tukey’s post-hoc analysis. Mean \pm SEM, *p<0.05; **p<0.01. (B) One-way ANOVA did not show significantly different CV_{fast}/CV_{slow} across heart failure phenotypes, but this was lower than donor samples. (C) Regression analysis investigating the relationship between the average conduction velocity, as measured using the velocities from two perpendicular sides of the slice, and the collagen coverage. There was no correlation between these variables.

9.2.3 Heart failure phenotypes and EGM morphology

A total of 1440 unipolar extracellular EGMs were acquired across all tissue slices, which were derived either from a healthy heart (n = 2 slices) or from a failing heart (n = 22 slices). The data were obtained using the MEA system (section 2.3) from slices correlating to a healthy heart and four HF phenotypes: IHD, DCM, HCM, LM (Table 9.1). Automated feature extraction was then carried out on these recordings.

Table 9.1: Number of tissue slices used for EGM recordings and the total number of EGMs on which feature extraction was carried out.

phenotypes	Number of tissue slices	Number of analysed EGMs
Donor heart	2	37
IHD	8	155
DCM	11	189
HCM	2	61
Lymphocyte myocarditis	2	45

It was found that 23 out of 34 features, described in Chapter 3, changed. Compared to control EGMs, which were obtained from donor heart slices, 21 out of these 23 morphological descriptors were increased to at least one of the HF phenotypes (Figure 9.3 – 9.4). The frequencies of maximum and minimum energy extracted during the time-frequency analysis were the only reduced characteristics compared, to donor heart EGMs (frequency of max energy: donor, $29 \times 10^4 \pm 35 \times 10^3$; IHD, $15 \times 10^4 \pm 7.2 \times 10^3$; DCM, $10.3 \times 10^4 \pm 65 \times 10^3$; HCM, $13.5 \times 10^4 \pm 5.6 \times 10^3$; LM, $66 \times 10^4 \pm 3.8 \times 10^3$; $p < 0.0001$ / frequency of min energy: donor, $1.9 \times 10^0 \pm 2 \times 10^0$; IHD, $1 \times 10^0 \pm 3.6 \times 10^0$; DCM, $1.05 \times 10^0 \pm 3.9 \times 10^0$; HCM, $7.6 \times 10^0 \pm 1.5 \times 10^0$; LM, $8.7 \times 10^0 \pm 2.6 \times 10^0$; $p < 0.0001$) (Figure 9.4C-D).

Despite the increased average collagen coverage in heart failure phenotypes, the average EGM amplitude was lower in the donor slices (donor: $1029 \pm 109 \mu\text{V}$, IHD: $1797 \pm 131 \mu\text{V}$, DCM: $2234 \pm 175 \mu\text{V}$, HCM: $2341 \pm 304 \mu\text{V}$, LM: $2229 \pm 229 \mu\text{V}$; $p = 0.003$) (Figure 9.3B). However, there was no correlation between collagen coverage and voltage ($r^2 = 0.13$, $p = 0.46$). EGM duration was prolonged in heart failure samples (control: $32.5 \pm 1 \text{ms}$, IHD: $35.9 \pm 0.7 \text{ms}$, DCM: $38.33 \pm 9.7 \text{ms}$, HCM: $36 \pm 1.2 \text{ms}$, LM: $34.6 \pm 1.8 \text{ms}$; $p < 0.0001$) (Figure 9.3H) and fractionation index was increased (control: 2.65 ± 0.5 , IHD: 4.9 ± 0.3 , DCM: 3.9 ± 0.2 , HCM: 3.8 ± 0.3 , LM: 8.1 ± 0.4 ; $p < 0.0001$) (Figure 9.3 J).

It was observed that HF-derived signals were similarly modified in general, but there were a few exceptions which made the morphology obtained from each phenotype distinguishable. HCM-derived EGMs showed a significant increase of the mean frequency of PSD estimates (HCM: 108.1 ± 6.8 ; donor:

23.72±2; IHD: 39.78±1.7; DCM: 27.38±1; LM: 27.16±1.5; p<0.0001)) (Figure 9.4G) and median frequency of PSD estimates (HCM: 47.9± 4; donor: 3.6± 0.34; IHD: 10.54± 0.8; DCM: 8.9± 0.5; LM: 10.5± 0.9; p<0.0001) (Figure 9.4H). These changes were opposed to the rest of HF phenotypes which remained at the same levels with the donor heart samples. In addition, HCM and IHD signals showed increased S-peak-to-EGM offset gradient (HCM: 231.7±33; IHD: 308.5±38; donor: 97.3±19; DCM: 157.7±12; LM: 131.4±17; p=0.003) (Figure 9.3E). Some features detected in HCM-derived data though seemed to have remained unmodified compared to the control data, and while these features were modified in the rest of phenotypes. These features included Shannon entropy (HCM: 2.8e+023±6.3e+022; donor heart: 1.9e+023±3.4e+022; p=0.63) (Figure 9.3K), maximum modulus (HCM: 8.04e+009±1.1e+009; donor heart: 5.33e+009±4.8e+008; p=0.59) (Figure 9.4A), variance of energy (HCM: 5.7e+025±1.5e+025; donor heart: 2.4e+025±3.9e+024; p=0.08) (Figure 9.4B) and S.D. PSD estimates (HCM: 7.44e+013±1.7e+013; donor heart: 1.3e+014±2.2e+013; p=0.08) (Figure 9.4I). The same features were significantly increased for the rest of phenotypes.

Other specific subtype changes included the unmodified peak ratio in DCM recordings (DCM: 3±0.6; donor: 3.3±0.6; p=0.06) (Figure 9.3I), the much greater increase of fractionation index (LM: 8.1± 0.6; donor: 2.6± 0.5; IHD: 4.9± 0.3; DCM: 3.9± 0.2; HCM: 3.8± 0.4; p<0.0001) (Figure 9.3J) and the scale of maximum energy (LM: 0.33±0.01; donor: 0.13±0.02; IHD: 0.18±0.01; DCM: 0.25±0.01; HCM: 0.16±0.004; p<0.0001) (Figure 9.4E) for LM data, and the significant decrease of the frequency of maximum energy for the LM recordings (LM: 66,009±3,819; donor: 294,608±35,042; IHD: 149,823±7,283; DCM: 103,213±5,791; HCM: 134,887±5,609; p<0.0001) (Figure 9.4C).

Linear regression analysis showed no correlation between conduction velocity and EGM duration ($r^2 = 0.002$, slope = -0.04 ± 0.2 , $p = 0.84$).

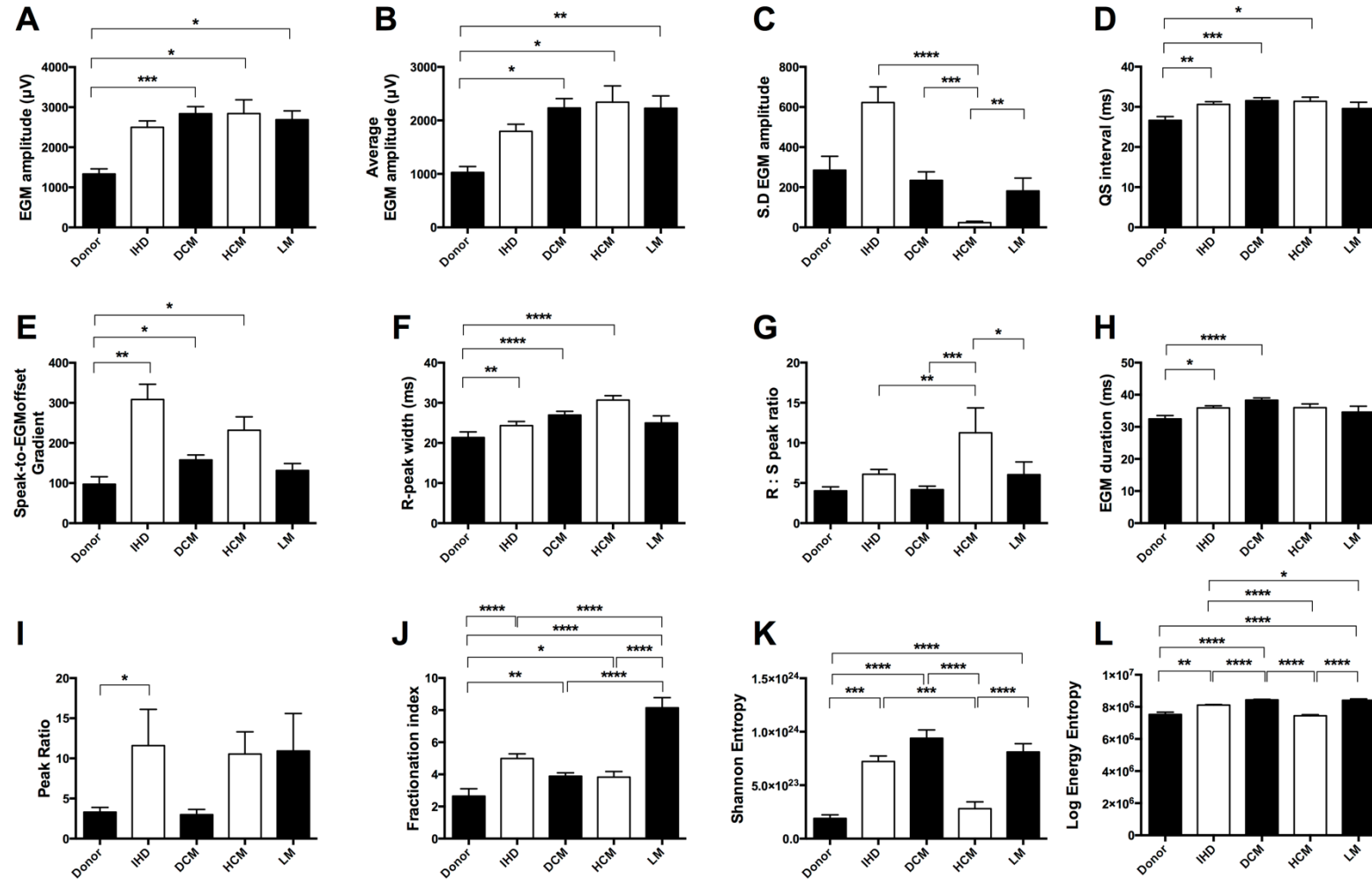


Figure 9.3: Time-domain features analysed using *ex vivo* data, which were affected by HF phenotypes. Compared to the EGM morphology recorded from healthy slices, all features were enhanced in at least one of the HF conditions. All bar charts represent mean \pm SEM. * $p \leq 0.5$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

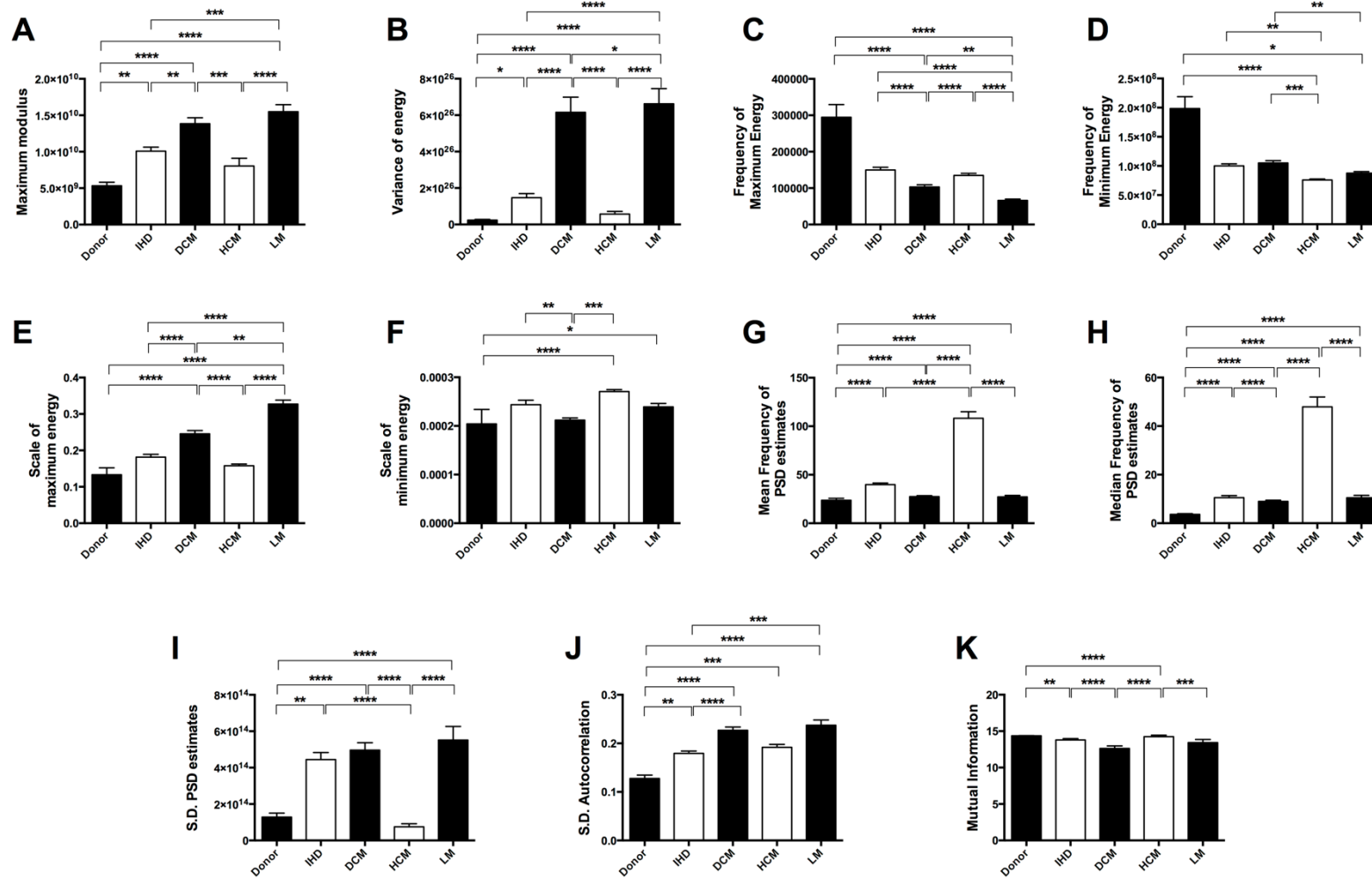


Figure 9.4: Morphological descriptors modified due to HF phenotypes. These features were derived from time-frequency and frequency-domain analysis and the quantification of information-theory based features. Apart from the frequency of maximum/minimum energy, the rest of features were significantly increased in all HF phenotypes. All bar charts represent mean \pm SEM. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

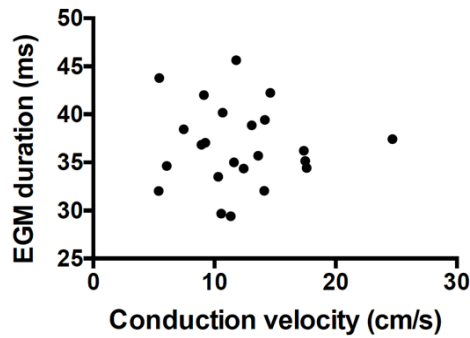


Figure 9.5: Linear regression analysis for the relationship between the conduction velocity, measured as the average conduction velocity from two perpendicular sides of a tissue slice, and EGM duration. There was not any correlation observed between the variables.

9.2.3 Feature selection

Even though, EGMs obtained from healthy and LM slices were assessed for changes in morphology alongside the rest of HF phenotypes, the data collected from these two conditions were not enough to be part of the classification process. Thus, the feature-based supervised machine learning process was applied using data obtained from IHD, DCM and HCM samples. After the rejection of noisy recordings, a dataset of 405 EGMs was created consisting of data for the three HF phenotypes. EGMs were randomly selected from this database to generate the dataset used for training the classifier (153 EGMs in total) and for validating the computed model (252 EGMs in total). The SFS method was applied for automated feature selection using the Tree Bagger algorithm and 10-fold cross-validation. The feature selection indicated 4 features as the essential ones to be used for classification training. These features were the following: maximum modulus, mutual information, frequency of maximum/minimum energy.

All these features were found to be significantly different among classes. In specific, maximum modulus was higher for the heart failure phenotypes compared to the donor sample and it was even higher for DCM and LM (Figure 9.4A). Mutual information was different among groups and it was found to be significantly lower for DCM (Figure 9.4K). Also, the frequency of maximum energy was lower for the heart failure phenotypes compared to the donor heart-derived data with a variability among heart failure phenotypes (Figure 9.4C). A similar pattern of results was also found for the frequency of minimum energy (Figure 9.4D).

9.2.4 Training an accurate heart failure phenotype classification algorithm

The Bagging ensemble method appeared to have the best performance from classification training. It gave a classification accuracy of 89.8%, which was the highest among 20 methods that were tested using the classification application of MATLAB (2017b) (Table 9.2). In order to optimise the classification algorithm, the out-of-bag classification error was investigated using 500 decision trees (Figure 9.6A). Fifty decision trees were sufficient to achieve stable out-of-bag classification error. In addition, leaf size equal to 1 was responsible for the lowest out-of-bag error. After the application of the optimised classification conditions for the bagging ensemble method, the classification accuracy increased to 89.8%. The area under the curve (AUC) of ROC curves was 0.94 for IHD, 0.92 for DCM and 0.98 for HCM, indicating high prediction accuracy for each heart failure subtype (Figure 9.6C-E).

Table 9.2. Classification accuracy of prediction methods using the same training dataset. The classification training was carried out with 10-fold cross-validation.

Prediction method	Accuracy (%)	Prediction method	Accuracy (%)
Decision Tree (fine)	78.9	KNN (fine)	83.7
Decision Tree (medium)	78.9	KNN (medium)	78
Decision Tree (coarse)	82.1	KNN (coarse)	69.1
Linear Discriminant Analysis	75.6	KNN (cosine)	81.3
Quadratic Discriminant Analysis	75.6	KNN (cubic)	74.8
SVM (Linear)	81.3	KNN (weighted)	80.5
SVM (Quadratic)	81.3	Ensemble (boosted)	32.5
SVM (Cubic)	82.9	Ensemble (bagging)	89.8
SVM (Fine Gaussian)	55.3	Ensemble (subspace discriminant)	79.7
Ensemble (subspace KNN)	87	Ensemble (RUSBoosted Trees)	32.5

Table 9.3. Systematic analysis of classification performance with Bagging Decision Trees. Statistical indices were estimated either for the overall performance (average classification accuracy, average error rate, average precision) or separately for each one of the HF phenotype classes (specificity, sensitivity, PPV, NPV).

	IHD	DCM	HCM
Sensitivity (%)	82.9	78	95.1
Specificity (%)	91	92.4	92.9
Positive Predictive Value (PPV) (%)	82.9	84.2	88.6
Negative Predictive Value (NPV) (%)	91	89	97
Precision (%)	85.3		
Error rate (%)	10.2		
Classification accuracy (%)	89.8		

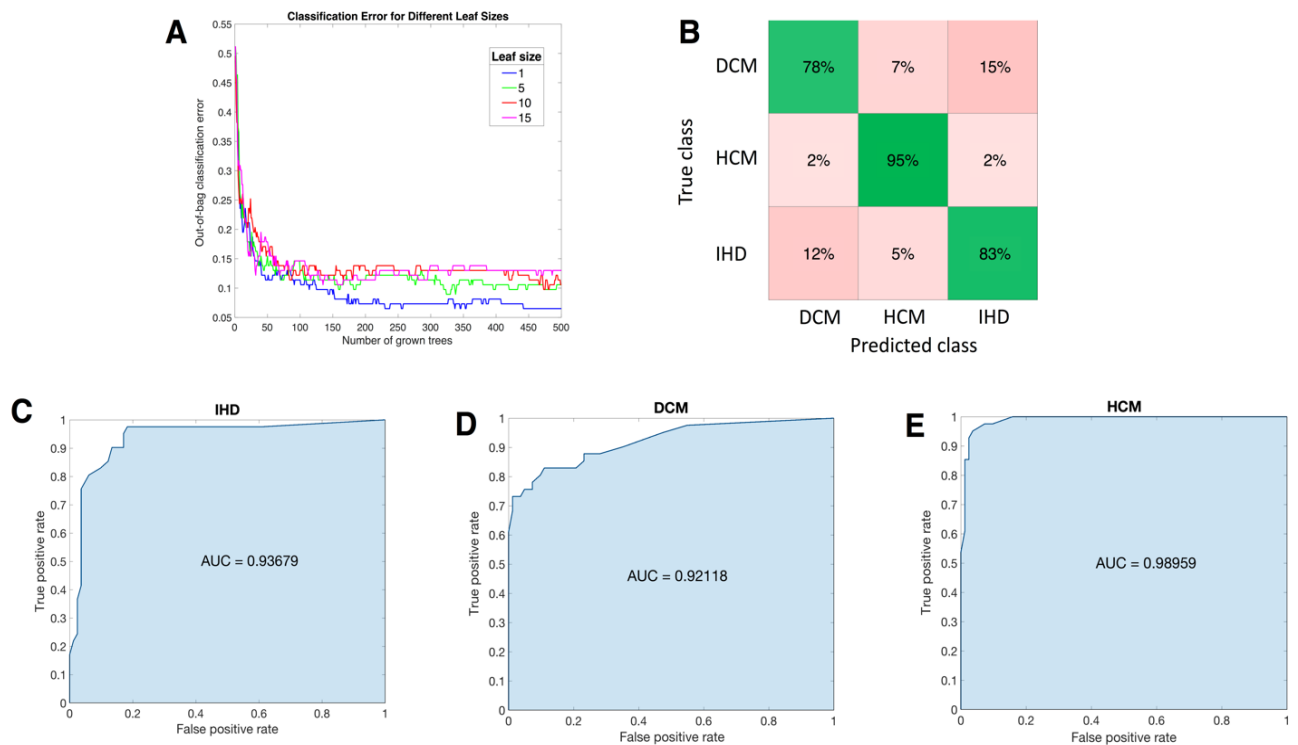


Figure 9.6: Classification training results. (A) Relative classification errors as a function of number of grown decision trees. (B) Confusion matrix presenting the classification accuracy of training dataset and comparing the predictability of classes. Diagonal cells (green) show the percentage of correctly classified EGMs. (C-E) ROC curves presenting the true positive rate versus false positive rate for each one of the classes when using the optimised bagging ensemble method. Based on ROC curves, the AUC value was also calculated for each one of them.

9.2.5 Trained algorithm reliably classifies EGMs according to heart failure phenotype

The prediction model based on the previously trained algorithm was used with a randomly selected validation dataset (n = 252 EGMs). The dataset was comprised of 53.8% IHD data, 41.04% DCM data and 5.16% HCM data. The model achieved a sensitivity between 78.1 – 88.2%, as that was measured separately for each class, a specificity of 89 – 93.6%, PPV of 86.9 – 94.3% and NPV of 74.5 – 89.2% (Table 9.4). These results are similar to the classification training performance.

Table 9.4. Evaluation performance of prediction model using a test dataset. Statistical indices (specificity, sensitivity, PPV, NPV) calculated for each class and average indices measured for the overall performance (precision, error rate).

(%)	IHD	DCM	HCM
Sensitivity	78.1	83.9	88.2
Specificity	91.3	89	93.6
PPV	92.7	86.9	94.3
NPV	74.5	86.4	89.2
Precision	89.8		
Error rate	14.9		
Accuracy	84.1		

9.3 Discussion

This chapter aimed to investigate how the EGM morphology obtained from a variety of end-stage heart failure samples is different compared to slices derived from donor transplants, resulting to a distinct EGM morphology. It was also investigated whether each heart failure condition is responsible for phenotype-specific EGM modifications. Even though conduction velocity was not significantly different between groups, there were significant differences in EGM morphology among groups for a number of features. The differences in EGM morphology helped the development of a prediction model which can successfully predict heart failure phenotypes from unipolar EGM morphology using supervised machine learning algorithms.

9.3.1 The effects of heart failure in EGM morphology

Our knowledge about the unipolar extracellular EGM morphology obtained from heart failure cases is very limited, as this is generally investigated in clinical practice when simultaneous arrhythmic conditions occur (440,458). In addition, ECGs are the main source of studying the electrophysiological activity in heart failure (444,458). The analyses of wavefront propagation and EGM morphology that were carried out in this chapter showed that each heart failure phenotype is linked to a distinct EGM morphology. In many cases, a number of time-domain, time-frequency and frequency-domain characteristics that form an EGM morphology are different among phenotypes. Since the majority of these 25 features are significantly different from features analysed in EGMs obtained from donor heart slices, then it could be suggested that the EGM features obtained from heart failure samples are modified and the subsequent overall EGM morphology is abnormal. An interesting finding was the fact that EGMs obtained from different heart failure subtypes may lead to similar changes, even though to a different extent, while other features were not modified or they changed to the opposite direction. This suggests that it is possible to distinguish heart failure subtypes between them, as well as from data obtained from rejected heart transplants.

One reason for the variable EGM morphology that was observed in this study may be ventricular remodelling, which often leads to heart failure. Ventricular remodelling is complex, entailing changes at many levels. At the cellular level, that involves modifications in the morphology, function and distribution of the cardiac myocyte. Changes in the type, structure and amount of collagen also occur (436). Other organ-level changes during ventricular remodelling may involve increased ventricular volume, hypertrophy, abnormal contractility and relaxation and enhanced ventricular compliance reflecting changes in collagen structure (436). However, the present *ex vivo* model excludes any parameters relevant to 3D architecture and thus, we are only able to make correlations between the amount of fibrosis in each heart failure phenotype and EGM morphology. The data analysis showed no correlation between collagen coverage on slices and conduction velocity or between conduction velocity and EGM duration or amplitude. This enhances our findings from Chapter 4 that such relationships are not clear in more complex experimental models.

Furthermore, this study showed that a number of time-frequency and frequency-domain features are significantly modified in heart failure. These changes are observed not only comparing to the donor heart samples, but they are different even among heart failure phenotypes. It is hard to translate these changes though in a biological setting. These features have been mostly used as ECG descriptors for heartbeat classification (459,460), an additional way for understanding AF (461) or as alternative features for increasing the efficiency of ablating procedures using intact EGMs (391). However,

considering the complexity of mechanisms that occur during heart failure or arrhythmic conditions, an investigation on a relatively simplistic model, such as the one presented here, can help towards the understanding of cellular-level mechanisms and the identification of specific EGM features could lead to hypothesis generation.

9.3.2 Classification of heart failure

This chapter presented the development of a model for efficiently predicting heart failure phenotypes using the EGM morphology. The Bagging ensemble algorithm was trained to discriminate IHD, HCM and DCM using 34 EGM features. The classification accuracy of the final model was 89.8%, and alongside the high classification performance, as described by the rest of indices, it allowed progress towards its evaluation using unlabelled EGMs. It was then confirmed that the prediction model was accurate, as it could correctly predict the heart failure phenotype with sensitivity 78.1 - 88.2% and specificity 89 - 93.6%. However, the small number of observations per class may be responsible for overfitting, and subsequently the relatively high classification accuracy, even though this is not clear due to the validation performance results.

There are multiple publications which present machine learning models for the classification of heart failure (462). The model presented by Austin *et al.* (463) utilised demographic characteristics, presenting signs and symptoms, laboratory data and previous medical history of patients to develop machine learning models based on multiple supervised machine learning algorithms, such as classification trees, random forests and boosted classification trees. These models aimed to predict and classify heart failure subtypes. As part of that study, it was shown that modern tree-based methods improved performance over conventional classification trees (463). Comparing the performance of all methods which are available within MATLAB, the same conclusion was addressed during this work. Shah *et al.* (464) used unsupervised machine learning algorithms for classifying heart failure phenotypes. They applied clustering algorithms on clinical, laboratory and ECG data for describing the patients.

However, there are not any published prediction models based on EGM morphology for classifying heart failure and the availability of an appropriate database may be a reason for this. This Chapter presents the first known attempt for application of supervised machine learning techniques on EGM morphology for the prediction of heart failure subtypes. Data were obtained from tissue slices, instead of heart failure patients, excluding some important parameters that affect EGM morphology. But the

demonstration of this application resulting to a precise prediction model shows that a similar approach using clinical data can be promising for the improvement of heart failure diagnosis and management.

9.4 Limitations

However, it is necessary to remember that the *ex vivo* model used in this study does not incorporate the entirety of the real clinical conditions under which a heart failure phenotype occurs, and therefore its limitations must be kept in mind. HCM is characterised by increased wall thickness (440), while extended scar areas and different types of fibrosis are quite common in DCM and IHD (434,438). These factors, and especially the 3D architecture, which can be crucial for EGM morphology, cannot be implemented easily in a lab setting and they were, of course, absent from our *ex vivo* model. In addition, infarcted hearts may have regional differences by distance from the infarct area (465). Therefore, it can be assumed that the samples used in this chapter do not take this into account, due to their small size.

Ventricular remodelling is complex and it may manifest itself in different ways in each heart failure phenotype. This fact combined with patients' clinical history, i.e. diabetes, hypertension prior to heart failure, which may be responsible for the way that ventricular remodelling is expressed out (434,436), increase the number of conditions that need to be correlated to a particular EGM morphology. Therefore, it is necessary to study a larger group of samples for each one of the heart failure phenotypes presented in this chapter. Moreover, functional abnormalities, such as dysfunctional ion channel activity, gap junction uncoupling, that might be present in our samples and affect EGM morphology were not studied in correlation to heart failure phenotypes.

9.5 Conclusion

Accurate classification of disease states, such as in heart failure, permits more accurate assessment of patient prognosis. Using a combination of EGM morphology features, derived from time-domain and time-frequency analysis, and applying supervised machine learning algorithms, a model for accurately predicting heart failure phenotypes was developed. Furthermore, the data analysis showed that these phenotypes have a distinct morphology from EGMs obtained from a donor heart.

Chapter 10

Conclusions

10.1 Summary of key findings

Understanding the way in which the cardiac electro-architecture affects the cellular and tissue electrophysiology, and subsequently the EGM morphology, is essential for improvements in diagnosis and personalised treatment of cardiac arrhythmias. One of the main contributions of the work presented in this thesis is the collection of a large number of unipolar EGMs under a variety of circumstances, in order to create unipolar EGM datasets, which not only are currently unavailable, but also were precious for the novel study of EGM morphology and the application of supervised machine learning. Another important contribution of this thesis is the ability to predict transmembrane currents, intracellular coupling, structural determinants of the action potential propagation and heart failure disease from the EGM. In order to come to this conclusion, contact EGMs were recorded and analysed from multicellular preparations, including cell monolayers, tissue slices and clinical data, under different modulations and control conditions. Supervised machine learning methods were then used to identify the relationship between the EGM morphology, previously recorded, and its functional and structural determinants.

The experimental process presented in Figure 2.1 was followed in this thesis. A range of 24 functional and structural modifications were applied on the *in vitro* model and their effects on cellular and tissue electrophysiology were validated and measured using a number of laboratory and clinical techniques. Collagen disposition and the presence of scar regions respectively were the only parameters researched in *ex vivo* and *in vivo* models. Optical mapping was used for obtaining action potential data from cell cultures and tissue slices. Structural characteristics were measured by immunofluorescence in cell cultures and two-photon microscopy (SHG technique) in tissue slices. Moreover, extracellular unipolar EGMs were recorded from both cell culture and tissue models using a MEA system, while paced unipolar recordings were collected from patients with a history of persistent AF using an electroanatomic mapping system. Algorithms were developed for the analysis of EGM morphology and its decomposition into a lower-dimensional feature-space representation was carried out to make it more amenable to machine learning techniques. The EGMs obtained from each experimental model were split into two datasets and feature selection algorithms were used to select the most informative subset of features. Twenty supervised machine learning algorithms were investigated, and the most effective method used to train a model

which discriminates the EGMs recorded from the different modulations in the experimental models. A number of predictive models were created which could be used for accurately predicting a range of structural and functional modulations on unlabelled data.

The added benefit of this work is:

- The production of a dataset consisting of unipolar EGMs, which are representative of a variety of experimental conditions related to heart diseases. Opposed to practical difficulties in obtaining clinical EGMs and given the lack of systematically characterised unipolar EGM datasets which can be used for the development of new algorithms and the in-depth investigation of EGM morphology, the data presented in this thesis can be considered valuable for future investigations. In addition, these data were obtained from different models with increasing complexity – cell cultures, tissue slices, patients – using controlled *in vitro* or *ex vivo* experimental conditions, in order to investigate how the increasing number of variables can affect the EGM morphology.
- The presentation of a novel method to analyse and quantify the structural heterogeneity of *ex vivo* tissue slices using SHG microscopy results. It is important being able to analyse *ex vivo* models because of their complexity, which is closer to the whole organ level, and the opportunity to delve in structural abnormalities related to heart diseases.
- Each functional and structural change to the electrophysiology at the cellular, tissue and whole organ level affects EGM morphology in a distinguishable manner. Apart from a number of overlapping EGM morphology characteristics that were identified across modulations on the same experimental model there were no changes that resulted in the same overall EGM morphology.
- Supervised machine learning can be used to accurately predict functional and structural determinants from unipolar EGM morphology.
- The breadth of features generated through time-domain, time-frequency and frequency-based analyses is important for generating accurate prediction models.
- Electrical recordings obtained at the single-cell level showed that gap junction uncoupling results to APD dispersion in cells located across an electrode with concurrent EGM fractionation.

The work presented in this thesis is also novel because of the following clinically relevant findings:

- Heart failure phenotypes, namely ischemic heart disease, dilated cardiomyopathy and hypertrophic cardiomyopathy, are responsible for distinct EGM morphologies which can be used inversely for predicting these phenotypes.
- Although scar regions can be predicted, voltage alone cannot be used as an *in vivo* marker of scar regions. The use of voltage and fractionation is inadequate for efficiently making assumptions about the fibrotic content of the substrate.

10.1.1 Electrogram morphology effects due to functional and structural changes

There is limited literature on the application of machine learning with unipolar EGMs and for this reason a number of signal analysis techniques were adapted from either the ECG morphology analysis (171,423) or other biomedical and engineering fields, such the analysis of EEG (426). However, the work presented in this thesis advanced the field by combining not only features that had been previously published or were clinically relevant, such as voltage and fractionation, but also a variety of other signal processing approaches, which were found critically important for successful discrimination of EGM morphology. Feature extraction was carried out by applying time-domain, time-frequency and frequency-domain analysis, and extracting features based on information theory. In total, 34 features could be computed and were used to represent EGMs obtained from different experimental models. The combination of features that was used in this thesis achieved accurate predictions during classification training and the successful validation of prediction models with test data. This agreed with previous studies that reported that the combination of different types of signal analysis leads to improved classification performance (213).

Pharmacological modulations were applied to cell culture, either individually or in combinations, in order to better understand how EGM morphology is affected by specific changes in ion channel and gap junction function, as well as the presence of fibrosis and scar. These experiments provided controlled information about the effects of reduced ion channel

and gap junction function on EGM morphology. Concurrent optical mapping, which was carried out for the majority of experimental conditions in cell culture, confirmed the effects of each pharmacological agent on EGM morphology by correlating AP characteristics with morphological EGM features. The reason for this is that AP morphology has been more thoroughly investigated following the blockade of ion currents, in comparison to EGMs, thereby providing validation of the specificity of action.

Na⁺ channel blockade was achieved by using lidocaine, the transient outward K⁺ current was blocked by 4-aminopyridine, nifedipine was used to target L-type Ca²⁺ channels, E-4031 and HMR-1556 were used to modify the rapidly and the slowly activating delayed rectifier K⁺ currents respectively, and pinacidil was used to open the K_{ATP} channels. A combination of lidocaine and nifedipine was also used for a double ion current blockade effect and carbenoxolone was applied on multicellular preparations for gap junction uncoupling. Even though there were a few similarities in the effects of some modulations on EGM morphology, the group of features which were significantly modified, and the degree of changes, were unique for each modulation. It was observed that the enhancement of the K⁺ current had opposing effects on some features compared to various types of K⁺ blockade, while there were cases that drugs with opposing effects altered features in the same way. The latter potentially indicated low specificity, especially for pinacidil, which had been found to reduce I_{to} in relatively low concentrations in canine ventricular myocytes (251). Moreover, the blockade of different types of K⁺ currents affected similar EGM features, but the degree of these effects were variable suggesting that feature extraction could be used for refined discrimination of ion currents.

Gap junction uncoupling affected all features computed from EGMs obtained from cell cultures. This led to conduction slowing and increased fractionation, which agree with the findings from humans (318,320) and isolated rabbit ventricular myocytes (314). However, increased amount of cultured fibroblasts in cell culture seemed to prohibit the increase of fractionation. The MEA and optical mapping concurrent experiments showed that AP morphology was also modified, but the absence of previous experience using NRVM and the contradictory findings with some animal models proposes the species- and age-specific effects of carbenoxolone. A number of features were also modified in tissue slices derived from a human DCM heart sample. FPD results agreed with the APD₉₀ change, but inverse and counter-

intuitive results were found for some of the rest features. These results could be explained by the presence of patchy and interstitial fibrosis on slices and the anti-arrhythmic effects of carbenoxolone in an area where the “source/sink mismatch” phenomenon is possible to occur. EGM features could be modified in that way, due to the protective response offered by gap junction uncoupling.

In addition to functional modulation, the effects of structural determinants on EGM morphology were investigated in this thesis. Even though different experimental models were used, it was observed in all cases, that the non-excitable regions of multicellular preparations which are responsible for discontinuous conduction – fibroblast rich, collagen-covered and scar regions – affect a variety of features and lead to conduction slowing. Feature modifications can be dependent on the amount of cellular scar in cell culture or the size of collagen covered area in tissue slices. There were a few exceptions to this though. For example, fractionation was significantly increased in +40%FB cell cultures, but remained stable and similar to the control in the other types of cell culture, which is in agreement with computational simulations that showed increase of fractionation in models with significantly high amount of fibrosis (466). The amount of fibroblasts in NC and +20%FB cell cultures seemed to be insufficient to create non-excitable regions that could lead to discontinuous conduction and the presence of a threshold effect of the fibroblast burden was proposed, in order to explain the differences between cell groups. A counter-intuitive *in vitro* finding was that increased amounts of the cellular component of scar increased the amplitude of the signal. Similar results were found in the tissue slice model. This is inverse to our knowledge from *in vivo* studies showing an inverse relationship between endocardial voltage of pacing or AF data and fibrosis (6,365). In order to investigate whether the increased signal voltage was not a result of enhanced NRVM/cultured fibroblast coupling, NRVMs were co-cultured with HeLa cells. Cx43 is not expressed in HeLa cells (366) and similar area of the cell culture was covered by this cell type as with fibroblasts. It was observed that the EGM amplitude was significantly decreased under these conditions suggesting that the absence of gap junctions *in vitro* does reduce the EGM amplitude and increases fractionation, which was similar to the +40%FB cell group.

A two-photon microscopy technique was used (372) in order to capture the area covered by collagen and subsequently measure the area covered by extracellular matrix in slices derived from heart failure human hearts and one rejected heart transplant. Features extracted from

different types of analysis were influenced by the amount of collagen, and FPD was among them. Dual modality experiments showed that the APD recorded from heart failure slices, and the concurrently recorded FPD, were similar to data previously recorded in ventricular slices (20) or wedges (349) and papillary muscles (375) in similar conditions. The FPD recorded in control samples was more abbreviated compared to heart failure samples which also agreed with clinical studies (349,376).

The analysis of pacing EGMs obtained from the LA of patients with a history of persistent AF was also carried out in this thesis. EGMs obtained from scar and non-scar regions, as defined according to a method of quantifying the S.D. above the mean blood pool intensity of LGE-MRI data (174), were compared. Even though it is known from clinical studies that low voltage is associated with scar areas (99,100) and the posterior LA is also related to low-voltage areas in patients with AF (101), it was found that EGM amplitude was not significantly reduced in scar regions. However, fractionation was increased and that was in agreement with previous studies about increased fractionation in LA regions coinciding with areas of slow conduction or pivot points where the activation wavefront turns around at the end of a conduction block (104). Apart from these, a number of other time-frequency features were also modified due to the presence of scar in posterior LA.

10.1.2 Supervised machine learning for predicting functional and structural modifications

Unipolar extracellular EGMs have been previously used to predict different AF phenotypes based on the presence of scar, but they ignored that electrical and structural remodelling may also take place (154,391). Therefore, one of the novelties of the work presented in this thesis was the study of a variety of isolated abnormalities that are associated with cardiac arrhythmias and heart failure diseases and the use of the decomposed EGM for training classification algorithms to predict those abnormalities. In general, the combination of features extracted from different types of analysis increased the efficiency of classification compared to the use of features extracted from only one type of analysis. Clinical studies conducted for simulation modelling (125) or human (396) and animal (397) models have come to the same conclusion about the use of unipolar EGM features in classification training.

The complexity of the classification problem was reduced by selecting a feature subset using an automated feature selection algorithm. It was difficult to make conclusions about the biological meaning of the subsets derived from *in vitro* data, due to the large number of classes included in each dataset and the inclusion of multiple features which were found to be affected by a number of abnormalities. However, the feature subset, which was obtained from *in vitro* data and was used in the end to predict functional modulations in isolation, did not overlap with the subset for predicting structural modifications or the combination of cell culture abnormalities. The Bagging algorithm, which is a supervised decision trees-based technique, was found to be the most effective learning method. However, the predictability of scarred regions in posterior LA should be conceived with scepticism, due to the low index values during the model validation step. A reason for this could be the small number of EGMs in the training dataset, which is responsible for overfitting and subsequently weak prediction models. The same problem may occur with the *ex vivo* data for predicting the different percentages of collagen coverage, but the validation results were closer to the training performance in that case. Using multiple folds of cross validation, as well as separate datasets for training and validation, is an approach followed in this thesis in order to eliminate the effects of overfitting.

10.2 Clinical relevance of findings

EGMs are routinely used in the clinic to examine the electrophysiological activity of the heart and diagnose any abnormalities. A number of ventricular or atrial arrhythmias are detected by observing the EGM morphology, as already discussed. However, EGM morphology is not adequately characterised not providing enough information about the content of the signal to aid diagnosis (2,3). Some simple morphological features, such as the QT interval, are used for the identification of a few types of arrhythmia. Voltage and fractionation are two other features that are clinically used. As part of this thesis, it has been presented that voltage and fractionation cannot be sufficient for discriminating *ex vivo* obtained EGMs according to the amount of collagen coverage, because the classification accuracy was only 54.6%. However, this increased to 87.6%, when the selected feature subset was added, and increased even more to 90% when fractionation was removed from the final feature subset. Despite the use of these two features as markers of scar, based on clinical findings (416), the findings of this thesis suggest that low voltage and high fractionation are not necessarily associated with high amount

of extracellular matrix of scar regions. Therefore, there needs to be a reconsideration of the features that could best describe the EGM signal obtained from scar regions.

10.3 Future directions

A number of mathematical models have been developed for the automated detection and segmentation of the T-wave based on trapezium's area approach (238) and pattern recognition techniques (239). Another recent study suggests an algorithm for the detection of T-wave based on a mathematical model given by a skewed Gaussian function (240). The advantage of these techniques is that they allow not only the detection of the T-wave peak, but also the location of T-wave end, giving the opportunity for the extraction of features, such as T-wave width, the slope of the descending part of T-wave, and scoring a T-wave according to the morphology of the peak, i.e. asymmetry, notch, flatness (241). These approaches could therefore be implemented in the future for the automated analysis of T-waves with accuracy and enriching that way the characterisation of an EGM.

An electrode picks up changes in electrical potential generated by the wavefront when it is still at a distance. That means that the electrode records the activity not only underneath the electrode, but also in the surrounding myocardium. This is the “field of view” of an electrode (4). The experimental models presented in this thesis – cell cultures and tissue slices – were two-dimensional. Therefore, a more realistic experimental model that can provide clinically-relevant data from intact myocardium is necessary to be used for studying how a variety of functional and structural modifications can affect EGM morphology. For this purpose, the isolated, retrograde-perfused Langendorff heart could be used. This experimental system is already used to extend our knowledge of ischemia-reperfusion injury, the cardiac effects of drugs and cell-based therapies for damaged myocardium. They are primarily used for small mammals, such as guinea pig, rat or rabbit (467).

However, more clinically-relevant data can be used from larger animal models, such as pigs, and humans. These are less frequently used models due to higher cost, greater biological variability, larger requirements in terms of perfusion solutions and associated equipment.

Moreover, data collection is more challenging when using large mammalian hearts. Due to all these complexities, clinically-relevant isolated porcine and human heart models are rarely used, but they are the next step towards cardiovascular translational research (467). However, protocols and adjustments to the Langendorff isolated heart perfusion apparatus (Figure 10.1A) and working heart model have been published, in order to facilitate electrophysiological experiments on large mammalian hearts (Figure 10.1B). The unipolar extracellular EGM data obtained from such models by applying clinical catheters on them could be compared with the *in vitro* and *ex vivo* data presented in this thesis. Furthermore, the algorithms for EGM morphology analysis and the supervised machine learning process presented in this thesis could be validated and be extended, in order to account for the additional complexity of the model. The validation results could indicate how the findings from this thesis are different in a complex experimental setting, how the local activation on the heart can be disturbed by remote activation disturbances in another part and how the cardiac tissue, which is anatomically and electrically anisotropic, influences the EGM morphology.

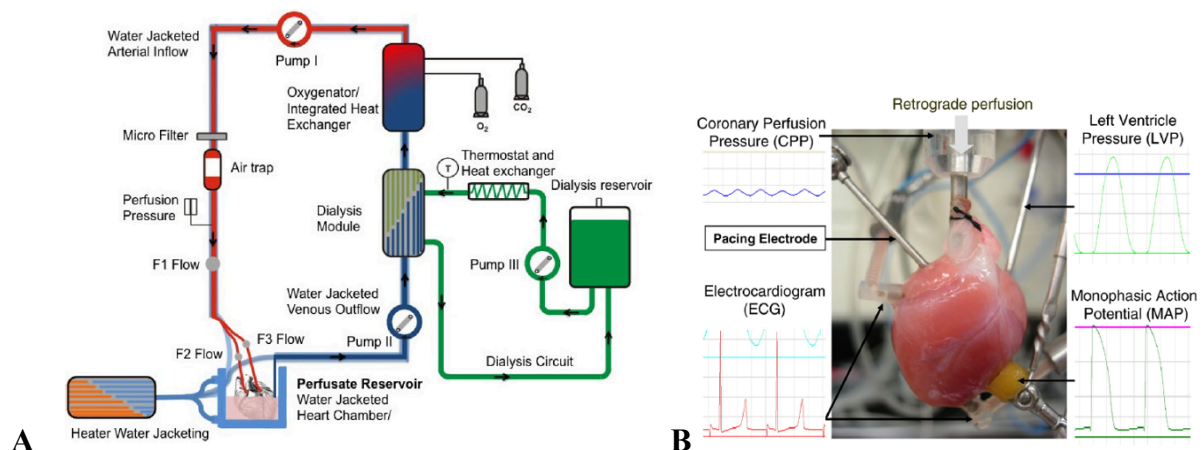


Figure 10.1: (A) Schematic representation of a Langendorff perfusion system. Adapted from Schuster et al. (468) and with permission of the rights holder, BioMed Central Ltd., under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>). (B) The posterior side of an isolated guinea pig heart mounted on a Langendorff apparatus with attached electrodes and probes on it for obtaining ECG, monophasic AP, left ventricle pressure and coronary perfusion pressure recordings continuously throughout perfusion. Image reproduced from Guo *et al.* (469) and with permission of the rights holder, Elsevier.

Apart from the sinus rhythm data obtained from the experimental models presented in this thesis and isolated, perfused Langendorff hearts, intracardiac unipolar EGMs obtained from

patients with functional abnormalities should be used for validating the algorithms presented in this thesis for EGM morphology analysis and the development of predictive models. Specifically, sinus rhythm data from patients with channelopathies, such as Brugada syndrome (BrS) and progressive cardiac conduction disease (PCCD), could be used. The formation of abnormal action potential, due to the variety of mechanisms contributing to arrhythmogenesis, is well characterised in such diseases (470). However, a few obvious morphological EGM changes that are characteristic for these diseases, such as the ST elevation and negative T-waves in BrS or the prolonged QT interval in Long QT syndrome, are only well known and used for diagnosis. PCCD is associated with more complex phenotypes and EGM morphology which overlap with BrS and Long QT syndrome (470). Therefore, the analysis of these data using the feature extraction algorithms presented in this thesis could indicate whether there are features uniquely modified on each cardiac disease distinguishing them from the rest.

10.4 Concluding remarks

In conclusion, the work presented in this thesis made systematic associations of a variety of factors related to types of cardiac arrhythmia and heart failure subtypes with specific changes of the unipolar extracellular EGM morphology. The electrical signal was recorded from NRVM cell cultures, tissue slices obtained from human failing hearts or rejected human heart transplants, and patients with persistent AF history. Optical mapping recordings were obtained concurrently to EGM recordings in cell monolayers and tissue slices, in order to confirm EGM modifications from AP correlates. This study, alongside a feature extraction methodology that was developed, enabled the characterisation of EGM changes in detail and the association of each functional and structural variable with specific feature changes. In addition, the development of prediction models was investigated. This enabled the accurate prediction of functional and structural variables or heart failure phenotypes, following a thorough analysis of optimising the classification process of the machine learning algorithms. Through the increasing complexity of the experimental models, it was revealed that supervised machine learning can be applied on EGMs obtained using different recording systems and under a variety of conditions. However, certain requirements must be met, i.e. biological and machine learning settings, in order to achieve high predictability. This can ensure the translational application of the process which was presented in this thesis for the improvement of the clinical diagnosis and treatment of cardiac diseases.

References

1. Plonsey R, Barr RC. *Bioelectricity: A quantitative approach*. 3rd ed. New York: Springer Science+Business Media, LLC; 2007. 531 p.
2. Kanagaratnam P, Kojodjojo P, Peters NS. Electrophysiological abnormalities occur prior to the development of clinical episodes of atrial fibrillation: Observations from human epicardial mapping. *PACE - Pacing Clin Electrophysiol*. 2008;31(4):443–53.
3. Schilling RJ, Davies DW, Peters NS. Characteristics of sinus rhythm electrograms at sites of ablation of ventricular tachycardia relative to all other sites: a noncontact mapping study of the entire left ventricle. *J Cardiovasc Electrophysiol*. 1998;9(9):921–33.
4. De Bakker JMT, Wittkampf FHM. The pathophysiologic basis of fractionated and complex electrograms and the impact of recording techniques on their detection and interpretation. Vol. 3, *Circulation: Arrhythmia and Electrophysiology*. 2010. p. 204–13.
5. Noorman M, Hakim S, Kessler E, Groeneweg J a., Cox MG PJ, Asimaki A, et al. Remodeling of the cardiac sodium channel, connexin43, and plakoglobin at the intercalated disk in patients with arrhythmogenic cardiomyopathy. *Hear Rhythm*. 2013;10(3):412–9.
6. Nademanee K, Lockwood E, Oketani N, Gidney B. Catheter ablation of atrial fibrillation guided by complex fractionated atrial electrogram mapping of atrial fibrillation substrate. *J Cardiol*. 2010;55.
7. Jadidi AS, Cochet H, Shah AJ, Kim SJ, Duncan E, Miyazaki S, et al. Inverse Relationship Between Fractionated Electrograms and Atrial Fibrosis in Persistent Atrial Fibrillation Combined Magnetic Resonance Imaging and High-Density Mapping. *J Am Coll Cardiol*. 62:802–12.
8. Alonso-Atienza F, Morgado E, Fernández-Martínez L, García-Alberola A, Rojo-Álvarez JL, Fernández-Martínez L, et al. Detection of life-threatening arrhythmias using feature selection and support vector machines. *IEEE Trans Biomed Eng*. 2014;61(3):832–40.
9. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920–30.
10. Kotsiantis SB. Supervised machine learning: A review of classification techniques. *Informatica*. 2007;31:249–68.
11. Nur F, Francis A, Abas FS, Besar R. ECG Classification using Wavelet Transform and Discriminant Analysis. In: *International Conference on Biomedical Engineering (ICoBE)*. 2012. p. 191–6.
12. Jekova I, Bortolan G, Christov I. Assessment and comparison of different methods for heartbeat classification. *Med Eng Phys*. 2008;30(2):248–57.
13. Nerbonne JM, Kass RS. Molecular Physiology of Cardiac Repolarization. *Physiol Rev*. 2005;85:1205–53.
14. Nanthakumar K, Jalife J, Masse S, Downar E, Pop M, Asta J, et al. Optical mapping of Langendorff-perfused human hearts : establishing a model for the study of ventricular fibrillation in humans. *Am J Physiol Hear Circ Physiol*. 2007;293:875–80.
15. Shenasa M, Hindricks G, Borggrefe M, Breithardt G, Josephson ME. *Cardiac Mapping*. 4th editio. Wiley-Blackwell; 2013. 5, 7 p.
16. Cantwell CD, Roney CH, Ng FS, Siggers JH, Sherwin SJ, Peters NS. Techniques for automated local activation time annotation and conduction velocity estimation in cardiac mapping. *Comput Biol Med*. 2015;65:229–42.
17. Spach MS, Kootsey JM. Relating the Sodium Current and Conductance to the Shape of Transmembrane and Extracellular Potentials by Simulation: Effects of Propagation Boundaries. *IEEE Trans Biomed Eng*. 1985;BME-32(10):743–55.
18. Coronel R, de Bakker JMT, Wilms-Schopman FJG, Opthof T, Linnenbank AC, Belterman CN, et al. Monophasic action potentials and activation recovery intervals as measures of ventricular action potential duration: Experimental evidence to resolve some controversies. *Hear Rhythm*. 2006;3(9):1043–50.
19. Vigmond EJ, Tsoi V, Yin Y, Pagé P, Vinet A. Estimating atrial action potential duration from electrograms. *IEEE Trans Biomed Eng*. 2009;56(5):1546–55.

20. Camelliti P, Al-Saud SA, Smolenski RT, Al-Ayoubi S, Bussek A, Wettwer E, et al. Adult human heart slices are a multicellular system suitable for electrophysiological and pharmacological studies. *J Mol Cell Cardiol.* 2011;51(3):390–8.
21. Roney CH. Mathematical techniques for assessing cardiac wavefront dynamics. 2015;
22. Biermann M, Shenasa M, Borggrete M, Hindricks G, Haverkamp W, Breithardt G. The interpretation of cardiac electrograms in “Cardiac Mapping.” 2nd editio. Elmsford, NY: Blackwell Publishing/Futura Division; 2003. 15–39 p.
23. Ndrepepa G, Caref EB, Yin H, El-Sherif N, Restivo M. Activation time determination by high-resolution unipolar and bipolar extracellular electrograms. 1995;6(3):174–88.
24. Levick RJ. An Introduction to Cardiovascular Physiology. CRC Press; 2009. 8 p.
25. Debbas NMG, Jackson SHD, De Jonghe D, Robert A, Camm AJ. Human atrial repolarization: Effects of sinus rate, pacing and drugs on the surface electrocardiogram. *J Am Coll Cardiol.* 1999;33(2):358–65.
26. Theres H, Stadler RW, Stylos LEE, Glos M, Leuthold T, Baumann G, et al. Comparison of Electrocardiogram and Intrathoracic Electrogram Signals for Detection of Ischemic ST Segment Changes During Normal Sinus and Ventricular Paced Rhythms. 2002;13(10):990–5.
27. Spach MS, Dolber PC. Relating extracellular potentials and their derivatives to anisotropic propagation at a microscopic level in human cardiac muscle: Evidence for electrical uncoupling of side-to-side fiber connections with increasing age. *Circ Res.* 1986;58(3):356–71.
28. Spach MS, Miller WT, Miller-jones E. Extracellular Potentials Related to Intracellular Action Potentials during Impulse Conduction in Anisotropic Canine Cardiac Muscle. 1978;188–205.
29. Spach MS, Miller WT, Dolber PC, Kootsey JM, Sommer JR, Mosher CE. The functional role of structural complexities in the propagation of depolarization in the atrium of the dog. Cardiac conduction disturbances due to discontinuities of effective axial resistivity. *Circ Res.* 1982;50(2):175–91.
30. Derksen R, Van Rijen HVM, Wilders R, Tasseron S, Hauer RNW, Rutten WLC, et al. Tissue discontinuities affect conduction velocity restitution: A mechanism by which structural barriers may promote wave break. *Circulation.* 2003;108(7):882–8.
31. Gaudesius G, Miragoli M, Thomas SP, Rohr S. Coupling of cardiac electrical activity over extended distances by fibroblasts of cardiac origin. *Circ Res.* 2003;93(5):421–8.
32. Rog-Zielinska EA, Norris RA, Kohl P, Markwald R. The living scar – cardiac fibroblasts and the injured heart. *Trends Mol Med.* 2016;In press.(2):99–114.
33. Janse MJ, Wit AL. Electrophysiological mechanisms of ventricular arrhythmias resulting from myocardial ischemia and infarction. *Physiol Rev.* 1989;69(4):1049–169.
34. Nguyen TP, Qu Z, Weiss JN. Cardiac fibrosis and arrhythmogenesis: The road to repair is paved with perils. *J Mol Cell Cardiol.* 2014;70:83–91.
35. Bakker JMT De, Capelle FJL Van, Janse MJ, de Bakker JM, van Capelle FJ, Janse MJ, et al. Slow Conduction in the Infarcted Human Heart ‘ Zigzag ’ Course of Activation. *Circulation.* 1993;88(3):915–26.
36. Soejima K, Stevenson WG, Maisel WH, Sapp JL, Epstein LM. Electrically unexcitable scar mapping based on pacing threshold for identification of the reentry circuit isthmus: Feasibility for guiding ventricular tachycardia ablation. *Circulation.* 2002;106(13):1678–83.
37. Nademanee K, McKenzie J, Kosar E, Schwab M, Sunsaneewitayakul B, Vasavakul T, et al. A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate. *J Am Coll Cardiol.* 2004;43(11):2044–53.
38. Kostin S, Klein G, Szalay Z, Hein S, Bauer EP, Schaper J. Structural correlate of atrial fibrillation in human patients. *Cardiovasc Res.* 2002;54(2):361–79.
39. Tanaka M, Fujiwara H, Onodera T, Wu DJ, Hamashima Y, Kawai C. Quantitative analysis of myocardial fibrosis in normals, hypertensive hearts, and hypertrophic cardiomyopathy. *Heart.* 1986;55(6):575–81.
40. Chowdhury RA, Tzortzis KN, Dupont E, Selvadurai S, Perbellini F, Cantwell CD, et al. Concurrent micro-to macro-cardiac electrophysiology in myocyte cultures and human heart slices. *Sci Rep.* 2018;8(1):6947.
41. Rohr S. Role of gap junctions in the propagation of the cardiac action potential. *Cardiovasc Res.*

- 2004;62(2):309–22.
42. Kléber AG, Rudy Y. Basic mechanisms of cardiac impulse propagation and associated arrhythmias. *Physiol Rev.* 2004;84(2):431–88.
 43. Severs NJ, Bruce AF, Dupont E, Rothery S. Remodelling of gap junctions and connexin expression in diseased myocardium. *Cardiovasc Res.* 2008;80(1):9–19.
 44. Severs NJ, Coppens SR, Dupont E, Yeh HI, Ko YS, Matsushita T. Gap junction alterations in human cardiac disease. *Cardiovasc Res.* 2004;62(2):368–77.
 45. Kanno S, Saffitz JE. The role of myocardial gap junctions in electrical conduction and arrhythmogenesis. *Cardiovasc Pathol.* 2001;10(4):169–77.
 46. Moreno AP. Biophysical properties of homomeric and heteromultimeric channels formed by cardiac connexins. *Cardiovasc Res.* 2004;62(2):276–86.
 47. Lesh MD, Pring M, Spear JF. Cellular uncoupling can unmask dispersion of action potential duration in ventricular myocardium. A computer modeling study. *Circ Res.* 1989;65(5):1426–40.
 48. De Groot JR, Coronel R. Acute ischemia-induced gap junctional uncoupling and arrhythmogenesis. *Cardiovasc Res.* 2004;62(2):323–34.
 49. Beardslee MA, Lerner DL, Tadros PN, Laing JG, Beyer EC, Yamada KA, et al. Dephosphorylation and intracellular redistribution of ventricular connexin43 during electrical uncoupling induced by ischemia. *Circ Res.* 2000;87(8):656–62.
 50. Peters NS, Coromilas J, Severs NJ, Wit AL. Disturbed connexin43 gap junction distribution correlates with the location of reentrant circuits in the epicardial border zone of healing canine infarcts that cause ventricular tachycardia. *Circulation.* 1997 Feb;95(4):988–996.
 51. Severs N. Gap junction alterations in the failing heart. *Eur Heart J.* 1994;15 Suppl D:53–7.
 52. Ng FS. *The Effects of Gap Junction Modulation on Myocardial Structure and Function.* Imperial College London; 2011.
 53. Miller D, Wang L, Zhong J. *Sodium Channels, Cardiac Arrhythmia, and Therapeutic Strategy.* 1st ed. Vol. 70, *Pharmacology & Therapeutics of Constitutively Active Receptors.* Elsevier Inc.; 2014. 367–392 p.
 54. Olson TM, Michels V V, Ballew JD, Reyna SP, Karst ML, Herron KJ, et al. Sodium Channel Mutations and Susceptibility to Heart Failure and Atrial Fibrillation. *JAMA J Am Med Assoc.* 2005;293(4):447–54.
 55. Perez-Cortes EJ, Islas AA, Arevalo JP, Mancilla C, Monjaraz E, Salinas-Stefanon EM. Modulation of the transient outward current (I_{to}) in rat cardiac myocytes and human Kv4.3 channels by mefloquine. *Toxicol Appl Pharmacol.* 2015;288(2):203–12.
 56. Sah R, Ramirez RJ, Oudit GY, Gidrewicz D, Trivieri MG, Zobel C, et al. Regulation of cardiac excitation-contraction coupling by action potential repolarization: Role of the transient outward potassium current (I_{to}). *J Physiol.* 2003;546(1):5–18.
 57. Benitah JP, Alvarez JL, Gómez AM. L-type Ca²⁺ current in ventricular cardiomyocytes. *J Mol Cell Cardiol.* 2010;48(1):26–36.
 58. Michael G, Xiao L, Qi XY, Dobrev D, Nattel S. Remodelling of cardiac repolarization: How homeostatic responses can lead to arrhythmogenesis. *Cardiovasc Res.* 2009;81(3):491–9.
 59. Li GR, Dong MQ. Pharmacology of Cardiac Potassium Channels. *Adv Pharmacol.* 2010;59(C):93–134.
 60. Heeringa J, Van Der Kuip DAM, Hofman A, Kors JA, Van Herpen G, Stricker BHC, et al. Prevalence, incidence and lifetime risk of atrial fibrillation: The Rotterdam study. *Eur Heart J.* 2006;27(8):949–53.
 61. Ezekowitz MD, Levine JH. Atrial Fibrillation. *Nurse Pract.* 1999;24(11):108.
 62. Mujovic N, Marinkovic M, Lenarczyk R, Titz R, Potpara TS. Catheter Ablation of Atrial Fibrillation : An Overview for Clinicians. *Adv Ther.* 2017;34:1897–917.
 63. Vaquero M, Calvo D, Jalife J. Cardiac fibrillation: From ion channels to rotors in the human heart. *Hear Rhythm.* 2008;5(6):872–9.
 64. Jalife J, Berenfeld O, Mansour M. Mother rotors and fibrillatory conduction: A mechanism of atrial fibrillation. *Cardiovasc Res.* 2002;54(2):204–16.
 65. Groot NMS de S De, Houben RPMM, Smeets JL, Boersma E, Schotten U, Schalij MJ, et al. Electropathological Substrate of Long-Standing Persistent Atrial Fibrillation in Patients With Structural

- Heart Disease Longitudinal Dissociation. *Circ Arrhythmia Electrophysiol.* 2010;3(6):606–15.
66. Davidenko JM, Kent PF, Chialvo DR, Michaels DC, Jalife J. Sustained vortex-like waves in normal isolated ventricular muscle. *Proc Natl Acad Sci.* 1990;87(22):8785–9.
 67. Narayan SM, Krummen DE, Rappel WJ. Clinical mapping approach to diagnose electrical rotors and focal impulse sources for human atrial fibrillation. *J Cardiovasc Electrophysiol.* 2012;23(5):447–54.
 68. Steinberg JS, Shah Y, Bhatt A. Focal impulse and rotor modulation : Acute procedural observations and extended clinical follow-up. *Hear Rhythm.* 2016;14(2):192–7.
 69. Haissaguerre M, Hocini M, Shah AJ, Derval N, Sacher F, Jais P, et al. Noninvasive panoramic mapping of human atrial fibrillation mechanisms: A feasibility report. *J Cardiovasc Electrophysiol.* 2013;24(6):711–7.
 70. Moe GK, Rheinboldt WC, Abildskov JA. A computer model of atrial fibrillation. *Am Heart J.* 1964;67(2):200–20.
 71. Pandit S V., Berenfeld O, Anumonwo JMB, Zaritski RM, Kneller J, Nattel S, et al. Ionic determinants of functional reentry in a 2-D model of human atrial cells during simulated chronic atrial fibrillation. *Biophys J.* 2005;88(6):3806–21.
 72. Sanchez C, Corrias A, Bueno-Orovio A, Davies M, Swinton J, Jacobson I, et al. The Na⁺/K⁺ pump is an important modulator of refractoriness and rotor dynamics in human atrial tissue. *AJP Hear Circ Physiol.* 2012;302(5):H1146–59.
 73. Pruvot EJ, Katra RP, Rosenbaum DS, Laurita KR. Role of Calcium Cycling Versus Restitution in the Mechanism of Repolarization Alternans. *Circ Res.* 2004;94(8):1083–90.
 74. Bayer JD, Narayan SM, Lalani GG, Trayanova NA. Rate-Dependent Action Potential Alternans in Human Heart Failure Implicate Abnormal Intracellular Calcium Handling. *Hear Rhythm.* 2010;7(8):1093–101.
 75. Wan X, Laurita KR, Pruvot EJ, Rosenbaum DS. Molecular correlates of repolarization alternans in cardiac myocytes. *J Mol Cell Cardiol.* 2005;39(3):419–28.
 76. Ausma J, Wijffels MC, Thone F, Wouters L, Allessie M, Borgers M. Structural Changes of Atrial Myocardium due to Sustained Atrial Fibrillation in the Goat. *Circulation.* 1997;96(9):3157–63.
 77. Frustaci A, Chimenti C, Bellocci F, Morgante E, Russo MA, Maseri A. Histological Substrate of Atrial Biopsies in Patients With Lone Atrial Fibrillation. *Circulation.* 1997;96(4):1180–4.
 78. Aldhoon B, Melenovský V, Peichl P, Kautzner J. New insights into mechanisms of atrial fibrillation. *Physiol Res.* 2010;59(1):1–12.
 79. Koduri H, Cokic I, Aistrup GL, Gordon D, Wasserstrom A, Kadish AH, et al. Contribution of fibrosis and the autonomic nervous system to atrial fibrillation electrograms in heart failure. Vol. 5, *Circulation: Arrhythmia and Electrophysiology.* 2012. 640–649 p.
 80. de Jong S, van Veen T a B, van Rijen HVM V, de Bakker JMT. Fibrosis and cardiac arrhythmias. *J Cardiovasc Pharmacol.* 2011;57(6):630–8.
 81. Ten Tusscher KHWJ, Panfilov A V. Reentry in heterogeneous cardiac tissue described by the Luo-Rudy ventricular action potential model. *Am J Physiol Circ Physiol.* 2003;284(2):H542–8.
 82. Engelman ZJ, Trew ML, Smaill BH. Structural heterogeneity alone is a sufficient substrate for dynamic instability and altered restitution. *Circ Arrhythmia Electrophysiol.* 2010;3(2):195–203.
 83. Zipes DP, Camm AJ, Borggrefe M, Buxton AE, Chaitman B, Fromer M, et al. ACC/AHA/ESC 2006 Guidelines for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death. Vol. 48, *Journal of the American College of Cardiology.* 2006. 247–346 p.
 84. Zheng Z, Croft JB, Giles WH, Mensah GA. Clinical Investigation and Reports Sudden Cardiac Death in the United States , 1989 to 1998. *Circ.* 2001;104.18:2158–63.
 85. Morgan JM, Cowan JC, Camm AJ, McComb JM. Sudden cardiac death: Opportunities for prevention. *Heart.* 2006;92(6):721–3.
 86. Huikuri H V, Castellanos A, Myerburg RJ. Sudden Death Due to Cardiac Arrhythmias. *N Engl J Med.* 2001 Nov 15;345(20):1473–82.
 87. Zipes DP, Wellens HJJ. Sudden Cardiac Death. *Circulation.* 1998;98:2334–51.
 88. Curtis MJ, Hancox JC, Farkas A, Wainwright CL, Stables CL, Saint DA, et al. The lambeth conventions (II): Guidelines for the study of animal and human ventricular and supraventricular arrhythmias.

- Pharmacol Ther. 2013;139(2):213–48.
89. Antzelevitch C. Basic mechanisms of reentrant arrhythmias. *Curr Opin Cardiol*. 2001;16(1):1–7.
 90. Shah M, Akar FG, Tomaselli GF. Molecular basis of arrhythmias. *Circulation*. 2005;112(16):2517–29.
 91. Wolf C, Berul C. Molecular Mechanisms of Inherited Arrhythmias. *Curr Genomics*. 2008;9(3):160–8.
 92. Alessie, MA, Bonke, FI, Schopman F. Circus Movement in Rabbit Atrial Muscle as a Mechanism of Tachycardia III . The " Leading Circle " Concept : A New Model of Circus Movement. *Circ Res*. 1977;41:9–18.
 93. Davidenko JM, Pertsov AM, Salomonsz R, Baxter WT, Jalife J. Stationary and drifting spiral waves of excitation in isolated cardiac muscle. *Nature*. 1992;355:349–51.
 94. January CT, Riddle JM. Early Afterdepolarizations: Mechanism of Induction and Block. *Circ Res*. 1989;64(5):977–91.
 95. Wit AL, Boyden PA. Triggered activity and atrial fibrillation. *Hear Rhythm*. 2007;4(3 SUPPL.):1–12.
 96. Boller M, Pott L. β -Adrenergic modulation of transient inward current in guinea-pig cardiac myocytes. *Pflugers Arch Eur J Physiol*. 1989;415:276–88.
 97. Brown HF, Difrancesco D, Noble SJ. How does adrenaline accelerate the heart? [13]. *Nature*. 1979;280(5719):235–6.
 98. Verma A, Wazni OM, Marrouche NF, Martin DO, Kilicaslan F, Minor S, et al. Pre-existent left atrial scarring in patients undergoing pulmonary vein antrum isolation: An independent predictor of procedural failure. *J Am Coll Cardiol*. 2005;45(2):285–92.
 99. TEH AW, KISTLER PM, LEE G, MEDI C, HECK PM, SPENCE SJ, et al. Electroanatomic Remodeling of the Left Atrium in Paroxysmal and Persistent Atrial Fibrillation Patients Without Structural Heart Disease. *J Cardiovasc Electrophysiol*. 2012;23(3):232–8.
 100. Fiala M, Wichterle D, Chovančik J, Bulková V, Wojnarová D, Nevřalová R, et al. Left atrial voltage during atrial fibrillation in paroxysmal and persistent atrial fibrillation patients. *PACE - Pacing Clin Electrophysiol*. 2010;33(5):541–8.
 101. Marcus GM, Yang Y, Varosy PD, Ordovas K, Tseng ZH, Badhwar N, et al. Regional left atrial voltage in patients with atrial fibrillation. *Hear Rhythm*. 2007;4(2):138–44.
 102. Anter E, Josephson ME. Bipolar voltage amplitude: What does it really mean? *Hear Rhythm*. 2015;108–9.
 103. Bailin SJ, Hoffman C, West R, Weers N. Substrate-Based Atrial Fibrillation Ablation Guided by Voltage Gradient Mapping: A New Approach for Atrial Fibrillation Ablation. *J Innov Card Rhythm Manag*. 2012;3(December):1034–41.
 104. Konings KTS, Kirchhof CJHJ, Smeets JRLM, Wellens HJJ, Penn OC, Allessie MA. High-density mapping of electrically induced atrial fibrillation in humans. *Circulation*. 1994;89(4):1665–80.
 105. Konings KT, Smeets JL, Penn OC, Wellens HJ, Allessie M a. Configuration of unipolar atrial electrograms during electrically induced atrial fibrillation in humans. *Circulation*. 1997;95(5):1231–41.
 106. Nademanee K, McKenzie J, Kosar E, Schwab M, Sunsaneewitayakul B, Vasavakul T, et al. A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate. *J Am Coll Cardiol*. 2004;43(11):2044–53.
 107. Oral H, Chugh A, Good E, Wimmer A, Dey S, Gadeela N, et al. Radiofrequency catheter ablation of chronic atrial fibrillation guided by complex electrograms. *Circulation*. 2007;115(20):2606–12.
 108. Porter M, Spear W, Akar JG, Helms R, Brysiewicz N, Santucci P, et al. Prospective study of atrial fibrillation termination during ablation guided by automated detection of fractionated electrograms. *J Cardiovasc Electrophysiol*. 2008;19(6):613–20.
 109. Verma A, Novak P, Macle L, Whaley B, Beardsall M, Wulffhart Z, et al. A prospective, multicenter evaluation of ablating complex fractionated electrograms (CFEs) during atrial fibrillation (AF) identified by an automated mapping algorithm: Acute effects on AF and efficacy as an adjuvant strategy. *Hear Rhythm*. 2008;5(2):198–205.
 110. Li WJ, Bai YY, Zhang HY, Tang RB, Miao CL, Sang CH, et al. Additional ablation of complex fractionated atrial electrograms after pulmonary vein isolation in patients with atrial fibrillation a meta-analysis. *Circ Arrhythmia Electrophysiol*. 2011;4(2):143–8.
 111. Narayan SM, Shivkumar K, Krummen DE, Miller JM, Rappel WJ. Panoramic electrophysiological

- mapping but not electrogram morphology identifies stable sources for human atrial fibrillation: Stable atrial fibrillation rotors and focal sources relate poorly to fractionated electrograms. *Circ Arrhythmia Electrophysiol.* 2013;6(1):58–67.
112. Viles-Gonzalez JF, Gomes JA, Miller MA, Dukkipati SR, Koruth JS, Eggert C, et al. Areas with complex fractionated atrial electrograms recorded after pulmonary vein isolation represent normal voltage and conduction velocity in sinus rhythm. *Europace.* 2013;15(3):339–46.
 113. Stiles MK, Brooks AG, John B, Shashidhar, Wilson L, Kuklik P, et al. The effect of electrogram duration on quantification of complex fractionated atrial electrograms and dominant frequency. *J Cardiovasc Electrophysiol.* 2008;19(3):252–8.
 114. Pantos I, Katritsis G, Zografos T, Camm AJ, Katritsis DG. Temporal stability of atrial electrogram fractionation in patients with paroxysmal atrial fibrillation. *Am J Cardiol.* 2013;111(6):863–8.
 115. Telgärsky R. Dominant Frequency Extraction. *arXiv.* 2013;1–12.
 116. Sanders P, Berenfeld O, Hocini M, Jaïs P, Vaidyanathan R, Hsu LF, et al. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. *Circulation.* 2005;112(6):789–97.
 117. Kumagai K, Sakamoto T, Nakamura K, Nishiuchi S, Hayano M, Hayashi T, et al. Combined dominant frequency and complex fractionated atrial electrogram ablation after circumferential pulmonary vein isolation of atrial fibrillation. *J Cardiovasc Electrophysiol.* 2013;24(9):975–83.
 118. Ng J, Goldberger JJ. Understanding and interpreting dominant frequency analysis of AF electrograms. *J Cardiovasc Electrophysiol.* 2007;18(6):680–5.
 119. Schuessler RB, Kay MW, Melby SJ, Branham BH, Boineau JP, Damiano RJ. Spatial and temporal stability of the dominant frequency of activation in human atrial fibrillation. *J Electrocardiol.* 2006;39(4 SUPPL.):7–12.
 120. Mandapati R, Skanes A, Chen J, Berenfeld O, Jalife J. Stable microreentrant sources as a mechanism of atrial fibrillation in the isolated sheep heart. *Circulation.* 2000;101(2):194–9.
 121. Atienza F, Almendral J, Moreno J, Vaidyanathan R, Talkachou A, Kalifa J, et al. Activation of inward rectifier potassium channels accelerates atrial fibrillation in humans: Evidence for a reentrant mechanism. *Circulation.* 2006;114(23):2434–42.
 122. Atienza F, Almendral J, Jalife J, Zlochiver S, Ploutz-Snyder R, Torrecilla EG, et al. Real-time dominant frequency mapping and ablation of dominant frequency sites in atrial fibrillation with left-to-right frequency gradients predicts long-term maintenance of sinus rhythm. *Hear Rhythm.* 2009;6(1):33–40.
 123. Verma A, Lakkireddy D, Wulffhart Z, Pillarisetti J, Farina D, Beardsall M, et al. Relationship between complex fractionated electrograms (CFE) and dominant frequency (DF) sites and prospective assessment of adding DF-guided ablation to pulmonary vein isolation in persistent atrial fibrillation (AF). *J Cardiovasc Electrophysiol.* 2011;22(12):1309–16.
 124. Katritsis DG, Pantos I, Efstathopoulos EP. Catheter ablation of atrial fibrillation guided by electrogram fractionation and dominant frequency analysis. *Expert Rev Cardiovasc Ther.* 2011;9(5):631–6.
 125. Ganesan AN, Kuklik P, Lau DH, Brooks AG, Baumert M, Lim WW, et al. Bipolar electrogram Shannon entropy at sites of rotational activation implications for ablation of atrial fibrillation. *Circ Arrhythmia Electrophysiol.* 2013;6(1):48–57.
 126. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948;27(July 1928):379–423.
 127. Ng J, Borodyanskiy AI, Chang ET, Villuendas R, Dibs S, Kadish AH, et al. Measuring the complexity of atrial fibrillation electrograms. *J Cardiovasc Electrophysiol.* 2010;21(6):649–55.
 128. Hutchinson MD, Gerstenfeld E, Desjardins B, Bala R, Riley MP, Garcia FC, et al. Endocardial Unipolar Voltage Mapping to Detect Epicardial VT Substrate in Patients with Nonischemic Left Ventricular Cardiomyopathy. *Circ Arrhythmia Electrophysiol.* 2011;4(1):49–55.
 129. Perin EC, Silva G V., Sarmiento-Leite R, Sousa ALS, Howell M, Muthupillai R, et al. Assessing myocardial viability and infarct transmuralty with left ventricular electromechanical mapping in patients with stable coronary artery disease: Validation by delayed-enhancement magnetic resonance imaging. *Circulation.* 2002;106(8):957–61.
 130. Daccarett M, Badger TJ, Akoum N, Burgon NS, Mahnkopf C, Vergara G, et al. Association of left atrial fibrosis detected by delayed-enhancement magnetic resonance imaging and the risk of stroke in patients with atrial fibrillation. *J Am Coll Cardiol.* 2011;57(7):831–8.

131. McGann C, Akoum N, Patel A, Kholmovski E, Revelo P, Damal K, et al. Atrial Fibrillation Ablation Outcome is Predicted by Left Atrial Remodelling on MRI. *Circ Arrhythmia Electrophysiol.* 2014;7(7):23–30.
132. Mitchell TM. *Machine Learning*. McGraw-Hill Science/Engineering/Math; 1997. 421 p.
133. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. *Science* (80-). 2015;349(6245):255–60.
134. Cantwell CD, Mohamied Y, Tzortzis KN, Garasto S, Houston C, Chowdhury RA, et al. Rethinking multiscale cardiac electrophysiology with machine learning and predictive modelling. *Comput Biol Med.* 2019;104:339–51.
135. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. Stanford, CA: Springer US; 2011. 109–116 p.
136. Yuan G, Ho C, Lin C. Recent Advances of Large-scale Linear Classification. *Proc IEEE.* 2012;100(9):2584–603.
137. Burges CJC. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min Knowl Discov.* 1998;2(2):121–67.
138. Raffenburgh RH. *Linear Discriminant Analysis*. Virginia Polytechnic Institute; 1957.
139. Breiman L. Bagging predictors. *Mach Learn.* 1996;24(2):123–40.
140. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32.
141. Schmidhuber J. Deep learning in neural networks – An overview. *Neural Networks.* 2015;61(10):85–117.
142. Bengio Y. Learning Deep Architectures for AI. *Found Trends® Mach Learn.* 2009;2(1):1–127.
143. Hinton GE, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks. *Science* (80-). 2006;313(July):504–7.
144. Murphy KP. *Machine Learning: A Probabilistic Perspective*. 1st editio. The MIT Press. The MIT Press; 2012.
145. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature.* 2015 Feb 25;518:529.
146. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. 2nd editio. The MIT Press. The MIT Press; 2017.
147. Criminisi A, Robertson D, Konukoglu E, Shotton J, Pathak S, White S, et al. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Med Image Anal.* 2013;17(8):1293–303.
148. Muenzing SEA, van Ginneken B, Murphy K, Pluim JPW. Supervised quality assessment of medical image registration: Application to intra-patient CT lung registration. *Med Image Anal.* 2012;16(8):1521–31.
149. Margeta J. *Machine Learning for Simplifying the Use of Cardiac Image Databases l ’ École nationale supérieure des mines de Paris Spécialité “ Contrôle , optimisation et prospective ” Apprentissage automatique pour simplifier l ’ utilisation de banques d ’ images car.* Paris Institute of Technology; 2016.
150. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. *J Am Coll Cardiol.* 2016;68(21):2287–95.
151. Froese T, Hadjiloucas S, Galvao RKH, Becerra VM, Coelho CJ. Comparison of extrasystolic ECG signal classifiers using discrete wavelet transforms. *Pattern Recognit Lett.* 2005;27:393–407.
152. Rajpurkar P, Hannun AY, Haghpanahi M, Bourn C, Ng AY. *Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks.* 2017;
153. Orozco-Duque A, Bustamante J, Castellanos-Dominguez G. Semi-supervised clustering of fractionated electrograms for electroanatomical atrial mapping. *Biomed Eng Online.* 2016;15(1):44.
154. Orozco-Duque A, Martinez-Vargas JD, Novak D, Bustamante J, Castellanos-Dominguez G. Feature Selection for Discrimination of Fractionation Levels in Atrial Electrograms. *36th Annu Int Conf IEEE Eng Med Biol Soc.* 2014;1595–8.
155. Heylman C, Datta R, Sobrino A, George S, Gratton E. Supervised machine learning for classification of

- the electrophysiological effects of chronotropic drugs on human induced pluripotent stem cell-derived cardiomyocytes. *PLoS One*. 2015;10(12):1–15.
156. Cantwell CD, Mohamied Y, Tzortzis KN, Garasto S, Houston C, Chowdhury RA, et al. Rethinking multiscale cardiac electrophysiology with machine learning and predictive modelling. *Comput Biol Med*. 2018;(June).
 157. Duque SI, Orozco-Duque A, Kremen V, Novak D, Tobón C, Bustamante J. Feature subset selection and classification of intracardiac electrograms during atrial fibrillation. *Biomed Signal Process Control*. 2017 Sep;38:182–90.
 158. McGillivray MF, Cheng W, Peters NS, Christensen K. Machine learning methods for locating re-entrant drivers from electrograms in a model of atrial fibrillation. 2017;1–14.
 159. Clifford GD, Liu C, Moody B, Lehman LH, Silva I, Li Q, et al. AF Classification from a Short Single Lead ECG Recording: the PhysioNet/Computing in Cardiology Challenge 2017. *Comput Cardiol*. 2017;2010(1):48–55.
 160. Hong S, Wu M, Zhou Y, Wang Q, Shang J, Li H, et al. ENCASE: an ENsemble CIASsifiEr for ECG Classification Using Expert Features and Deep Neural Networks. 2017;44:2–5.
 161. Zihlmann M, Perekrestenko D, Tschannen M. Convolutional Recurrent Neural Networks for Electrocardiogram Classification. 2018;1–4.
 162. Xiong Z, Stiles MK, Zhao J. Robust ECG Signal Classification for Detection of Atrial Fibrillation Using a Novel Neural Network. *Comput Cardiol (2010)*. 2017;44.
 163. Teijeiro T, García CA, Castro D, Félix P. Arrhythmia Classification from the Abductive Interpretation of Short Single-Lead ECG Records. 2017;
 164. Bin G, Shao M, Bin G, Huang J, Zheng D, Wu S. Detection of Atrial Fibrillation Using Decision Tree Ensemble Detection of Atrial Fibrillation Using Decision Tree Ensemble Method. *Comput Cardiol (2010)*. 2017;44.
 165. Datta S, Puri C, Mukherjee A, Banerjee R, Choudhury AD. Identifying Normal , AF and other Abnormal ECG Rhythms using a Cascaded Binary Classifier Identifying Normal , AF and other Abnormal ECG Rhythms using a Cascaded Binary Classifier. *Comput Cardiol (2010)*. 2017;44(September).
 166. Datta S, Puri C, Mukherjee A, Banerjee R, Dutta Choudhury A, Singh R, et al. Identifying Normal, AF and other Abnormal ECG Rhythms using a Cascaded Binary Classifier. In 2017.
 167. Sutton JR, Mahajan R, Akbilgic O, Kamaleswaran R. PhysOnline : An Open Source Machine Learning Pipeline for Real-Time Analysis of Streaming Physiological Waveform. *IEEE J Biomed Heal Informatics*. 2018;PP(c):1.
 168. Sengupta PP, Kulkarni H, Narula J. Prediction of Abnormal Myocardial Relaxation From Signal Processed Surface ECG. *J Am Coll Cardiol*. 2018;71(15):1650–60.
 169. Hajimolahoseini H, Hashemi J, Gazor S, Redfearn D. Inflection point analysis: A machine learning approach for extraction of IEGM active intervals during atrial fibrillation. *Artif Intell Med*. 2018;85:7–15.
 170. Jekova I. Shock advisory tool: Detection of life-threatening cardiac ar- rhythmias and shock success prediction by means of a common parameter set. *Biomed Signal Process Control*. 2007;2:25–33.
 171. Li Q, Rajagopalan C, Clifford GD. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Trans Biomed Eng*. 2014;61(6):1607–13.
 172. Chlopcíková S, Psotová J, Miketová P, Chlopcikova S, Psotova J, Miketova P. Neonatal rat cardiomyocytes - a model for the study of morphological, biochemical and electrophysiological characteristics of the heart. *Biomed Pap*. 2001;145(2):49–55.
 173. Ali RL, Cantwell CD, Qureshi NA, Roney CH, Lim PB, Sherwin SJ, et al. Automated fiducial point selection for reducing registration error in the co-localisation of left atrium electroanatomic and imaging data. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2015;2015-Novem:1989–92.
 174. Malcolm-Lawes LC, Juli C, Karim R, Bai W, Quest R, Lim PB, et al. Automated analysis of atrial late gadolinium enhancement imaging that correlates with endocardial voltage and clinical outcomes: A 2-center study. *Hear Rhythm*. 2013;10(8):1184–9.
 175. Biermann M, Rubart M, Moreno a, Wu J, Josiah-Durant a, Zipes DP. Differential effects of cytochalasin D and 2,3 butanedione monoxime on isometric twitch force and transmembrane action

- potential in isolated ventricular muscle: implications for optical measurements of cardiac repolarization. *J Cardiovasc Electrophysiol*. 1998;9(12):1348–57.
176. Cheng Y, Li L, Nikolski V, Wallick DW, Efimov IR. Shock-induced arrhythmogenesis is enhanced by 2,3-butanedione monoxime compared with cytochalasin D. *Am J Physiol Heart Circ Physiol*. 2004;286(1):H310–8.
 177. Kettlewell S, Walker NL, Cobbe SM, Burton FL, Smith GL. The electrophysiological mechanical effects of 2,3-butane-dione monoxime and cytochalasin-D in the Langendorff perfused rabbit heart. *Exp Physiol*. 2004;89(2):163–72.
 178. Straight AF, Cheung A, Limouze J, Chen I, Westwood NJ, Sellers JR, et al. Dissecting Temporal and Spatial Control of Cytokinesis with a Myosin II Inhibitor. *Science* (80-). 2011;299(2003):1743–7.
 179. Fedorov V V., Lozinsky IT, Sosunov EA, Anyukhovskiy EP, Rosen MR, Balke CW, et al. Application of blebbistatin as an excitation-contraction uncoupler for electrophysiologic study of rat and rabbit hearts. *Hear Rhythm*. 2007;4(5):619–26.
 180. Brack KE, Narang R, Winter J, Ng GA. The mechanical uncoupler blebbistatin is associated with significant electrophysiological effects in the isolated rabbit heart. *Exp Physiol*. 2013;98(5):1009–27.
 181. Efimov IR, Nikolski VP, Salama G. Optical imaging of the heart. *Circ Res*. 2004;95(1):21–33.
 182. Cox G, Kable E. Second-harmonic imaging of collagen. *Methods Mol Biol*. 2006;319(1):15–35.
 183. Williams RM, Zipfel WR, Webb WW. Interpreting second-harmonic generation images of collagen I fibrils. *Biophys J*. 2005;88(2):1377–86.
 184. Dey A, Learning AS. Machine Learning Algorithms : A Review. *Int J Comput Sci Inf Technol*. 2016;
 185. Zhang S, Zhang C, Yang Q. Data preparation for data mining. *Appl Artif Intell*. 2003;17(5–6):375–81.
 186. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13(1):21–7.
 187. Ye R, Le Z, Suganthan PN. K-nearest neighbor based bagging SVM pruning. *Proc 2013 IEEE Symp Comput Intell Ensemble Learn CIEL 2013 - 2013 IEEE Symp Ser Comput Intell SSCI 2013*. 2013;(April):25–30.
 188. Kemp RA, Macaulay C, Palcic B. Opening the black box: The relationship between neural networks and linear discriminant functions. *Anal Cell Pathol*. 1997;14(1):19–30.
 189. Hautaniemi S, Kharait S, Iwabu A, Wells A, Lauffenburger DA. Modeling of signal-response cascades using decision tree analysis. *Bioinformatics*. 2005;21(9):2027–35.
 190. He X, Chaney NW, Schleiss M, Sheffield J. Spatial downscaling of precipitation using adaptable random forests. *Water Resour Res*. 2016;52(10):8217–37.
 191. Peters J, De Baets B, Verhoest NEC, Samson R, Degroevae S, Becker P De, et al. Random forests as a tool for ecohydrological distribution modelling. *Ecol Modell*. 2007;207(2–4):304–18.
 192. Carlisle DM, Falcone J, Wolock DM, Meador MR, Norris RH. Predicting the natural flow regime: MODels for assessing hydrological alteration in streams. *River Res Appl*. 2010;26:118–36.
 193. Wang Z, Srinivasan R. Homogeneous Ensemble Model for Building Energy Prediction : A Case Study Using Ensemble Regression Tree. *ACEEE Summer Study Energy Effic Build*. 2016;1–12.
 194. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning. Vol. 7, Springer Texts in Statistics. 2000. 316–324 p.
 195. Seiffert C, Khoshgoftaar TM, Van Hulse J, Napolitano A. RUSBoost: Improving classification performance when training data is skewed. 2008 19th Int Conf Pattern Recognit. 2008;(December):1–4.
 196. Yu L, Liu H. Efficient Feature Selection via Analysis of Relevance and Redundancy. *J Mach Learn Res*. 2004;5:1205–1224.
 197. Whitney AW. A Direct Method of Nonparametric Measurement Selection. *IEEE Trans Comput*. 1971;C-20(9):1100–3.
 198. Geurts P, Irtuthum A, Wehenkel L. Supervised learning with decision tree-based methods in computational and systems biology. *Mol Biosyst*. 2009;5(12):1593–605.
 199. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag*. 2009;45(4):427–37.
 200. Zhou X, Ding H, Ung B, Pickwell-MacPherson E, Zhang Y. Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy. *Biomed Eng Online*. 2014;13(1):18.

201. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006;27(8):861–74.
202. Japkowicz N, Stephen S. The class imbalance problem: A systematic study. *Intell Data Anal.* 2002;6(5):429–49.
203. Smith SW. *The Scientist and Engineers's Guide to Digital Signal Processing.* California Technical Publishing, 1999. 688 p.
204. Tan L, Jiang J. *Digital Signal Processing: Fundamentals and Applications.* 2nd ed. Oxford, UK: Elsevier; 2013. 1–13 p.
205. Venkatachalam KL, Herbrandson JE, Asirvatham SJ. Signals and signal processing for the electrophysiologist: Part II: Signal processing and artifact. *Circ Arrhythmia Electrophysiol.* 2011;4(6):974–81.
206. Almeida TP, Chu GS, Salinet JL, Vanheusden FJ, Li X, Tuan JH, et al. Minimizing discordances in automated classification of fractionated electrograms in human persistent atrial fibrillation. *Med Biol Eng Comput.* 2016;54(11):1695–706.
207. Schmitt C, Estner H, Hecher B, Luik A, Kolb C, Karch M, et al. Radiofrequency ablation of complex fractionated atrial electrograms (CFAE): Preferential sites of acute termination and regularization in paroxysmal and persistent atrial fibrillation. *J Cardiovasc Electrophysiol.* 2007;18(10):1039–46.
208. Estner HL, Hessling G, Ndrepepa G, Luik A, Schmitt C, Konietzko A, et al. Acute Effects and Long-Term Outcome of Pulmonary Vein Isolation in Combination With Electrogram-Guided Substrate Ablation for Persistent Atrial Fibrillation. *Am J Cardiol.* 2008;101(3):332–7.
209. Sanders P, Berenfeld O, Hocini M, Jaïs P, Vaidyanathan R, Hsu LF, et al. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. *Circulation.* 2005;112(6):789–97.
210. Takahashi Y, Sanders P, Jaïs P, Hocini M, Dubois R, Rotter M, et al. Organization of frequency spectra of atrial fibrillation: Relevance to radiofrequency catheter ablation. *J Cardiovasc Electrophysiol.* 2006;17(4):382–8.
211. Lin YJ, Tsao HM, Chang SL, Lo LW, Hu YF, Chang CJ, et al. Role of high dominant frequency sites in nonparoxysmal atrial fibrillation patients: Insights from high-density frequency and fractionation mapping. *Hear Rhythm.* 2010;7(9):1255–62.
212. Ravelli F, Mase M. Computational mapping in atrial fibrillation: How the integration of signal-derived maps may guide the localization of critical sources. *Europace.* 2014;16(5):714–23.
213. Schilling C, Keller M, Scherr D, Oesterlein T, Haïssaguerre M, Schmitt C, et al. Fuzzy decision tree to classify complex fractionated atrial electrograms. *Biomed Tech.* 2015;60(3):245–55.
214. Liu J, Li S, Li X, Klein C, Rymer WZ, Zhou P. Suppression of stimulus artifact contaminating electrically evoked electromyography. *NeuroRehabilitation.* 2014;34(2):381–9.
215. Rosenbaum DS, Jackson LE, Smith JM, Garan H, Ruski JN, Cohen RJ. Electrical alternans and vulnerability to ventricular arrhythmias. *N Engl J Med.* 1994;327(27):1893–8.
216. Estes NAM, Michaud G, Zipes DP, El-Sherif N, Venditti FJ, Rosenbaum DS, et al. Electrical alternans during rest and exercise as predictors of vulnerability to ventricular arrhythmias. *Am J Cardiol.* 1997;80(10):1314–8.
217. Markovitch S, Rosenstein D. Feature generation using general constructor functions. *Mach Learn.* 2002;49(1):59–98.
218. Chavan MS, Agarwala RA, Uplane MD. Comparative Study of Chebyshev I and Chebyshev II Filter used For Noise Reduction in ECG Signal. *Int J Circuits, Syst Signal Process.* 2008;2(1):1–17.
219. Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal Chem.* 1964;36(8):1627–39.
220. Ng J, Kadish AH, Goldberger JJ. Technical considerations for dominant frequency analysis. *J Cardiovasc Electrophysiol.* 2007;18(7):757–64.
221. Hall P, River US. Spectral Analysis of Signals. *Spectral Element Method in Structural Dynamics.* 2009. 4–7 p.
222. Engelberg S. *Digital Signal Processing: An Experimental Approach.* In: *Signals and Communication Technology.* 2008. p. 56.
223. Welch PD. The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans Audio Electroacoust.* 1967;(2):70–3.

224. Čisar P, Čisar SM. Skewness and kurtosis in function of selection of network traffic distribution. *Acta Polytech Hungarica*. 2010;7(2):95–106.
225. Addison PS. Wavelet transforms and the ECG: a review. *Physiol Meas*. 2005;26(5):R155–99.
226. Talbi M, Aouinet A. ECG Analysis based on Wavelet Transform and Modulus Maxima. *Int J Comput Sci Issues*. 2012;9(1):427–35.
227. Aydın S, Saraoğlu HM, Kara S. Log energy entropy-Based EEG classification with multilayer neural networks in seizure. *Ann Biomed Eng*. 2009;37(12):2626–30.
228. Coifman RR, Wickerhauser MV. Entropy Based Algorithms for Best Basis Selection. *IEEE Trans Inf Theory*. 1992;32(2):712–8.
229. Guillén SG, Arredondo MT, Martín G, Ferrero Corral JM. Ventricular fibrillation detection by autocorrelation function peak analysis. *J Electrocardiol*. 1990;22(SUPPL.):253–62.
230. BRAUNSTEIN JR, FRANKE EK. Autocorrelation of Ventricular Response in Atrial Fibrillation. *Circ Res*. 1961;9(2):300–304.
231. Throne RD. Detecting ventricular fibrillation using efficient techniques for computing a normalized autocorrelation. *Comput Biol Med*. 1993;23(4):317–25.
232. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Phys Rev A*. 1986;33(2):1134–40.
233. Zduniak A, Lusakowski J. Autocorrelation function and mutual information from short experimental timeseries. *Proc 23rd Int Sch Semicond Compd*. 1995;87(1):257–60.
234. Guvenir H a., Acar B, Demiroz G, Cekin a. A supervised machine learning algorithm for arrhythmia analysis. *Comput Cardiol* 1997. 1997;
235. Nollo G, Marconcini M, Faes L, Bovolo F, Ravelli F, Bruzzzone L. An Automatic System for the Analysis and Classification of Human Atrial Fibrillation Patterns from Intracardiac Electrograms. *IEEE Trans Biomed Eng*. 2008;55(9):2275–85.
236. Kadambe S, Murray R, Boudreaux-Bartels GF. Wavelet transform-based QRS complex detector. *IEEE Trans Biomed Eng*. 1999;46(7):838–48.
237. Köhler BU, Hennig C, Orglmeister R. The principles of software QRS detection. *IEEE Eng Med Biol Mag*. 2002;21(1):42–57.
238. Vázquez-Seisdedos CR, Neto J, Marañón Reyes EJ, Klautau A, Limão de Oliveira RC. New approach for T-wave end detection on electrocardiogram: Performance in noisy conditions. *Biomed Eng Online*. 2011;10(1):77.
239. Last T, Nugent CD, Owens FJ. Multi-component based cross correlation beat detection in electrocardiogram analysis. *Biomed Eng Online*. 2004;3:1–14.
240. Madeiro JP V, Nicolson WB, Cortez PC, Marques JAL, Vazquez-Seisdedos CR, Elangovan N, et al. New approach for T-wave peak detection and T-wave end location in 12-lead paced ECG signals based on a mathematical model. *Med Eng Phys*. 2013;35(8):1105–15.
241. Xue J, Gao W, Chen Y, Han X. Identify drug-induced T wave morphology changes by a cell-to-electrocardiogram model and validation with clinical trial data. *J Electrocardiol*. 2009;42(6):534–42.
242. De Bakker JMT, Wittkampf FHM. The pathophysiologic basis of fractionated and complex electrograms and the impact of recording techniques on their detection and interpretation. *Circ Arrhythmia Electrophysiol*. 2010;3(2):204–13.
243. Holmqvist F, Bahnson TD. Atrial fibrillation mapping strategies: The hunt is on. *Hear Rhythm*. 2018;15(9):1304–5.
244. Himel HD, Knisley SB. Comparison of optical and electrical mapping of fibrillation. *Physiol Meas*. 2007;28(6):707–19.
245. Knisley SB, Neuman MR. Simultaneous electrical and optical mapping in rabbit hearts. *Ann Biomed Eng*. 2003;31(1):32–41.
246. Spach MS, Barr RC, Serwer GA, Kootsey JM, Johnson EA. Extracellular potentials related to intracellular action potentials in the dog purkinje system. *Circ Res*. 1971;XIII(2):437–45.
247. Halbach MD, Egert U, Hescheler J. Estimation of action potential changes from field potential recordings in multicellular mouse cardiac myocyte cultures. *Cell Physiol Biochem*. 2003;13:271–84.
248. Xu L, Gutbrod SR, Bonifas AP, Su Y, Sulkin MS, Lu N, et al. 3D multifunctional integumentary

- membranes for spatiotemporal cardiac measurements and stimulation across the entire epicardium. *Nat Commun.* 2014;5:1–10.
249. Himmel HM, Bussek A, Hoffmann M, Beckmann R, Lohmann H, Schmidt M, et al. Field and action potential recordings in heart slices: Correlation with established in vitro and in vivo models. *Br J Pharmacol.* 2012;166(1):276–96.
 250. Lee H-C, Cai JJ, Arnar DO, Shibata EF, Martins JB. Mechanism of Alpha-2 adrenergic modulation of canine cardiac Purkinje action potential. *J Pharmacol Exper Ther.* 1996;278(2):597–606.
 251. Hayashi S, Kii Y, Tabo M, Fukuda H. QT PRODACT: A multi-site study of in vitro action potential assays on 21 compounds in isolated guinea-pig papillary muscles. *J Pharmacol Sci.* 2005;99:423–37.
 252. Dumotier BM, Adamantidis MM, Puisieux FL, Bastide MM, Dupuis BA. Repercussions of pharmacologic reduction in ionic currents on action potential configuration in rabbit Purkinje fibers: Are they indicative of proarrhythmic potential? *Drug Dev Res.* 1999;47(2):63–76.
 253. Salama G, Kanai a, Efimov IR. Subthreshold stimulation of Purkinje fibers interrupts ventricular tachycardia in intact hearts. Experimental study with voltage-sensitive dyes and imaging techniques. *Circ Res.* 1994;74(4):604–19.
 254. Ragsdale DS, McPhee JC, Scheuer T, Catterall WA. Common molecular determinants of local anesthetic, antiarrhythmic, and anticonvulsant block of voltage-gated Na⁺ channels. *Proc Natl Acad Sci.* 1996;93(17):9270–5.
 255. Bean BP, Cohen CJ, Tsien RW. Lidocaine block of cardiac sodium channels. *J Gen Physiol.* 1983;81(5):613–42.
 256. Anderson KP, Walker R, Lux RL, Ershler PR, Menlove R, Williams MR, et al. Conduction velocity depression and drug-induced ventricular tachyarrhythmias. Effects of lidocaine in the intact canine heart. *Circulation.* 1990;81(3):1024–38.
 257. Hondeghem LM, Cotner CL. Reproducible and uniform cardiac ischemia: effects of antiarrhythmic drugs. *Am J Physiol.* 1978;235(5):H574–80.
 258. Gray B, Semsarian C, Sy RW. Brugada syndrome: A heterogeneous disease with a common ECG phenotype? *J Cardiovasc Electrophysiol.* 2014;25(4):450–6.
 259. Nademanee K, Veerakul G, Chandanamattha P, Chaothawee L, Ariyachaipanich A, Jirasirojanakorn K, et al. Prevention of ventricular fibrillation episodes in brugada syndrome by catheter ablation over the anterior right ventricular outflow tract epicardium. *Circulation.* 2011;123(12):1270–9.
 260. Bassani R a. Transient outward potassium current and Ca²⁺ homeostasis in the heart: Beyond the action potential. *Brazilian J Med Biol Res.* 2006;39(3):393–403.
 261. Xu H, Guo W, Nerbonne JM. Four Kinetically Distinct Depolarization-activated K⁺ Currents in Adult Mouse Ventricular Myocytes. *J Gen Physiol.* 1999;113(5):661–78.
 262. Wickenden AD, Kaprielian R, Parker TG, Jones OT, Backx PH. Effects of Development and Thyroid Hormone on K⁺ Currents and K⁺ Channel Gene Expression in Rat Ventricle. *J Physiol.* 1997;504(2):271–86.
 263. Kilborn MJ, Fedida D. A study of the developmental changes in outward currents of rat ventricular myocytes. *J Physiol.* 1990;430:37–60.
 264. Walker KE, Lakatta EG, Houser SR. Age associated changes in membrane currents in rat ventricular myocytes. *CardiovascRes.* 1993;27(85):1968–77.
 265. Mitcheson JS, Hancox JC. Characteristics of a transient outward current (sensitive to 4-aminopyridine) in Ca²⁺-tolerant myocytes isolated from the rabbit atrioventricular node. *Pflugers Arch.* 1999;438(1):68–78.
 266. Li GR, Yang B, Sun H, Baumgarten CM. Existence of a transient outward K(+) current in guinea pig cardiac myocytes. *Am J Physiol Heart Circ Physiol.* 2000;279(1):H130–8.
 267. Wolk R, Cobbe SM, Hicks MN, Kane KA. Functional, structural, and dynamic basis of electrical heterogeneity in healthy and diseased cardiac muscle implications for arrhythmogenesis and anti-arrhythmic drug therapy. *Pharmacol Ther.* 1999;84(2):207–31.
 268. Shannon TR, Ginsburg KS, Bers DM. Potentiation of fractional sarcoplasmic reticulum calcium release by total and free intra-sarcoplasmic reticulum calcium concentration. *Biophys J.* 2000;78(1):334–43.
 269. Henry PD. Comparative pharmacology of calcium antagonists: nifedipine, verapamil and diltiazem. *Am J Cardiol.* 1980;46(6):1047–58.

270. Gussak I, Chaitman BR, Kopecky SL, Nerbonne JM. Rapid ventricular repolarization in rodents: electrocardiographic manifestations, molecular mechanisms, and clinical insights. *J Electrocardiol.* 2000;33(2):159–70.
271. Zhang S, Zhou Z, Gong Q, Makielski JC, January CT. Mechanism of block and identification of the verapamil binding domain to HERG potassium channels. *Circ Res.* 1999;84(9):989–98.
272. Sanguinetti MC, Jurkiewicz NK. Two components of cardiac delayed rectifier K⁺ current. Differential sensitivity to block by class III antiarrhythmic agents. *J Gen Physiol.* 1990;96(1):195–215.
273. Katoh H, Ogawa S, Furuno I, Sato Y, Yoh S, Saeki K, et al. Electrophysiologic effects of E-4031, a class III antiarrhythmic agent, on re-entrant ventricular arrhythmias in a canine 7-day-old myocardial infarction model. *J Pharmacol Exp Ther.* 1990;253(3):1077–82.
274. Lohmann H, Bussek A, Schmidt M, Wettwer E, Ravens U. Isolated Living Heart Slices From Adult Rats and Guinea Pigs. *Pharmacology.* 2006;372(March):2006.
275. Guo J, Massaeli H, Li W, Xu J, Luo T, Shaw J, et al. Identification of I_{Kr} and Its Trafficking Disruption Induced by Probucol in Cultured Neonatal Rat Cardiomyocytes. *Pharmacology.* 2007;321(3):911–20.
276. Sanguinetti MC, Tristani-Firouzi M. hERG potassium channels and cardiac arrhythmia. *Nature.* 2006;440(7083):463–9.
277. Yue L, Xie J, Nattel S. Molecular determinants of cardiac fibroblast electrical function and therapeutic implications for atrial fibrillation. *Cardiovasc Res.* 2011;89(4):744–53.
278. Rohr S. Myofibroblasts in diseased hearts: New players in cardiac arrhythmias? *Hear Rhythm.* 2009;6(6):848–56.
279. Vasquez C, Mohandas P, Louie KL, Benamer N, Bapat AC, Morley GE. Enhanced fibroblast-myocyte interactions in response to cardiac injury. *Circ Res.* 2010;107(8):1011–20.
280. Sanguinetti MC, Jurkiewicz NK. Delayed rectifier outward K⁺ current is composed of two currents in guinea pig atrial cells. *Am J Physiol.* 1991;260(2 Pt 2):H393-9.
281. Antzelevitch C, Shimizu W. Cellular mechanisms underlying the long QT syndrome. *Curr Opin Cardiol.* 2002;17(0268-4705 SB-IM):43–51.
282. Camm AJ, Yap YG. Clinical trials of antiarrhythmic drugs in postmyocardial infarction and congestive heart failure patients. *J Cardiovasc Pharmacol Ther.* 2001;6(1):99–106.
283. Thomas GP, Gerlach U, Antzelevitch C. HMR 1556, a potent and selective blocker of slowly activating delayed rectifier potassium current. *J Cardiovasc Pharmacol.* 2003;41(1):140–7.
284. Gerlach U, Brendel J, Lang HJ, Paulus EF, Weidmann K, Brüggemann A, et al. Synthesis and activity of novel and selective I_{Ks}-channel blockers. *J Med Chem.* 2001;44(23):3831–7.
285. Gogelein H, Brüggemann A, Gerlach U, Brendel J, Busch AE. Inhibition of I(K_s) channels by HMR 1556. *Naunyn Schmiedebergs Arch Pharmacol.* 2000;362(6):480–8.
286. Muñoz V, Grzeda KR, Desplantez T, Pandit S V., Mironov S, Taffet SM, et al. Adenoviral expression of I_{Ks} contributes to wavebreak and fibrillatory conduction in neonatal rat ventricular cardiomyocyte monolayers. *Circ Res.* 2007;101(5):475–83.
287. So PPS, Hu XD, Backx PH, Puglisi JL, Dorian P. Blockade of I_{Ks} by HMR 1556 increases the reverse rate-dependence of refractoriness prolongation by dofetilide in isolated rabbit ventricles. *Br J Pharmacol.* 2006;148(3):255–63.
288. Porta-Sánchez A, Spillane DR, Harris L, Xue J, Dorsey P, Care M, et al. T-Wave Morphology Analysis in Congenital Long QT Syndrome Discriminates Patients From Healthy Individuals. *JACC Clin Electrophysiol.* 2017;3(4):374–81.
289. Tseng GN, Hoffman BF. Actions of pinacidil on membrane currents in canine ventricular myocytes and their modulation by intracellular ATP and cAMP. *Pflügers Arch Eur J Physiol.* 1990;415(4):414–24.
290. Nichols CG, Singh GK, Grange DK. KATP Channels and Cardiovascular Disease: Suddenly a Syndrome. *Circ Res.* 2013;112(7):1059–72.
291. Yokoshiki H, Sunagawa M, Seki T, Sperelakis N. Antisense oligodeoxynucleotides of sulfonylurea receptors inhibit ATP-sensitive K⁺ channels in cultured neonatal rat ventricular cells. *Pflügers Arch.* 1999;437(3):400–8.
292. Neophytou N, Kyriakides A, Pitris C. ECG analysis in the Time-Frequency domain. 2012 IEEE 12th Int Conf Bioinforma Bioeng. 2012;(November):80–4.

293. Nattel S, Maguy A, Le Bouter S, Yeh Y-H. Arrhythmogenic ion-channel remodeling in the heart: heart failure, myocardial infarction, and atrial fibrillation. *Physiol* 2007;87(2):425–56.
294. Bosch RF, Zeng X, Grammer JB, Popovic K, Mewis C, Kühlkamp V. Ionic mechanisms of electrical remodeling in human atrial fibrillation. *Cardiovasc Res*. 1999;44(1):121–31.
295. Brundel BJM, Van Gelder IC, Henning RH, Tieleman RG, Tuinenburg AE, Wietses M, et al. Ion channel remodeling is related to intraoperative atrial effective refractory periods in patients with paroxysmal and persistent atrial fibrillation. *Circulation*. 2001;103(5):684–90.
296. Sun H, Chartier D, Leblanc N, Nattel S. Intracellular calcium changes and tachycardia-induced contractile dysfunction in canine atrial myocytes. *Cardiovasc Res*. 2001;49(4):751–61.
297. Nattel S. New ideas about atrial fibrillation 50 years on. 2002;415(January):219–26.
298. Van Wagoner DR, Pond AL, Lamorgese M, Rossie SS, McCarthy PM, Nerbonne JM. Atrial L-Type Ca²⁺ Currents and Human Atrial Fibrillation. *Circ Res*. 1999;85(5):428–36.
299. Toyoshima S, Kanno A, Kitayama T, Sekiya K, Nakai K, Haruna M, et al. QT PRODACT : In Vivo QT Assay in the Conscious Dog for Assessing the Potential for QT Interval Prolongation by Human Pharmaceuticals. *J Pharmacol Sci*. 2005;99(5):459–71.
300. Grant AO. Cardiac Ion Channels. *Circ Arrhythmia Electrophysiol*. 2009;2(2):185–94.
301. Peters NS, Green CR, Severs NJ. Reduced Content of Connexin43 Gap Junctions in Ventricular Myocardium From Hypertrophied and Ischemic Human Hearts. *Circulation*. 1993;88(3):864–75.
302. Van Der Velden HMW, Ausma J, Rook MB, Hellemons AJCGM, Van Veen T a a B, Allessie M a., et al. Gap junctional remodeling in relation to stabilization of atrial fibrillation in the goat. *Cardiovasc Res*. 2000;46(3):476–86.
303. Patel PM, Plotnikov A, Kanagaratnam P, Shvilkin A, Sheehan CT, Xiong W, et al. Altering ventricular activation remodels gap junction distribution in canine heart. *J Cardiovasc Electrophysiol*. 2001;12(5):570–7.
304. Akar FG, Spragg DD, Tunin RS, Kass DA, Tomaselli GF. Mechanisms underlying conduction slowing and arrhythmogenesis in nonischemic dilated cardiomyopathy. *Circ Res*. 2004;95(7):717–25.
305. Lampe PD, Cooper CD, King TJ, Burt JM. Analysis of Connexin43 phosphorylated at S325, S328 and S330 in normoxic and ischemic heart. *J Cell Sci*. 2006;119(16):3435–42.
306. Poelzing S, Rosenbaum DS. Altered connexin43 expression produces arrhythmia substrate in heart failure. *Am J Physiol Circ Physiol*. 2004;287(4):H1762–70.
307. van der Velden HM, van Kempen MJ, Wijffels MC, van Zijverden M, Groenewegen WA, Allessie MA, et al. Altered pattern of connexin40 distribution in persistent atrial fibrillation in the goat [see comments]. *J Cardiovasc Electrophysiol*. 1998;9(6):596–607.
308. Wit AL, Duffy HS. Drug development for treatment of cardiac arrhythmias: targeting the gap junctions. *Am J Physiol Heart Circ Physiol*. 2008;294(1):H16–8.
309. Burt JM, Massey KD, Minnich BN. Uncoupling of cardiac cells by fatty acids: structure-activity relationships. *Am J Physiol*. 1991;260(3 Pt 1):C439–48.
310. Dhein S, Krüsemann K, Schaefer T. Effects of the gap junction uncoupler palmitoleic acid on the activation and repolarization wavefronts in isolated rabbit hearts. *Br J Pharmacol*. 1999;128(7):1375–84.
311. Goldberg GS, Moreno AP, Bechberger JF, Hearn SS, Shivers RR, Macphree DJ, et al. Evidence that disruption of connexon particle arrangements in gap junction plaques is associated with inhibition of gap junctional communication by a glycyrrhetic acid derivative. *Exp Cell Res*. 1996;222(1):48–53.
312. Dhein S, Manicone N, Müller A, Gerwin R, Ziskoven U, Irankhahi A, et al. A new synthetic antiarrhythmic peptide reduces dispersion of epicardial activation recovery interval and diminishes alterations of epicardial activation patterns induced by regional ischemia - A mapping study. *Naunyn Schmiedebergs Arch Pharmacol*. 1994;350(2):174–84.
313. Hennen JK, Swillo RE, Morgan GA, Keith JC, Schaub RG, Smith RP, et al. Rotigaptide (ZP123) Prevents Spontaneous Ventricular Arrhythmias and Reduces Infarct Size During Myocardial Ischemia / Reperfusion Injury in Open-Chest Dogs. 2006;317(1):236–43.
314. DeGroot J, Veenstra T, Verkerk A, Wilders R, Smits J, Wilmsschopman F, et al. Conduction slowing by the gap junctional uncoupler carbenoxolone. *Cardiovasc Res*. 2003;60(2):288–97.
315. Kirubakaran S, Chowdhury R a, Hall MCS, Patel PM, Garratt CJ, Peters NS. Fractionation of electrograms is caused by colocalized conduction block and connexin disorganization in the absence of

- fibrosis as AF becomes persistent in the goat model. *Heart Rhythm*. 2014;12(2):1–12.
316. Stauffer BL, Sobus RD, Sucharov CC. Sex differences in cardiomyocyte connexin43 expression. *J Cardiovasc Pharmacol*. 2011;58(1):32–9.
 317. Vozzi C, Dupont E, Coppen SR, Hung-I Y, Severs NJ. Chamber-related differences in connexin expression in the human heart. *J Mol Cell Cardiol*. 1999;31(5):991–1003.
 318. Dhillon PS, Gray R, Kojodjojo P, Jabr R, Chowdhury R, Fry CH, et al. Relationship between gap-junctional conductance and conduction velocity in mammalian myocardium. *Circ Arrhythmia Electrophysiol*. 2013;6(6):1208–14.
 319. Howarth FC, Qureshi MA. Effects of carbenoxolone on heart rhythm, contractility and intracellular calcium in streptozotocin-induced diabetic rat. *Mol Cell Biochem*. 2006;289(1–2):21–9.
 320. Kojodjojo P, Kanagaratnam P, Segal OR, Hussain W, Peters NS. The Effects of Carbenoxolone on Human Myocardial Conduction. A Tool to Investigate the Role of Gap Junctional Uncoupling in Human Arrhythmogenesis. *J Am Coll Cardiol*. 2006;48(6):1242–9.
 321. Lin X, Gemel J, Glass A, Zemlin CW, Beyer EC, Veenstra RD. Connexin40 and Connexin43 determine gating properties of atrial gap junction channels. *J Mol Cell Cardiol*. 2010;48(1):238.
 322. Fast VG, Kléber a G. Microscopic conduction in cultured strands of neonatal rat heart cells measured with voltage-sensitive dyes. *Circ Res*. 1993;73(5):914–25.
 323. Bussek A, Wettwer E, Christ T, Lohmann H, Camelliti P, Ravens U. Tissue slices from adult mammalian hearts as a model for pharmacological drug testing. *Cell Physiol Biochem*. 2009;24:527–36.
 324. Wang K, Lee P, Mirams GR, Sarathchandra P, Borg TK, Gavaghan DJ, et al. Cardiac tissue slices: preparation, handling, and successful optical mapping. *Am J Physiol - Hear Circ Physiol*. 2015;
 325. Pillekamp F, Halbach M, Reppel M, Rubenchyk O, Pfannkuche K, Xi JY, et al. Neonatal murine heart slices. A robust model to study ventricular isometric contractions. *Cell Physiol Biochem*. 2007;20(6):837–46.
 326. Watson SA, Scigliano M, Bardi I, Ascione R, Terracciano CM, Perbellini F. Preparation of viable adult ventricular myocardial slices from large and small mammals. *Nat Protoc*. 2017;12(12):2623–39.
 327. Claycomb WC. Biochemical aspects of cardiac muscle differentiation. *J Biol Chem*. 1976;251(19):6082–9.
 328. Perbellini F, Watson SA, Scigliano M, Alayoubi S, Tkach S, Bardi I, et al. Investigation of cardiac fibroblasts using myocardial slices. *Cardiovasc Res*. 2017;(February):77–89.
 329. Perbellini F, Liu AKL, Watson SA, Bardi I, Rothery SM, Terracciano CM. Free-of-Acrylamide SDS-based Tissue Clearing (FASTClear) for three dimensional visualization of myocardial tissue. *Sci Rep*. 2017;7(1):5188.
 330. Carmeliet E. Conduction in cardiac tissue. Historical reflections. *Physiol Rep*. 2019;7(1):e13860.
 331. Roney CH, Whitaker J, Sim I, O’Neill L, Mukherjee RK, Razeghi O, et al. A technique for measuring anisotropy in atrial conduction to estimate conduction velocity and atrial fibre direction. *Comput Biol Med*. 2018;104(July 2018):278–90.
 332. Jadidi AS, Cochet H, Shah AJ, Kim SJ, Duncan E, Miyazaki S, et al. Inverse Relationship Between Fractionated Electrograms and Atrial Fibrosis in Persistent Atrial Fibrillation. *J Am Coll Cardiol*. 2013;62(9):802–12.
 333. Noorman M, van der Heyden MAG, van Veen TAB, Cox MGPJ, Hauer RNW, de Bakker JMT, et al. Cardiac cell-cell junctions in health and disease: Electrical versus mechanical coupling. *J Mol Cell Cardiol*. 2009;47(1):23–31.
 334. Severs NJ, Bruce AF, Dupont E, Rothery S. Remodelling of gap junctions and connexin expression in diseased myocardium. *Biochim Biophys Acta*. 2004;1662(1):138–48.
 335. Rohr S, Kucera JP, Fast VG, Kleber AG. Paradoxical Improvement of Impulse Conduction in Cardiac Tissue by Partial Cellular Uncoupling Paradoxical Improvement of Impulse Conduction in Cardiac Tissue by Partial Cellular Uncoupling. 1997;275:841–4.
 336. Verma A, Jiang CY, Betts TR, Chen J, Deisenhofer I, Mantovan R, et al. Approaches to catheter ablation for persistent atrial fibrillation. *N Engl J Med*. 2015;372(19):1812–22.
 337. Verma A, Sanders P, Champagne J, Macle L, Nair GM, Calkins H, et al. Selective complex fractionated atrial electrograms targeting for atrial fibrillation study (SELECT AF): A multicenter, randomized trial. *Circ Arrhythmia Electrophysiol*. 2014;7(1):55–62.

338. Rolf S, Kircher S, Arya A, Eitel C, Sommer P, Sergio R, et al. Tailored atrial substrate modification based on low-voltage areas in catheter ablation of atrial fibrillation. *Circ Arrhythmia Electrophysiol.* 2014;7(5):825–33.
339. Yang G, Yang B, Wei Y, Zhang F, Ju W, Chen H, et al. Catheter Ablation of Nonparoxysmal Atrial Fibrillation Using Electrophysiologically Guided Substrate Modification during Sinus Rhythm after Pulmonary Vein Isolation. *Circ Arrhythmia Electrophysiol.* 2016;9(2).
340. Ghoraani B, Dalvi R, Gizurarson S, Das M, Ha A, Suszko A, et al. Localized rotational activation in the left atrium during human atrial fibrillation: Relationship to complex fractionated atrial electrograms and low-voltage zones. *Hear Rhythm.* 2013;10(12):1830–8.
341. Sun Y, Kiani MF, Postlethwaite AE, Weber KT. Infarct scar as living tissue. *Basic Res Cardiol.* 2002;97(5):343–7.
342. Sun Y, Weber KT. Infarct Scar a Dynamic Tissue. 2000;46:250–6.
343. Pouliopoulos J, Chik WWB, Kanthan A, Sivagangabalan G, Barry MA, Fahmy PNA, et al. Intramyocardial adiposity after myocardial infarction: New implications of a substrate for ventricular tachycardia. *Circulation.* 2013;128(21):2296–308.
344. Ichikawa Y, Kitagawa K, Chino S, Ishida M, Matsuoka K, Tanigawa T, et al. Adipose Tissue Detected by Multislice Computed Tomography in Patients After Myocardial Infarction. *JACC Cardiovasc Imaging.* 2009;2(5):548–55.
345. Kohl P, Gourdie RG. Fibroblast-myocyte electrotonic coupling: Does it occur in native cardiac tissue? *J Mol Cell Cardiol.* 2014;70:37–46.
346. Pedrotty DM, Klinger RY, Kirkton RD, Bursac N. Cardiac fibroblast paracrine factors alter impulse conduction and ion channel expression of neonatal rat cardiomyocytes. *Cardiovasc Res.* 2009;83:688–97.
347. Guo W, Kamiya K, Yasui K, Kodama I, Toyama J. Paracrine hypertrophic factors from cardiac non-myocyte cells downregulate the transient outward current density and Kv4.2 K⁺ channel expression in cultured rat cardiomyocytes. *Cardiovasc Res.* 1999;41(1):157–65.
348. Srinivasan NT, Orini M, Providencia R, Dhinoja MB, Lowe MD, Ahsan SY, et al. Prolonged action potential duration and dynamic transmural action potential duration heterogeneity underlie vulnerability to ventricular tachycardia in patients undergoing ventricular tachycardia ablation. *EP Eur.* 2018;1–10.
349. Glukhov A V., Fedorov V V., Lou Q, Ravikumar VK, Kalish PW, Schuessler RB, et al. Transmural dispersion of repolarization in failing and non failing human ventricle. *Circ Res.* 2010;106(5):981–91.
350. Camelliti P, Borg TK, Kohl P. Structural and functional characterisation of cardiac fibroblasts. *Cardiovasc Res.* 2005;65(1):40–51.
351. Mahoney VM, Mezzano V, Morley GE. A review of the literature on cardiac electrical activity between fibroblasts and myocytes. *Prog Biophys Mol Biol.* 2016;120(1–3):128–33.
352. Kohl P, Camelliti P, Burton FL, Smith GL. Electrical coupling of fibroblasts and myocytes: Relevance for cardiac propagation. *J Electrocardiol.* 2005;38(4 SUPPL.):45–50.
353. Manabe I, Shindo T, Nagai R. Gene expression in fibroblasts and fibrosis involvement in cardiac hypertrophy. *Circ Res.* 2002;91(12):1103–13.
354. Brown RD, Ambler SK, Mitchell MD, Long CS. THE CARDIAC FIBROBLAST: Therapeutic Target in Myocardial Remodeling and Failure. *Annu Rev Pharmacol Toxicol.* 2005;45(1):657–87.
355. He X, Gao X, Peng L, Wang S, Zhu Y, Ma H, et al. Atrial fibrillation induces myocardial fibrosis through angiotensin II type 1 receptor-specific Arkadia-mediated downregulation of Smad7. *Circ Res.* 2011;108(2):164–75.
356. Miragoli M, Gaudesius G, Rohr S. Electrotonic modulation of cardiac impulse conduction by myofibroblasts. *Circ Res.* 2006;98(6):801–10.
357. Camelliti P, Devlin GP, Matthews KG, Kohl P, Green CR. Spatially and temporally distinct expression of fibroblast connexins after sheep ventricular infarction. *Cardiovasc Res.* 2004;62(2):415–25.
358. Baum J, Duffy HS. Fibroblasts and myofibroblasts: what are we talking about? *J Cardiovasc Pharmacol.* 2011;57(4):376–9.
359. McDowell KS, Arevalo HJ, Maleckar MM, Trayanova NA. Susceptibility to arrhythmia in the infarcted heart depends on myofibroblast density. *Biophys J.* 2011;101(6):1307–15.
360. Kohl P, Camelliti P. Fibroblast-myocyte connections in the heart. *Hear Rhythm.* 2012;9(3):461–4.

361. Xie Y, Garfinkel A, Camelliti P, Kohl P, Weiss JN, Qu Z. Effects of fibroblast-myocyte coupling on cardiac conduction and vulnerability to reentry: A computational study. *Hear Rhythm*. 2009;6(11):1641–9.
362. Fernandez NF, Gundersen GW, Rahman A, Grimes ML, Rikova K, Hornbeck P, et al. Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data. *Sci Data*. 2017;4:1–12.
363. Ashihara T, Haraguchi R, Nakazawa K, Namba T, Ikeda T, Nakazawa Y, et al. The role of fibroblasts in complex fractionated electrograms during persistent/permanent atrial fibrillation: Implications for electrogram-based catheter ablation. *Circ Res*. 2012;110(2):275–84.
364. Jacquemet V, Henriquez CS. Genesis of complex fractionated atrial electrograms in zones of slow conduction: A computer model of microfibrosis. *Hear Rhythm*. 2009;6(6):803–10.
365. Sau A, Qureshi N, Bai W, Kim S, Ariff B, Ng FS, et al. Late-gadolinium enhanced cardiac MRI defined scar is predominantly located on the left atrial septum and posterior wall in patients with persistent atrial fibrillation. In: *Annual Meeting of the European Society of Cardiology*. OXFORD UNIV PRESS; 2014. p. 799–799.
366. Antanavičiute I, Mildažiene V, Stankevičius E, Herdegen T, Skeberdis VA. Hyperthermia differently affects connexin43 expression and gap junction permeability in skeletal Myoblasts and Hela cells. *Mediators Inflamm*. 2014;2014.
367. Cartledge JE, Kane C, Dias P, Tesfom M, Clarke L, Mckee B, et al. Functional crosstalk between cardiac fibroblasts and adult cardiomyocytes by soluble mediators. *Cardiovasc Res*. 2015;105(3):260–70.
368. Weber KT, Sun Y, Bhattacharya SK, Ahokas R a, Gerling IC. Myofibroblast-mediated mechanisms of pathological remodelling of the heart. *Nat Rev Cardiol*. 2013;10(1):15–26.
369. Frangogiannis NG, Michael LH, Entman ML. Myofibroblasts in reperfused myocardial infarcts express the embryonic form of smooth muscle myosin heavy chain (SMemb). *Cardiovasc Res*. 2000;48(1):89–100.
370. Masur S, Dewal H, Dinh T, Erenburg I, Petridou S. Myofibroblasts differentiate from fibroblasts when plated at low density. *Proc Natl Acad Sci U S A*. 1996;93(April):4219–23.
371. Miragoli M, Salvarani N, Rohr S. Myofibroblasts induce ectopic activity in cardiac tissue. *Circ Res*. 2007;101(8):755–8.
372. Tsai M-R, Chiu Y-W, Lo MT, Sun C-K. Second-harmonic generation imaging of collagen fibers in myocardium for atrial fibrillation diagnosis. *J Biomed Opt*. 2014;15(April 2010):026002.
373. Chen X, Nadiarynkh O, Plotnikov S, Campagnola PJ. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat Protoc*. 2012;7(4):654–69.
374. Yap Y, Camm J. Risk of torsades de pointes with non-cardiac drugs. Doctors need to be aware that many drugs can cause QT prolongation. *BMJ*. 2000;320:1158–9.
375. Varró A, Baláti B, Iost N, Takács J, Virág L, Lathrop DA, et al. The role of the delayed rectifier component I(Ks) in dog ventricular muscle and Purkinje fibre repolarization. *J Physiol*. 2000;523(1):67–81.
376. Jost, N.; Virag L.; Bitay M.; Takacs, J.; Lengyel, C.; Biliczki, P.; Nagy, Z.; Bogats, G.; Lathrop, D.A.; Papp, J.G.; Varro A. Restricting Excessive Cardiac Action Potential and QT Prolongation: A Vital Role for IKs in Human Ventricular Muscle. *Circulation*. 2005;112(10):1392–9.
377. Terracciano CMN, Hardy J, Birks EJ, Khaghani A, Banner NR, Yacoub MH. Clinical recovery from end-stage heart failure using left-ventricular assist device and pharmacological therapy correlates with increased sarcoplasmic reticulum calcium content but not with regression of cellular hypertrophy. *Circulation*. 2004;109(19):2263–5.
378. Rubart M. Two-photon microscopy of cells and tissue. *Circ Res*. 2004;95(12):1154–66.
379. Wang BG, König K, Halhuber KJ. Two-photon microscopy of deep intravital tissues and its merits in clinical research. *J Microsc*. 2010;238(1):1–20.
380. Quinn TA, Camelliti P, Rog-Zielinska EA, Siedlecka U, Poggioli T, O’Toole ET, et al. Electrotonic coupling of excitable and nonexcitable cells in the heart revealed by optogenetics. *Proc Natl Acad Sci U S A*. 2016;1611184114-.
381. Golden HB, Gollapudi D, Gerilechaogetu F, Li J, Cristales RJ, Peng X, et al. Isolation of Cardiac

- Myocytes and Fibroblasts from Neonatal Rat Pups. In: Peng X, Antonyak M, editors. *Cardiovascular Development: Methods and Protocols*. Totowa, NJ: Humana Press; 2012. p. 205–14.
382. Rook MB, Van Ginneken ACG, de Jonge B, el Aoumari A, Gros D, Jongsma HJ, et al. Differences in gap junction channels between cardiac myocytes, fibroblasts, and heterologous pairs. *Am J Physiol*. 1992;263(5 Pt 1):C959-77.
 383. Chilton L, Ohya S, Freed D, George E, Drobic V, Shibukawa Y, et al. K⁺ currents regulate the resting membrane potential, proliferation, and contractile responses in ventricular fibroblasts and myofibroblasts. *AJP Hear Circ Physiol*. 2005;288(6):H2931–9.
 384. Kohl P. Heterogeneous cell coupling in the heart: An electrophysiological role for fibroblasts. *Circ Res*. 2003;93(5):381–3.
 385. MacCannell KA, Bazzazi H, Chilton L, Shibukawa Y, Clark RB, Giles WR. A mathematical model of electrotonic interactions between ventricular myocytes and fibroblasts. *Biophys J*. 2007;92(11):4121–32.
 386. Jacquemet V, Henriquez CS. Modeling cardiac fibroblasts, their interactions with myocytes and their impact on impulse propagation. *Europace*. 2007;9(Suppl 6):vi29–37.
 387. Lee EK, Tran DD, Keung W, Chan P, Wong G, Chan CW, et al. Machine Learning of Human Pluripotent Stem Cell-Derived Engineered Cardiac Tissue Contractility for Automated Drug Classification. *Stem Cell Reports*. 2017;9(5):1560–72.
 388. Li X, Zhang R, Zhao B, Lossin C, Cao Z. Cardiotoxicity screening: a review of rapid-throughput in vitro approaches. *Arch Toxicol*. 2016;90(8):1803–16.
 389. Tixier E, Raphel F, Lombardi D, Gerbeau JF. Composite biomarkers derived from Micro-Electrode Array measurements and computer simulations improve the classification of drug-induced channel block. *Front Physiol*. 2018;8(JAN):1–17.
 390. Lee EK, Kurokawa YK, Tu R, George SC, Khine M. Machine learning plus optical flow: A simple and sensitive method to detect cardioactive drugs. *Sci Rep*. 2015;5(July):1–12.
 391. Duque SI, Orozco-Duque A, Kremen V, Novak D, Tobon C, Bustamante J. Feature subset selection and classification of intracardiac electrograms during atrial fibrillation. *Biomed Signal Process Control*. 2017;38:182–90.
 392. Dempsey GT, Chaudhary KW, Atwater N, Nguyen C, Brown BS, McNeish JD, et al. Cardiotoxicity screening with simultaneous optogenetic pacing, voltage imaging and calcium imaging. *J Pharmacol Toxicol Methods*. 2016;81:240–50.
 393. Ravelli F, Masè M, Cristoforetti A, Marini M, Disertori M. The logical operator map identifies novel candidate markers for critical sites in patients with atrial fibrillation. *Prog Biophys Mol Biol*. 2014;115(2–3):186–97.
 394. Hunter RJ, Diab I, Thomas G, Duncan E, Abrams D, Dhinoja M, et al. Validation of a classification system to grade fractionation in atrial fibrillation and correlation with automated detection systems. *Europace*. 2009;11(12):1587–96.
 395. Ugarte JP, Orozco-Duque A, N CT, Kremen V, Novak D, Saiz J, et al. Dynamic approximate entropy electroanatomic maps detect rotors in a simulated atrial fibrillation model. *PLoS One*. 2014;9(12):1–19.
 396. Lin YJ, Lo MT, Lin C, Chang SL, Lo LW, Hu YF, et al. Prevalence, characteristics, mapping, and catheter ablation of potential rotors in nonparoxysmal atrial fibrillation. *Circ Arrhythmia Electrophysiol*. 2013;6(5):851–8.
 397. Chang SL, Chen YC, Hsu CP, Kao YH, Lin YK, Lin YJ, et al. Electrophysiological characteristics of complex fractionated electrograms and high frequency activity in atrial fibrillation. *Int J Cardiol*. 2013;168(3):2289–99.
 398. Patel SP, Campbell DL. Transient outward potassium current, “I_{to}”, phenotypes in the mammalian left ventricle: underlying molecular, cellular and biophysical mechanisms. *J Physiol*. 2005;569(Pt 1):7–39.
 399. Schoonderwoerd BA, Van Gelder IC, Van Veldhuisen DJ, Van Den Berg MP, Crijns HJGM. Electrical and structural remodeling: Role in the genesis and maintenance of atrial fibrillation. *Prog Cardiovasc Dis*. 2005;48(3):153–68.
 400. Papa P, Falc AX, Suzuki CTN. Supervised Pattern Classification based on Optimum-Path Forest. 2009;(i):1–14.
 401. Freund Y, Schapire RE. Experiments with a New Boosting Algorithm. In: *Machine Learning*:

- Proceedings of the 13th International Conference. 1996.
402. Verheule S, Eckstein J, Linz D, Maesen B, Bidar E, Gharaviri A, et al. Role of endo-epicardial dissociation of electrical activity and transmural conduction in the development of persistent atrial fibrillation. *Prog Biophys Mol Biol*. 2014;115(2–3):173–85.
 403. Narayan SM, Krummen DE, Shivkumar K, Clopton P, Rappel WJ, Miller JM. Treatment of Atrial Fibrillation By The Ablation Of Localized Sources: The Conventional Ablation for Atrial Fibrillation With or Without Focal Impulse and Rotors Modulation (CONFIRM) Trial. *J Am Coll Cardiol*. 2012;60(7):628–36.
 404. Roney CH, Cantwell CD, Bayer JD, Qureshi NA, Lim PB, Tweedy JH, et al. Spatial resolution requirements for accurate identification of driver of Atrial Fibrillation. *Circ Arrhythmia Electrophysiol*. 2017;10(5).
 405. Latchamsetty R, Kocheril AG. Review of dominant frequency analysis in atrial fibrillation. *J Atr Fibrillation*. 2009;1(9):531–4.
 406. Nollo G, Marconcini M, Faes L, Bovolo F, Ravelli F, Bruzzone L. An automatic system for the analysis and classification of human atrial fibrillation patterns from intracardiac electrograms. *IEEE Trans Biomed Eng*. 2008;55(9):2275–85.
 407. Ohler A, Ravens U. Effects of E-4031, almokalant and tedisamil on postrest action potential duration of human papillary muscles. *J Pharmacol Exp Ther*. 1994;270(2):460–5.
 408. Pillekamp F, Reppel M, Brockmeier K, Hescheler J. Impulse propagation in late-stage embryonic and neonatal murine ventricular slices. *J Electrocardiol*. 2006;39(4):16–9.
 409. Martis RJ, Chakraborty C, Ray AK. Wavelet-based Machine Learning Techniques for ECG Signal Analysis. In: *Machine Learning in Healthcare Informatics*. 2014.
 410. Jin Z, Sun Y, Cheng AC. Predicting cardiovascular disease from real-time electrocardiographic monitoring: An adaptive machine learning approach on a cell phone. *Proc 31st Annu Int Conf IEEE Eng Med Biol Soc Eng Futur Biomed EMBC 2009*. 2009;6889–92.
 411. Polat K, Güneş S. Detection of ECG Arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine. *Appl Math Comput*. 2007;186(1):898–906.
 412. Scherschel JA, Rubart M. Cardiovascular Imaging Using Two-Photon Microscopy. *Microsc Microanal*. 2008;14(6):492–506.
 413. Horn MA, Trafford AW. Aging and the cardiac collagen matrix: Novel mediators of fibrotic remodelling. *J Mol Cell Cardiol*. 2016;93:175–85.
 414. Fan D, Takawale A, Lee J, Kassiri Z. Cardiac fibroblasts, fibrosis and extracellular matrix remodeling in heart disease. *Fibrogenes Tissue Repair*. 2012;5(1):1.
 415. Rich L, Whittaker P. Collagen and Picrosirius Red Staining : a Polarized Light Assessment of Fibrillar Hue and Spatial Distribution. *Braz J Morphol Sci*. 2005;22(2):97–104.
 416. Hordof AJ, Edie R, Malm JR, Hoffman BF, Rosen MR. Electrophysiologic properties and response to pharmacologic agents of fibers from diseased human atria. *Circulation*. 1976;54(5):774–9.
 417. Gepstein L, Hayam G, Shpun S, Cohen D, Ben-Haim SA. Atrial linear ablations in pigs: Chronic effects on atrial electrophysiology and pathology. *Circulation*. 1999;100(4):419–26.
 418. Faes L, Nollo G, Antolini R, Gaita F, Ravelli F. A method for quantifying atrial fibrillation organization based on wave-morphology similarity. *IEEE Trans Biomed Eng*. 2002;49(12 I):1504–13.
 419. Scherr D, Dalal D, Cheema A, Cheng A, Henrikson CA, Spragg D, et al. Automated detection and characterization of complex fractionated atrial electrograms in human left atrium during atrial fibrillation. *Hear Rhythm*. 2007;4(8):1013–20.
 420. Lau DH, Maesen B, Zeemering S, Verheule S, Crijns HJ, Schotten U. Stability of complex fractionated atrial electrograms: A systematic review. *J Cardiovasc Electrophysiol*. 2012;23(9):980–7.
 421. Soejima K, Stevenson WG, Sapp JL, Selwyn AP, Couper G, Epstein LM. Endocardial and epicardial radiofrequency ablation of ventricular tachycardia associated with dilated cardiomyopathy: The importance of low-voltage scars. *J Am Coll Cardiol*. 2004;43(10):1834–42.
 422. Anderson RE, Jakobsson JG. Entropy of EEG during anaesthetic induction: A comparative study with propofol or nitrous oxide as sole agent. *Br J Anaesth*. 2004;92(2):167–70.
 423. Du X, Rao N, Qian M, Liu D, Li J, Feng W, et al. A novel method for real-time atrial fibrillation

- detection in electrocardiograms using multiple parameters. *Ann Noninvasive Electrocardiol.* 2014;19(3):217–25.
424. Glass L, Tateno K. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and dRR intervals. *Medical&Biological Eng.* 2001;39:664–71.
 425. Baldi P, Brunak S, Chauvin Y, Andersen C a, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics.* 2000;16(5):412–24.
 426. Jeong J, Gore JC, Peterson BS. Mutual information analysis of the EEG in patients with Alzheimer’s disease. *ClinNeurophysiol.* 2001;112:827–35.
 427. Harrison JL, Jensen HK, Peel SA, Chiribiri A, Grondal AK, Bloch LO, et al. Cardiac magnetic resonance and electroanatomical mapping of acute and chronic atrial ablation injury: A histological validation study. *Eur Heart J.* 2014;35(22):1486–95.
 428. Coronel R, Wilders R, Verkerk AO, Wiegerinck RF, Benoist D, Bernus O. Electrophysiological changes in heart failure and their implications for arrhythmogenesis. *Biochim Biophys Acta - Mol Basis Dis.* 2013;1832(12):2432–41.
 429. Metra M, Teerlink JR. Heart failure. *Lancet.* 2017;390(10106):1981–95.
 430. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. *Nat Rev Cardiol.* 2011;8(1):30–41.
 431. Braunwald E. Heart failure. *JACC Hear Fail.* 2013;1(1):1–20.
 432. Franssen C, Chen S, Unger A, Korkmaz HI, De Keulenaer GW, Tschöpe C, et al. Myocardial Microvascular Inflammatory Endothelial Activation in Heart Failure With Preserved Ejection Fraction. *JACC Hear Fail.* 2016;4(4):312–24.
 433. Ter Maaten JM, Damman K, Verhaar MC, Paulus WJ, Duncker DJ, Cheng C, et al. Connecting heart failure with preserved ejection fraction and renal dysfunction: the role of endothelial dysfunction and inflammation. *Eur J Heart Fail.* 2016;18(6):588–98.
 434. Cleland JG, McGowan J. Heart failure due to ischaemic heart disease: epidemiology, pathophysiology and progression. *J Cardiovasc Pharmacol.* 1999;33 Suppl 3(June 1999):S17-29.
 435. Vaccarino V, Badimon L, Corti R, De Wit C, Dorobantu M, Manfrini O, et al. Presentation, management, and outcomes of ischaemic heart disease in women. *Nat Rev Cardiol.* 2013;10(9):508–18.
 436. Cleland JG, Puri S. How do ACE inhibitors reduce mortality in patients with left ventricular dysfunction with and without heart failure: remodelling, resetting, or sudden death? *Br Hear J.* 1994;72(3 Suppl):S81-6.
 437. McDonagh TA, Morrison CE, Lawrence A, Ford I, Tunstall-Pedoe H, McMurray JJV, et al. Symptomatic and asymptomatic left-ventricular systolic dysfunction in an urban population. *Lancet.* 1997;350(9081):829–33.
 438. Assomull RG, Prasad SK, Lyne J, Smith G, Burman ED, Khan M, et al. Cardiovascular Magnetic Resonance, Fibrosis, and Prognosis in Dilated Cardiomyopathy. *J Am Coll Cardiol.* 2006;48(10):1977–85.
 439. Ferrari P, Bianchi G. The genomics of cardiovascular disorders: Therapeutic implications. *Drugs.* 2000;59(5):1025–42.
 440. Maron BJ, Maron MS. Hypertrophic cardiomyopathy. *Lancet.* 2013;381(9862):242–55.
 441. Klues HG, Maron BJ, Dollar AL, Roberts WC. Diversity of structural mitral valve alterations in hypertrophic cardiomyopathy. *Circulation.* 1992;85(5):1651–60.
 442. Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, et al. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines Developed in Collaboration with the American Ass. *J Am Coll Cardiol.* 2011;58(25):e212–60.
 443. Olivetto I, Maron MS. Hypertrophic Cardiomyopathy. 2017;64(1).
 444. Gruner C, Care M, Siminovitch K, Moravsky G, Wigle ED, Woo A, et al. Sarcomere protein gene mutations in patients with apical hypertrophic cardiomyopathy. *Circ Cardiovasc Genet.* 2011;4(3):288–95.
 445. Cooper LT. Myocarditis. *N Engl J Med.* 2009;360(15):1526–38.
 446. Garner WL, Starling C, Kuiper JJ, Roberts WC. Lymphocytic myocarditis as a cause of fulminant fatal

- heart failure. *Proc (Bayl Univ Med Cent)*. 2006;19(2):122–3.
447. Connolly SJ, Dorian P, Roberts RS, Gent M, Bailin S, Fain ES, et al. Comparison of β -blockers, amiodarone plus β -blockers, or sotalol for prevention of shocks from implantable cardioverter defibrillators - The OPTIC study: A randomized trial. *J Am Med Assoc*. 2006;295(2):165–71.
 448. Nakahara S, Tung R, Ramirez RJ, Michowitz Y, Vaseghi M, Buch E, et al. Characterization of the arrhythmogenic substrate in ischemic and nonischemic cardiomyopathy: implications for catheter ablation of hemodynamically unstable ventricular tachycardia. *J Am Coll Cardiol*. 2011;55(21):2355–65.
 449. Kinoshita O, Tomita T, Hanaoka T, Tsutsui H, Imamura H, Yazaki Y, et al. T-Wave Alternans in Patients with Right Ventricular Tachycardia. *Cardiology*. 2003;100(2):86–92.
 450. Vermeulen JT, Mcguire MA, Opthof T, Coronel R, De Bakker JMT, Klöpping C, et al. Triggered activity and automaticity in ventricular trabeculae of failing human and rabbit hearts. *Cardiovasc Res*. 1994 Oct 1;28(10):1547–54.
 451. Mines G. On circulating excitations in heart muscle and their possible relation to tachycardia and fibrillation. *Trans R Soc Can*. 1914;8:43–52.
 452. Han J, Moe G. Nonuniform Recovery of Excitability in Ventricular Muscle. *Circ Res*. 1964;14:44–60.
 453. Jansen JA, Noorman M, Musa H, Stein M, de Jong S, van der Nagel R, et al. Reduced heterogeneous expression of Cx43 results in decreased Nav1.5 expression and reduced sodium current that accounts for arrhythmia vulnerability in conditional Cx43 knockout mice. *Hear Rhythm*. 2012;9(4):608–9.
 454. Wilders R. Arrhythmogenic right ventricular cardiomyopathy: Considerations from in silico experiments. *Front Physiol*. 2012;3 MAY(May):1–11.
 455. Wiegerinck RF, Verkerk AO, Belterman CN, Van Veen TAB, Baartscheer A, Opthof T, et al. Larger cell size in rabbits with heart failure increases myocardial conduction velocity and QRS duration. *Circulation*. 2006;113(6):806–13.
 456. Wilders R, Wagner MB, Golod DA, Kumar R, Wang YG, Goolsby WN, et al. Effects of anisotropy on the development of cardiac arrhythmias associated with focal activity. *Pflugers Arch Eur J Physiol*. 2000;441(2–3):301–12.
 457. Hoogendijk MG, Potse M, Vinet A, de Bakker JMT, Coronel R. ST segment elevation by current-to-load mismatch: an experimental and computational study. *Hear Rhythm*. 2011 Jan 1;8(1):111–8.
 458. Wang TJ, Larson MG, Levy D, Vasani RS, Leip EP, Wolf PA, et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: The Framingham heart study. *Circulation*. 2003;107(23):2920–5.
 459. Christov I, Gómez-Herrero G, Krasteva V, Jekova I, Gotchev A, Egiazarian K. Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. *Med Eng Phys*. 2006;28(9):876–87.
 460. Owis MI, Abou-Zied AH, Youssef ABM, Kadah YM. Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification. *IEEE Trans Biomed Eng*. 2002;49(7):733–6.
 461. Ng J, Goldberger JJ. Understanding and interpreting dominant frequency analysis of AF electrograms. *J Cardiovasc Electrophysiol*. 2007;18(6):680–5.
 462. Awan SE, Sohail F, Sanfilippo FM, Bennamoun M, Dwivedi G. Machine learning in heart failure. *Curr Opin Cardiol*. 2018;33(2):190–5.
 463. Austin PC, Tu J V., Ho JE, Levy D, Lee DS. Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes. *J Clin Epidemiol*. 2013;66(4):398–407.
 464. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghide M, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation*. 2015;131(3):269–79.
 465. Ashikaga H, Mickelsen SR, Ennis DB, Rodriguez I, Kellman P, Wen H, et al. Electromechanical analysis of infarct border zone in chronic myocardial infarction. *Am J Physiol - Hear Circ Physiol*. 2005;289(3):H1099–105.
 466. Ashihara T, Haraguchi R, Nakazawa K, Namba T, Ikeda T, Nakazawa Y, et al. The role of fibroblasts in complex fractionated electrograms during persistent/permanent atrial fibrillation. *Circ Res*. 2012;110(2):275–84.
 467. Schechter MA, Southerland KW, Feger BJ, Linder D, Ali AA, Njoroge L, et al. An Isolated Working

- Heart System for Large Animal Models. *J Vis Exp*. 2014;(88):4–11.
468. Schuster A, Grünwald I, Chiribiri A, Southworth R, Ishida M, Hay G, et al. An isolated perfused pig heart model for the development, validation and translation of novel cardiovascular magnetic resonance techniques. *J Cardiovasc Magn Reson*. 2010;12:53.
 469. Guo L, Dong Z, Guthrie H. Validation of a guinea pig Langendorff heart model for assessing potential cardiovascular liability of drug candidates. *J Pharmacol Toxicol Methods*. 2009;60(2):130–51.
 470. Martin C a., Matthews GDK, Huang CL-H. Sudden cardiac death and inherited channelopathy: the basic electrophysiology of the myocyte and myocardium in ion channel disease. *Heart*. 2012;98(7):536–43.
 471. Berenfeld O, Mandapati R, Dixit S, Skanes A, Chen J, Mansour M, Jalife J. Spatially distributed dominant excitation frequencies reveal hidden organization in atrial fibrillation in the Langendorff-perfused sheep heart. *J Cardiovasc Electrophysiol*. 2000; 11(8):869-79
 472. Kui P, Orosz S, Takacs H, Sarusi A, Csik N, Rarosi F, Cseko C, Varro A, Papp J, Forster T, Farkas A, Farkas A. New in vitro model for proarrhythmia safety screening: IKs inhibition potentiates the QTc prolonging effect of IKr inhibitors in isolated guinea pig hearts. *J Pharmacol Toxicol Meth* 2016;80:26-34
 473. M. Nair. ECG feature extraction using time-frequency analysis. *Innov Comp Sci & Soft Eng* 2010;461-466
 474. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. p.181-184, 1st ed. 2013, Corr. 7th printing 2017
 475. Hodgkin AL & Huxley AFA. Quantitative description of membrane current and its applications to conduction and excitation in nerve. *J Physiol* 1952; 117, 500-544
 476. Unger LA, Oesterlein TG, Loewe A, Dossel O. Noise quantification and noise reduction for unipolar and bipolar electrograms. *CinC* 2019; 46, 288
 477. Tian X, Li Y, Zhou H, Li X, Chen L, Zhang X. Electrocardiogram signal denoising using extreme-point symmetric mode decomposition and nonlocal means. *Sensors* 2016; 16, 1584
 478. Rohr S, Kucera JP, Kleber, AG. Slow conduction in cardiac tissue I: Effects of a reduction of excitability versus a reduction of electrical coupling on microconduction. *Circ Res* 1998; 83(8), 781-794
 479. Jabr RI, Hatch FS, Salvage SC, Orłowski A, Lampe PD, Fry CH. Regulation of gap junction by calcineurin through Cx43 phosphorylation: Implications for action potential conduction. *Pflugers Arch*. 2016; 468(11-12), 1945-1955

Appendix A

Analysis of electrogram morphology modified
by individual functional and structural
modulations

I. The effects of Na⁺ channel blockade on EGM morphology

Following the administration of 50 μ M lidocaine in 4 MEA plates of MO monolayers, it was found from the automated analysis of data obtained before and after the Na⁺ channel blockade (Control: n=134 EGMs; Lidocaine: n=120 EGMs) that 16 out of 34 EGM features significantly changed (Figure A.1A). Even though the dual modality performance showed that EGM duration increased following the Na⁺ channel blockade, this feature did not significantly change when a larger number of EGMs were analysed (control: 4.05 \pm 0.3ms, lidocaine: 4.15 \pm 0.87ms, p=0.17). The average EGM amplitude was reduced (control: 502.9 \pm 19.2 μ V; lidocaine: 427 \pm 16.5 μ V; p=0.002). In total, 9 time-domain, 3 frequency-domain and 4 time-frequency features were significantly modified due to the lidocaine activity. The general EGM morphology profile (Figure A.1B) showed that a large number of features seemed to be modified, but in fact apart from the 16 features previously mentioned, the rest did not significantly change.

The automated analysis of EGM morphology for data obtained during MEA experiments carried out using NC monolayers showed that 11 features were significantly modified following the addition of 1 μ M lidocaine (Figure A.2A). Fractionation index (control: 1 \pm 0.04; lidocaine: 1.9 \pm 0.26; p<0.0001) and RR interval (control: 1027 \pm 7.8ms; lidocaine: 1011 \pm 5.7ms; p<0.0001) were the only time-domain features that were affected by Na⁺ channel blockade. The rest were time-frequency and information theory-based features. Moreover, all features increased following the Na⁺ channel blockade and none of them decreased (Figure A.2B). Regarding the co-cultures with +20% cultured fibroblasts, 13 features significantly changed after the addition of 1 μ M lidocaine (Figure A.3A). Shannon entropy increased by 460% compared to the control data (p<0.0001) and the median of PSD estimates decreased the most by 91.7% (control: 44.65 \pm 11.1; lidocaine: 3.7 \pm 1.4; p<0.0001). No significant modifications occurred on time-domain features. Twenty-one features were significantly modified following the administration of 10 μ M lidocaine in co-cultures of NRVM with +40% cultured fibroblasts (Figure A.4A). However, there were no dramatic changes in EGM morphology in overall, since 21 features changed by <10% compared to the control EGM morphology (Figure A.4B). Despite this fact, some of these changes were significant, such as in EGM duration (control: 3.5 \pm 0.13ms; lidocaine: 3.27 \pm 0.14ms; change: -6.73%; p=0.45) and logarithmic energy entropy (control: 7.7*10⁶ \pm 43,180; lidocaine: 7.96*10⁶ \pm 37,821; change: +3.28%; p<0.0001).

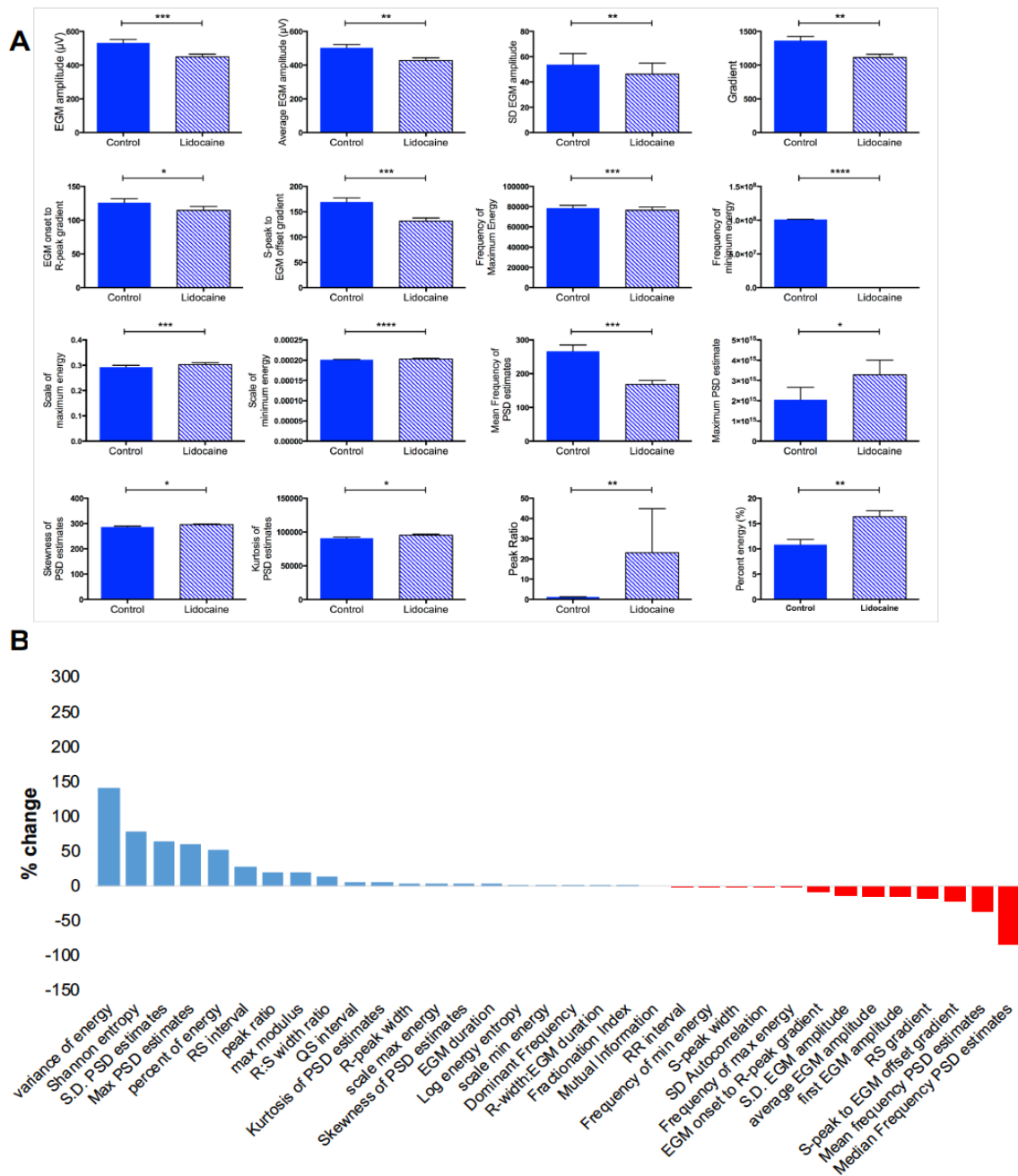


Figure A.1: EGM morphology modifications due to Na⁺ channel blockade when 50µM lidocaine were administered on NRVM cultures. (A) Features significantly changed due to lidocaine activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. (B) General profile of the EGM morphology following Na⁺ channel blockade as compared to the average baseline data. (Control: n = 134 EGMs; Lidocaine: n=120 EGMs).

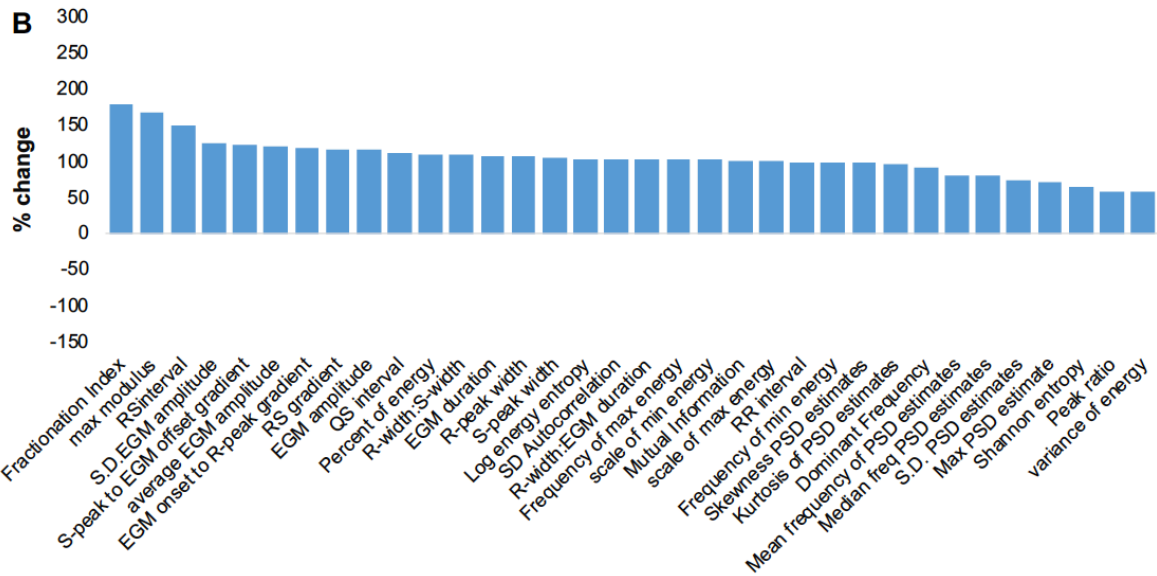
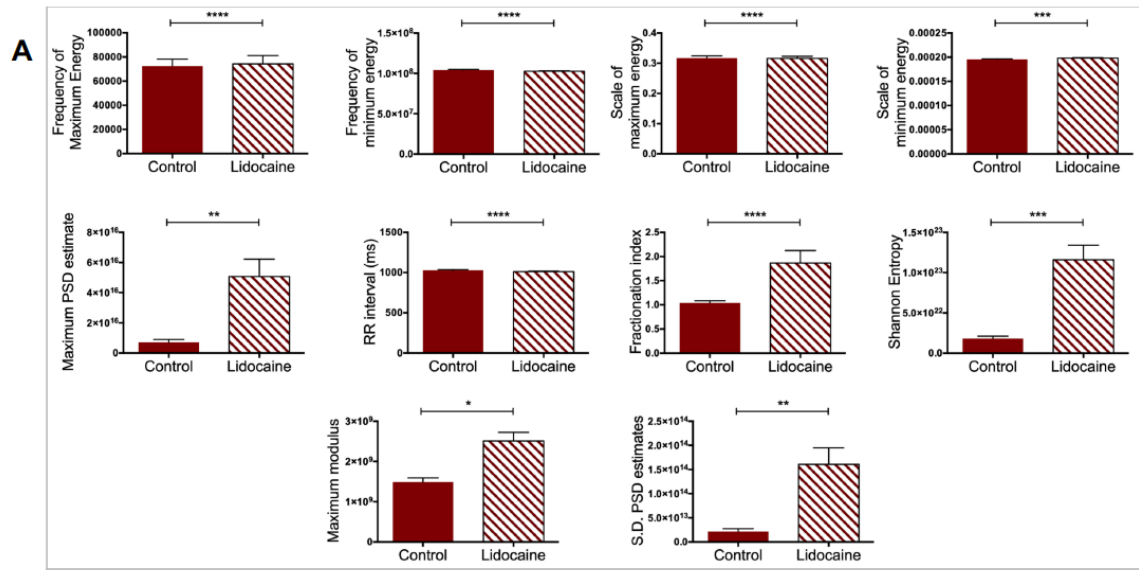
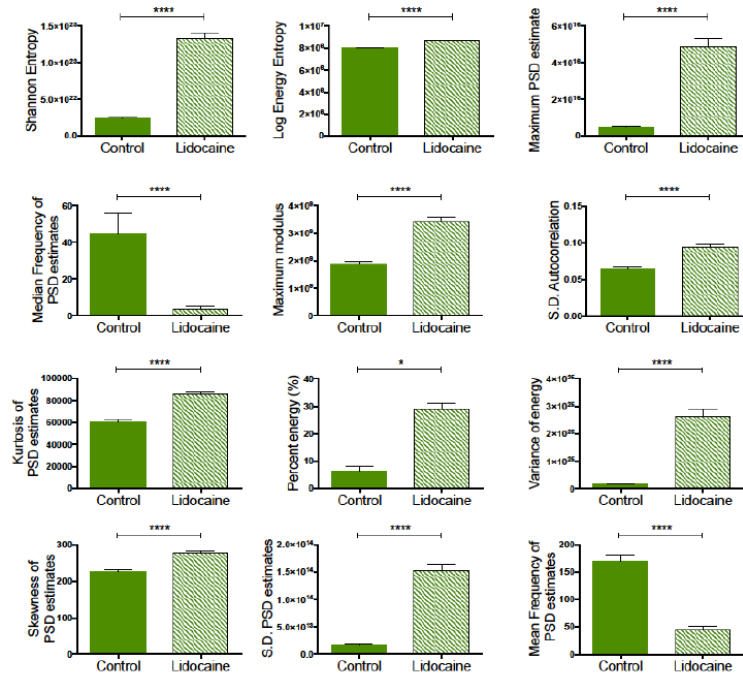


Figure A.2: EGM morphology modifications in NC cell cultures, due to the Na⁺ channel blockade using 1µM lidocaine. (A) Features significantly changed due to lidocaine activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. (B) Quantification of EGM feature change compared to the average control EGM morphology. (Control: n = 71 EGMs; Lidocaine: n = 83 EGMs).

A



B

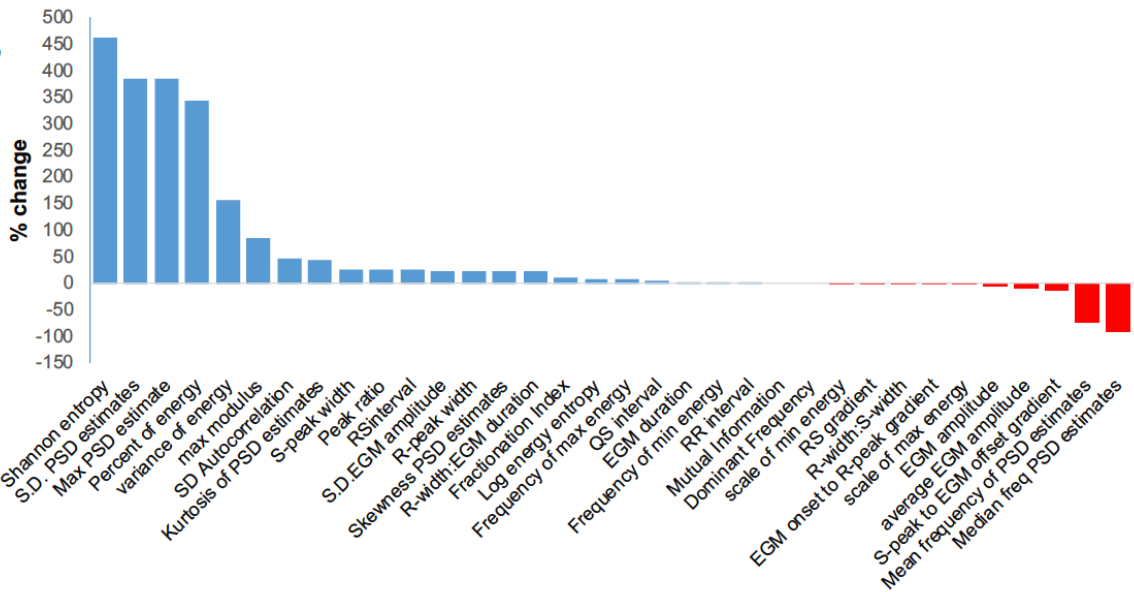


Figure A.3: The effects of Na⁺ channel blockade using 1μM lidocaine on EGM morphology recorded in co-cultures of NRVMs and 20% cultured fibroblasts. (A) Features significantly changed due to lidocaine activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. (B) Positive and negative change of EGM features compared to the average control EGM morphology. (Control: n=64 EGMs; Lidocaine: n=53 EGMs).

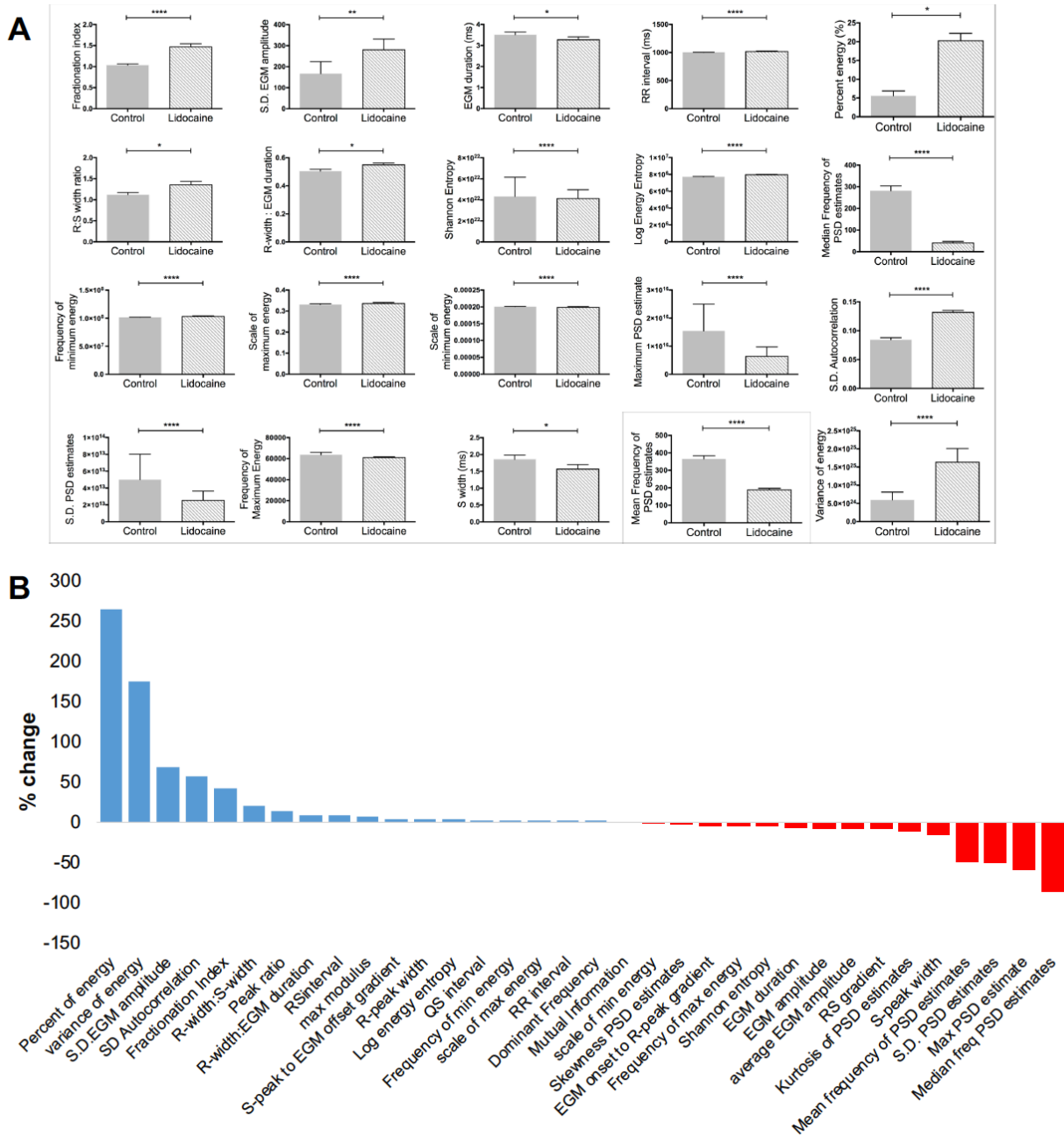


Figure A.4: EGM morphology modifications due to Na⁺ channel blockade when 10 μ M lidocaine were administered on co-cultures of NRVM with 40% cultured fibroblasts. (A) Features significantly changed due to lidocaine activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) General profile of the EGM morphology following Na⁺ channel blockade as compared to the average baseline data. (Control: n = 82 EGMs; Lidocaine: n=68 EGMs).

II. EGM morphology manifestations of transient outward K^+ current blockade

Data obtained from 3 MEAs with MO cell monolayers before and after the I_{to} blockade using 750 μ M 4-AP were analysed using the MATLAB script for automated feature extraction (Control: n=117 EGMs; 4-AP: n=72 EGMs). It was found that 25 features significantly changed due to the I_{to} blockade (Figure A.5A). These were features belonging to all categories described in Chapter 3, i.e. time-domain, time-frequency, frequency-domain, non-linear dynamics/information theory-based feature. Twenty features increased, 14 features decreased and one, the R-peak width/EGM duration ratio) remained stable (Figure A.5B).

EGM data collected from 3 MEA dishes of NC cell monolayers were analysed (control: n=105 EGMs; 4-AP: n=117 EGMs). The addition of 750 μ M 4-AP was followed by modifications on 20 EGM features (Figure A.6A). These included 10 time-domain features, 7 time-frequency, 2 non-linear dynamics and 1 frequency-domain features. The RR interval was significantly reduced under these conditions (control: 998.8 \pm 5.7ms; 4-AP: 834.9 \pm 29ms; $p < 0.0001$). The EGM alternans was also enhanced, as the S.D. of EGM amplitude was enhanced (control: 117 \pm 20; 4-AP: 204.9 \pm 19; $p < 0.0001$).

The data related to EGM morphology change after the addition of 300 μ M 4-AP on a +20% FB cell culture should be criticised with some precaution, as they were collected from only one cell culture (control: 29 EGMs; 4-AP: 25 EGMs; n=1). In that case 21 features changed (Figure A.7A). It can be observed that, although the EGM morphology changed, there were no dramatic modifications like the ones that took place in myocyte only and NC cell monolayers following the I_{to} blockade. All changes varied between +61% - -88% (Figure A.7B). As opposed to the previous case, the RR interval was increased by 7.6% (control: 999.9 \pm 0.16ms; 4-AP: 1276 \pm 73.4ms; $p < 0.0001$). Fractionation index was also significantly increased (control: 1.2 \pm 0.15; 4-AP: 3.3 \pm 0.4; $p < 0.0001$).

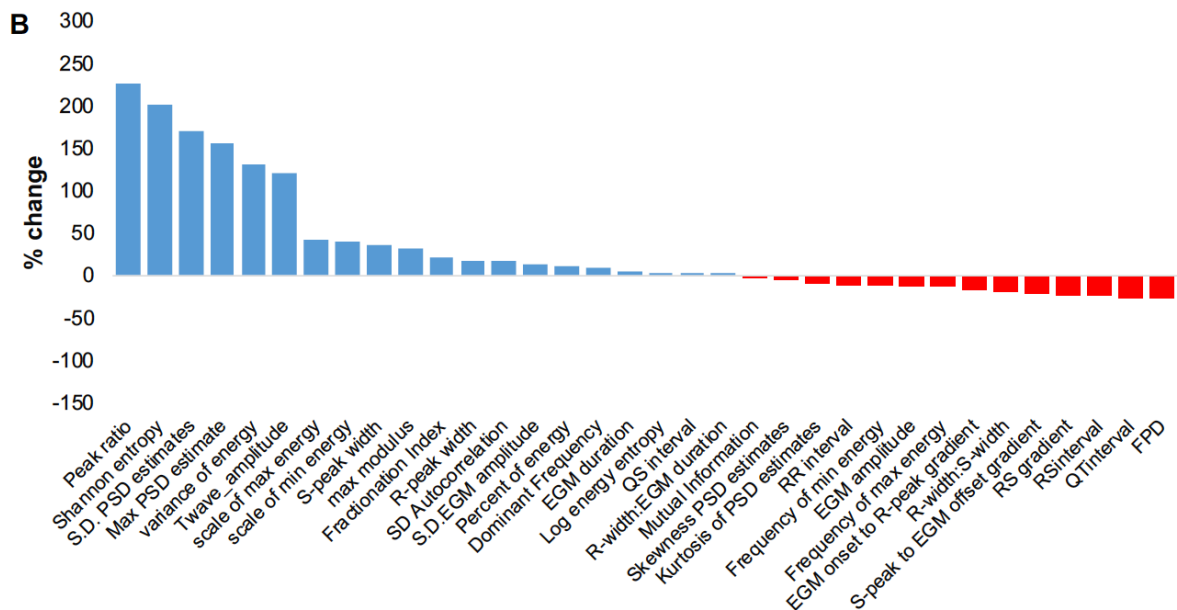
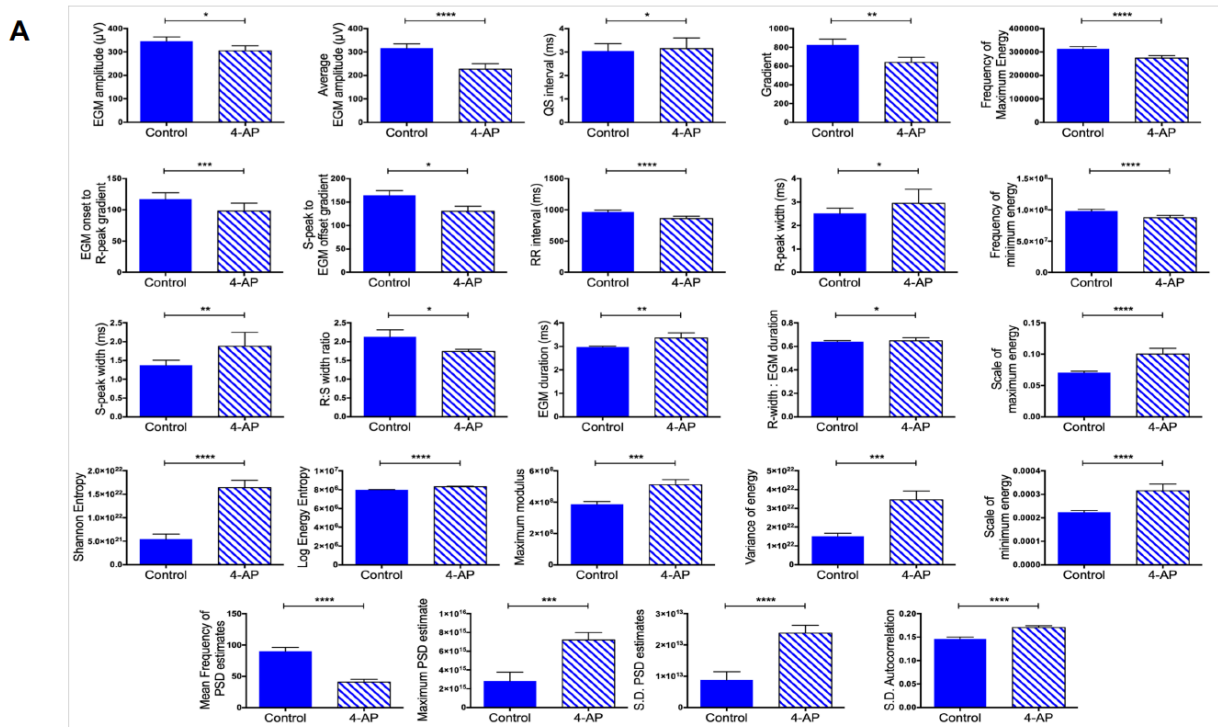
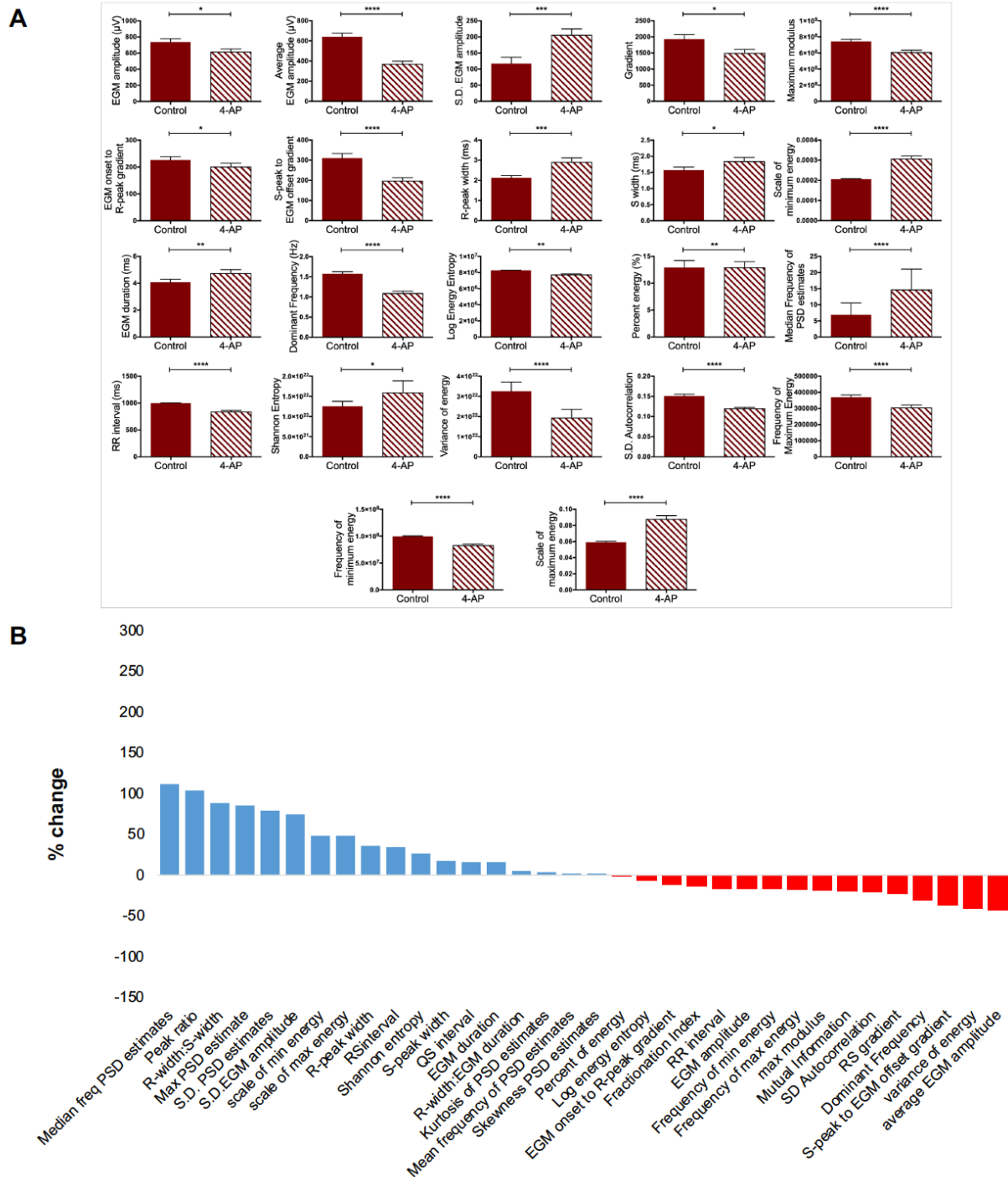
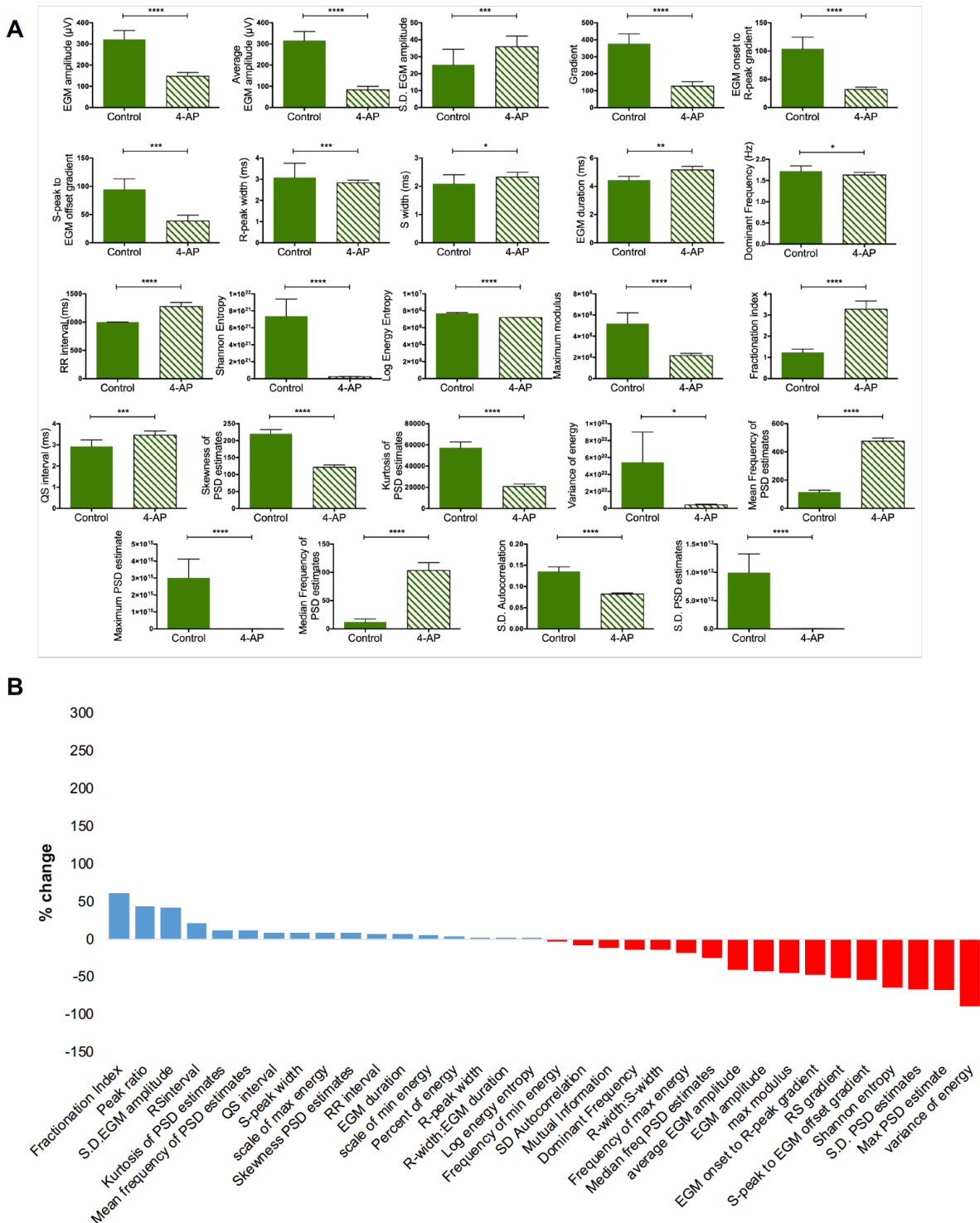


Figure A.5: EGM morphology modifications due to the I_o blockade when $750\mu\text{M}$ 4-AP were administered on NRVM cultures. (A) Features significantly changed due to 4-AP activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) General profile of the EGM morphology following I_o blockade as compared to the average baseline data. (Control: $n = 117$ EGMs; 4-AP: $n=72$ EGMs).





III. The effects of L-type Ca^{2+} current blockade on EGM morphology

EGMs collected from three MEA dishes of MO monolayers during dual modality performance experiments were analysed before and after the administration of IC_{50} of nifedipine (control: 86 EGMs; nifedipine: 77 EGMs). After the automated EGM morphology analysis for feature extraction, it was found that 15 features were significantly modified (Figure A.8A). These included 5 time-domain, 3 information theory, 2 non-linear dynamics and 5 time-frequency features.

The analysis of EGMs obtained from three MEAs of NC cell cultures showed that only 9 features were significantly modified due to the blockade of L-type Ca^{2+} channels using 100nM nifedipine on NC cultures (Figure A.9A). Shannon entropy had the largest increase among all features, as it increased by 300% following I_{CaL} blockade (control: $6.56\text{e}+021 \pm 3.9\text{e}+021$; nifedipine: $2.63\text{e}+022 \pm 1.8\text{e}+022$; $p = 0.014$). There were no dramatic changes of time-domain features in overall. Variance of energy decreased the most (-75.4%), but that was not a significant change due to the large variability of data values. The administration of 300nM nifedipine on four NRVM/+20% fibroblast co-cultures led to the modification of 18 features (Figure A.10A). Moreover, data obtained from two NRVM/+40% fibroblast co-cultures before and after the I_{CaL} blockade using 300nM nifedipine were analysed. Nine features significantly changed (Figure A.11A). The comparison to the average control EGM morphology showed a limited range of mean feature change (+7% - -10%) for 20 out of 35 features (Figure A.11B).

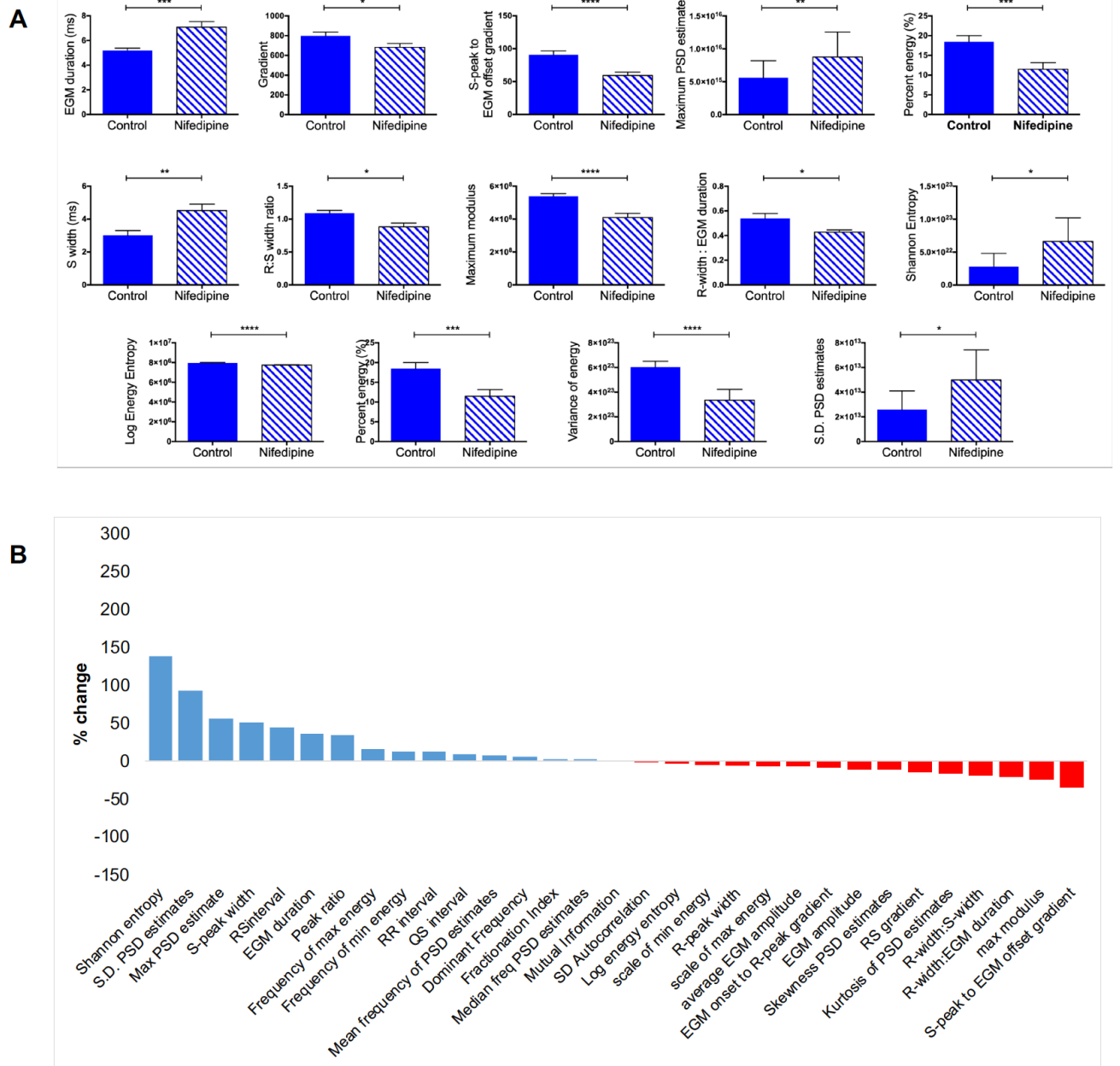


Figure A.8: EGM morphology modifications due to the I_{CaL} blockade when 772.2nM nifedipine were administered on MO cell monolayers. (A) Features significantly changed due to nifedipine activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) General profile of the EGM morphology following I_{CaL} blockade as compared to the average control EGM morphology. (Control: n = 86 EGMs; Nifedipine: n =77 EGMs)

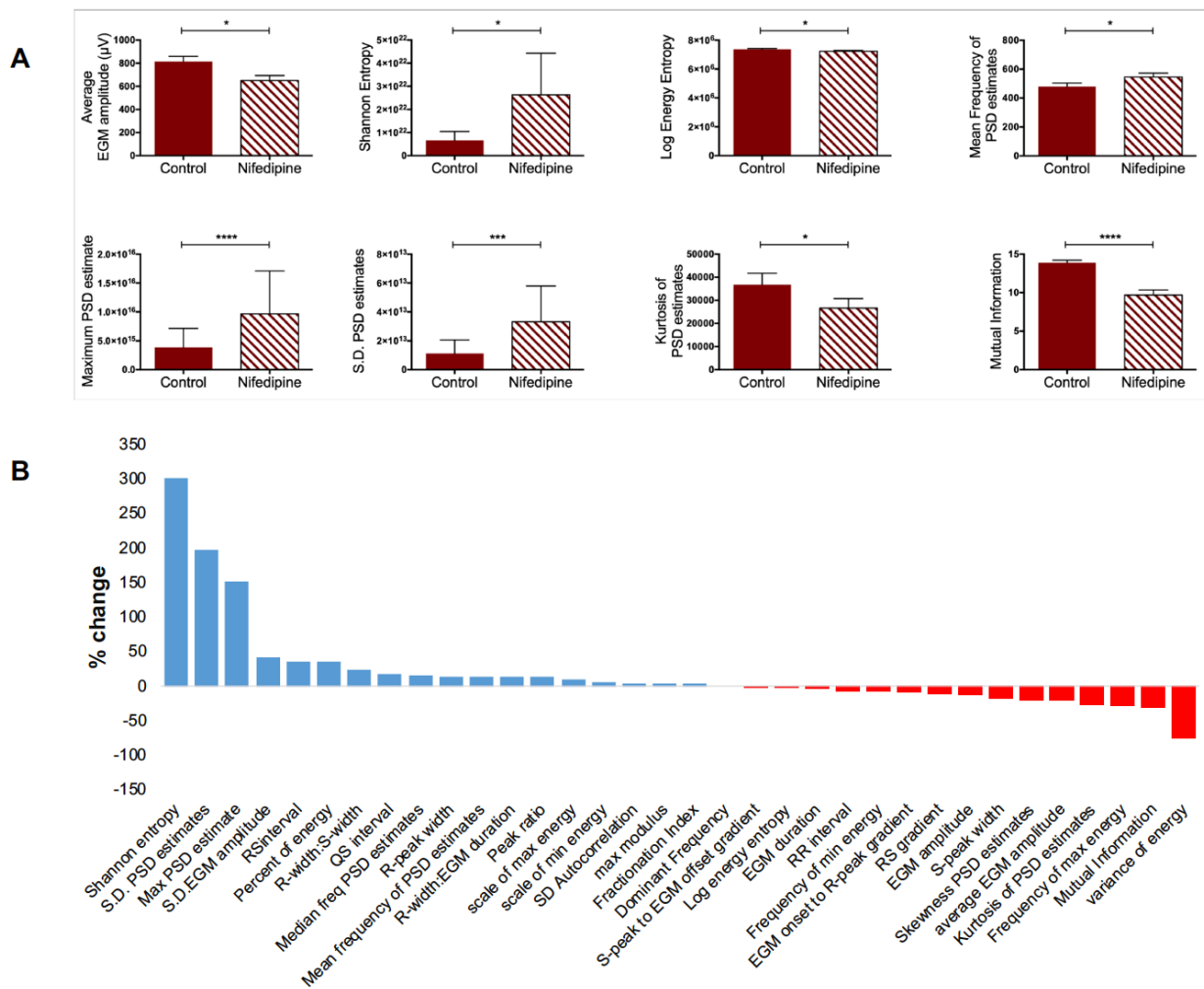
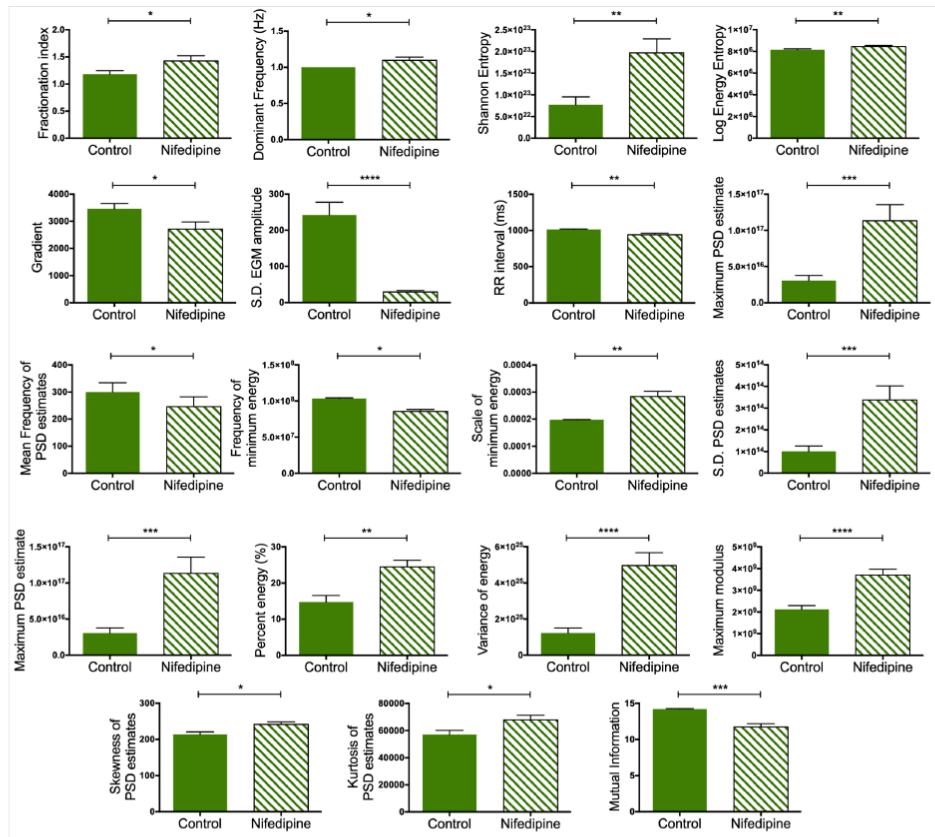


Figure A.9: EGM morphology modifications due to the I_{CaL} blockade when 100nM nifedipine were administered on NC monolayers. (A) Significant feature changes were observed due to nifedipine activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. (B) Quantification of EGM morphology changes following I_{CaL} blockade as compared to the average control EGM morphology. (Control: $n = 55$ EGMs; Nifedipine: $n = 65$ EGMs).

A



B

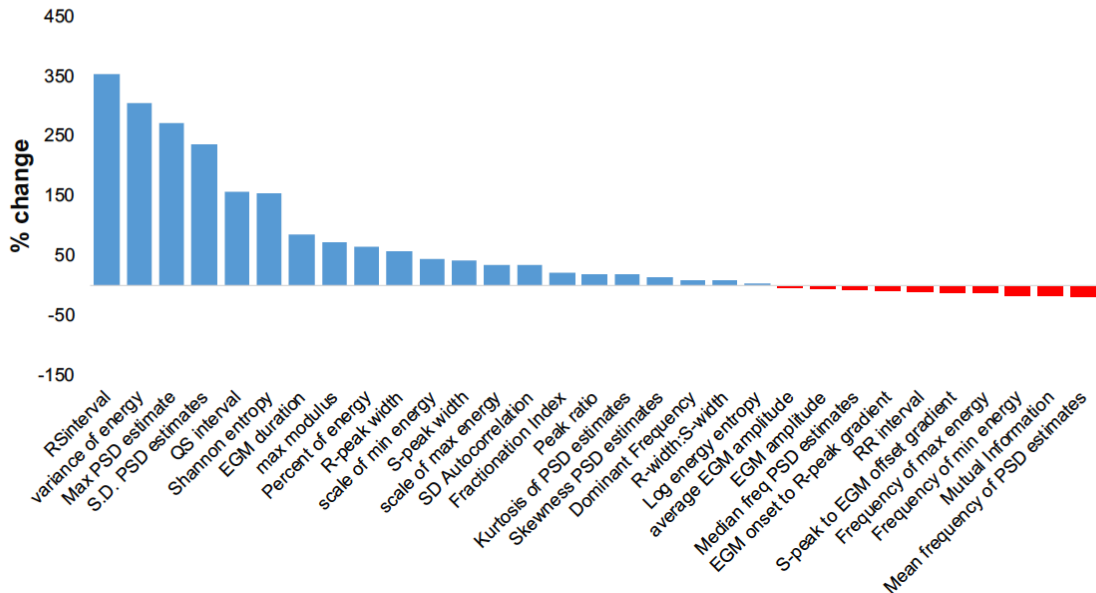


Figure A.10: EGM morphology modifications due to the I_{CaL} blockade when 300nM nifedipine were administered on co-cultures of NRVM and 20% cultured fibroblasts. (A) Features significantly changed due to nifedipine activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) Quantification of EGM morphology changes following I_{CaL} blockade as compared to the average control EGM morphology. (Control: n = 84 EGMs; Nifedipine: n =75 EGMs).

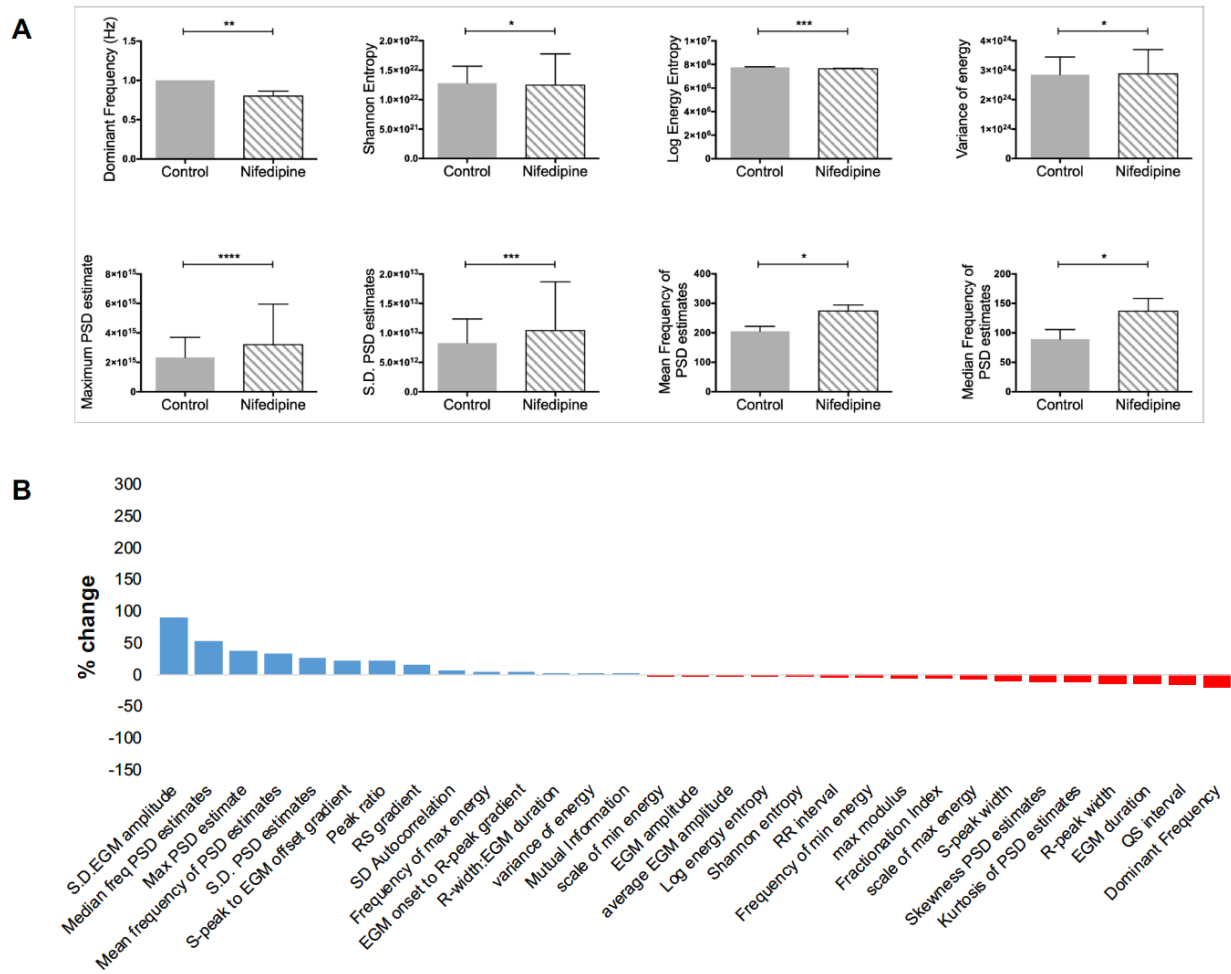


Figure A.11: Modifications on EGM morphology due to the I_{CaL} blockade when 300nM nifedipine were administered on co-cultures of NRVM and 40% cultured fibroblasts. (A) Features significantly changed due to nifedipine activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) Quantification of EGM morphology changes following I_{CaL} blockade as compared to the average control EGM morphology. (Control: n = 39 EGMs; Nifedipine: n =38 EGMs).

IV. I_{Kr} -related EGM morphology modifications

The analysis of EGM data obtained from three NRVM monolayers (Control: n=83 EGMs; E-4031: n=73 EGMs) before and after the blockade of I_{Kr} current using 300nM E-4031 showed significant changes in 23 features (Figure A.12A). These included 11 time-domain, 2 non-linear, 2 information theory, 3 frequency-domain and 5 time-frequency features. No particular preference was shown for any of the feature categories. Even though some of the features changed dramatically when compared to the average control EGM morphology, i.e. S.D. EGM amplitude, maximum PSD estimate, EGM onset-to-R peak gradient, these modifications were not significant, because the data values were too variable (Figure A.12B).

The automated EGM morphology analysis using data obtained from NC monolayers showed that 16 features were modified due to the I_{Kr} blockade using 300nM E-4031 (Figure A.13A). As part of these changes, the EGM amplitude was reduced (control: $2279 \pm 204.3 \mu\text{V}$; E-4031: $1270 \pm 153 \mu\text{V}$; $p < 0.0001$), the RS interval was significantly prolonged (control: $0.79 \pm 0.19 \text{ms}$; E-4031: $0.99 \pm 0.2 \text{ms}$; $p = 0.0002$) and the RR interval was decreased (control: $1000 \pm 0.02 \text{ms}$; E-4031: $979 \pm 15 \text{ms}$; $p < 0.0001$). Fibrillation was also observed in some of the studied cases, as there were non-uniform complexes of increased frequency EGMs amongst uniform EGM complexes. In overall, the majority of features was reduced after the I_{Kr} blockade, even though some of these changes were not significant (Figure A.13B).

Twelve features changed in EGMs recorded from NRVM/+20% fibroblast co-cultures after the I_{Kr} blockade using 100nM E-4031 (Figure A.14A). The average EGM amplitude was increased in this case (control: $885.3 \pm 68 \mu\text{V}$; E-4031: $1034 \pm 70 \mu\text{V}$; $p = 0.03$), as opposed to the NC cultures, and the RR interval, the RS interval and fractionation did not change. There were 19 features which changed by $\pm 5\%$ compared to the control data (Figure A.14B).

The I_{Kr} blockade using $1 \mu\text{M}$ E-4031 on NRVM/+40% fibroblast co-cultures caused significant modifications on 17 features (Figure A.15A). Among these changes, the average EGM amplitude was found to be reduced after the I_{Kr} blockade (control: $257 \pm 28 \mu\text{V}$; E-4031: $208 \pm 39 \mu\text{V}$; $p = 0.03$), the S.D. of EGM amplitude was remarkably increased (control: 17.6 ± 3 ; E-4031: 65.3 ± 9.6 ; $p < 0.0001$) and the RR interval was prolonged (control: $999.9 \pm 0.08 \text{ms}$; E-4031: $1341 \pm 62 \text{ms}$; $p = 0.003$).

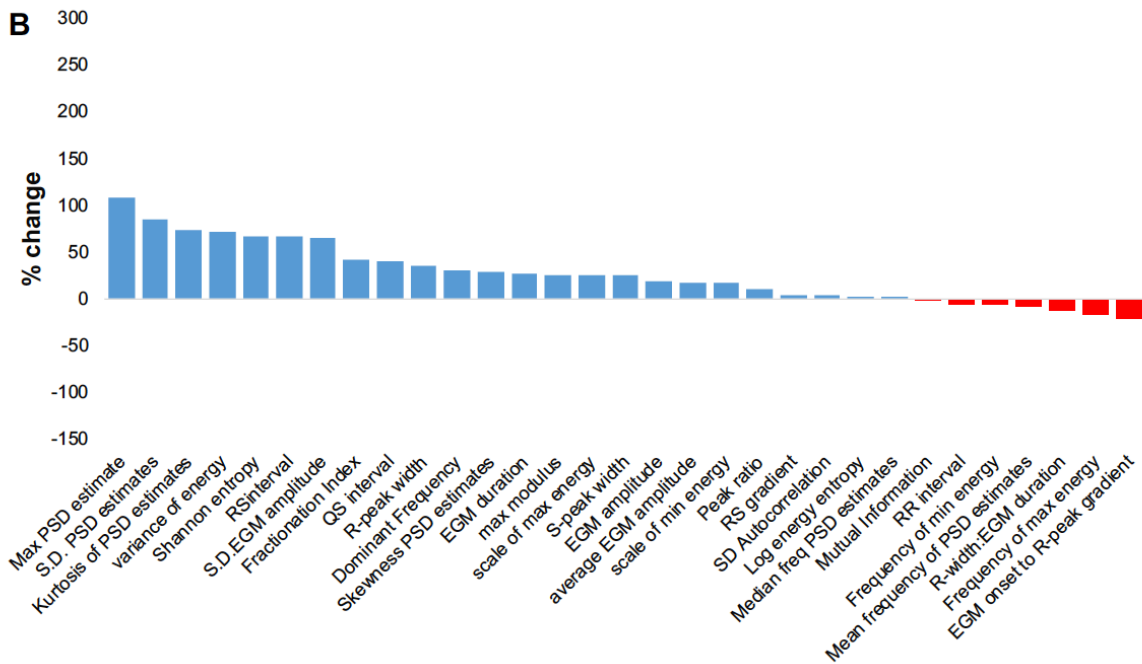
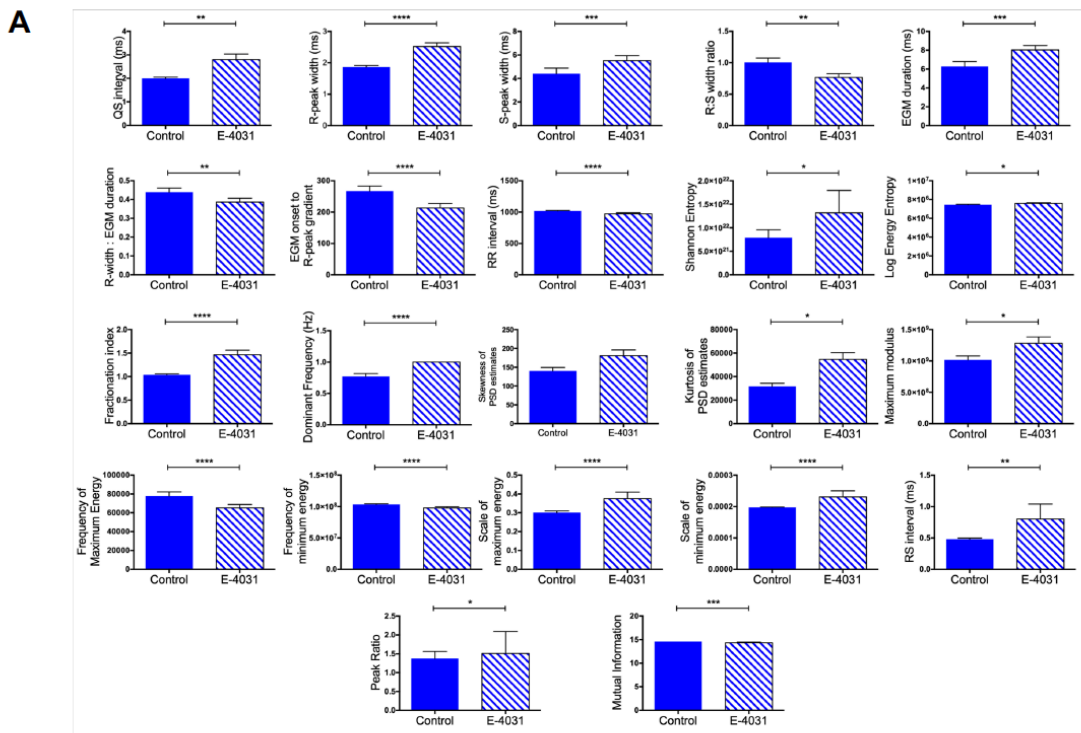


Figure A.12: EGM morphology changes due to the I_{Kr} blockade when 300nM E-4031 were administered on NRVM monolayers. (A) Significant EGM feature changes were observed due to the E-4031 activity. Mann-Whitney test analysis. (B) Quantification of EGM morphology changes following I_{Kr} blockade on NRVM cultures as compared to the baseline EGM morphology. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (Control: n = 83 EGMs; E-4031: n =73 EGMs).

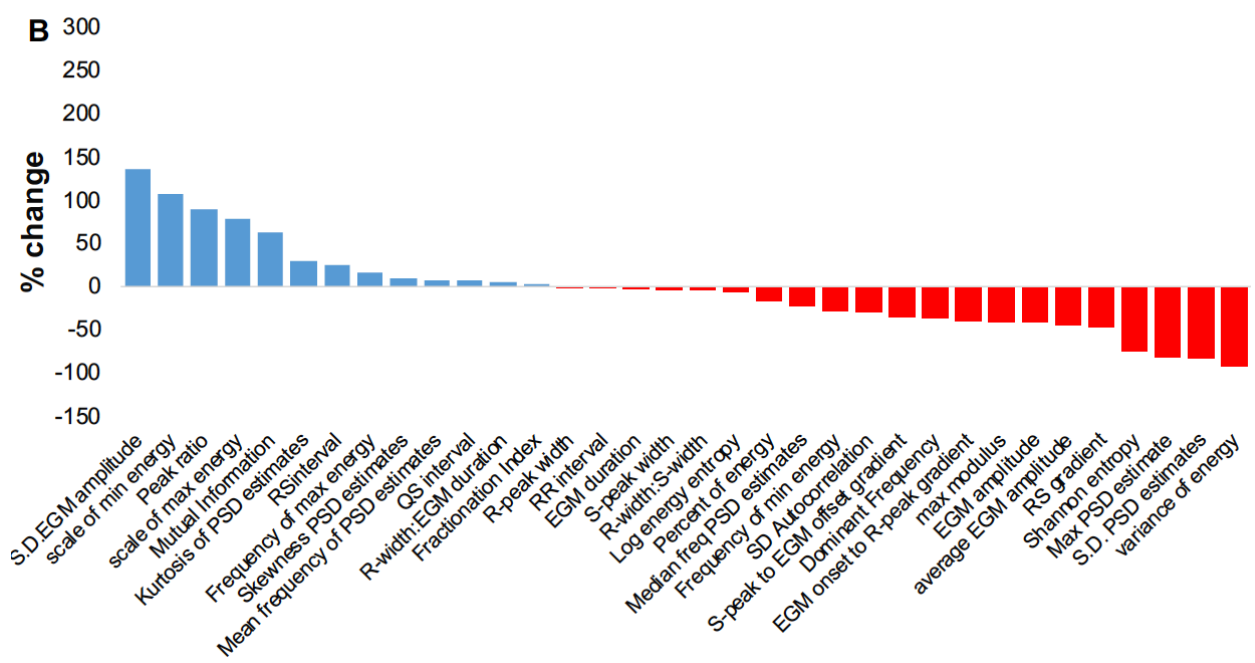
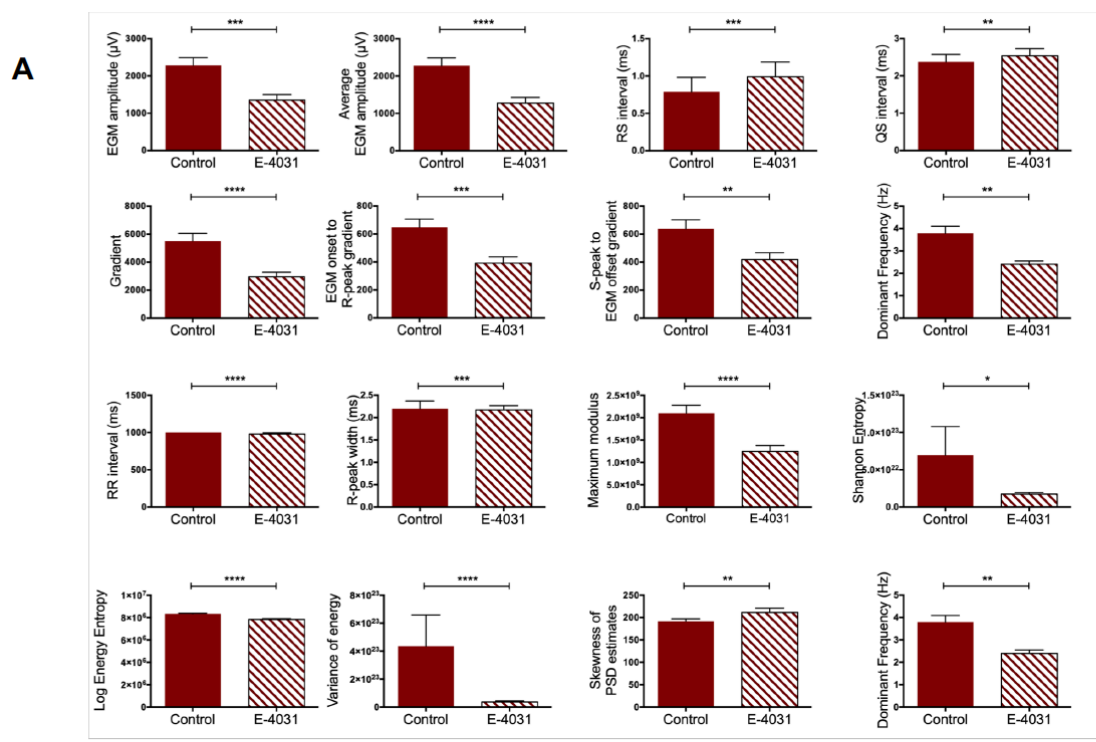


Figure A.13: EGM morphology modifications due to the I_{Kr} blockade when 300nM E-4031 were administered on NC cell monolayers. (A) Features significantly changed due to E-4031 activity. Mann-Whitney test analysis. All bar charts represent mean ± SEM; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. (B) Quantification of EGM morphology changes following I_{Kr} blockade as compared to the average control EGM morphology. (Control: n = 47 EGMs; E-4031: n = 73 EGMs).

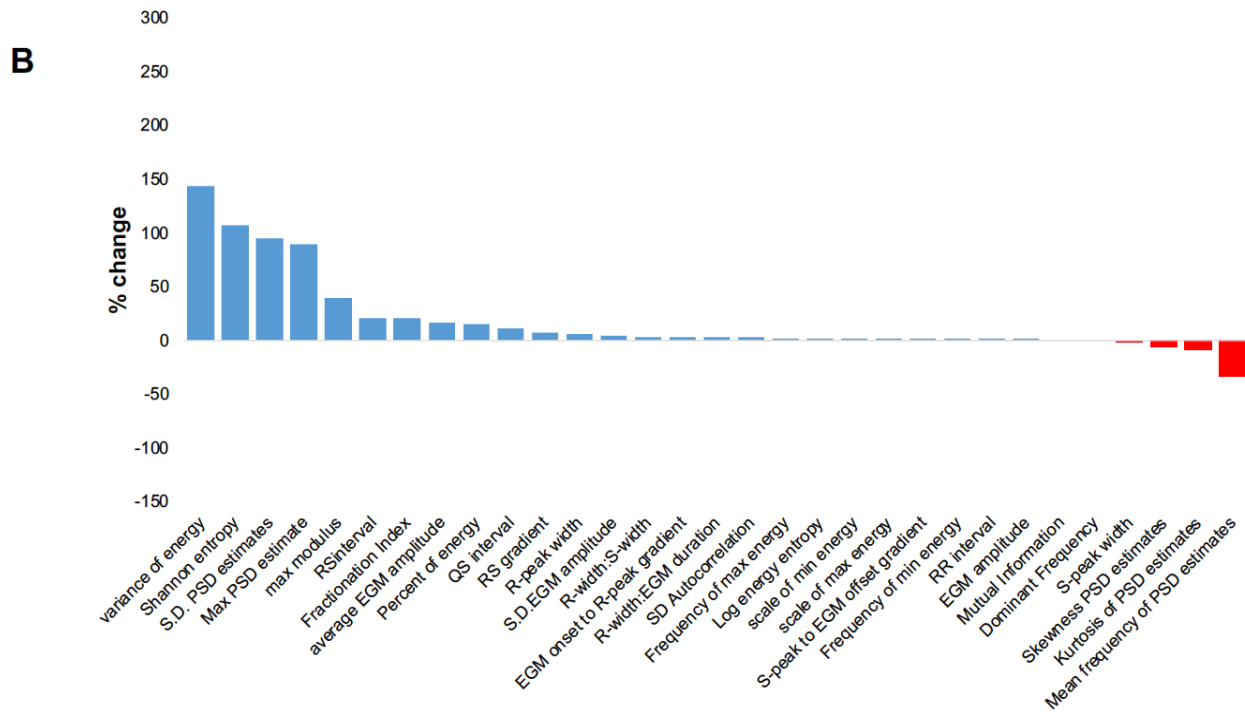
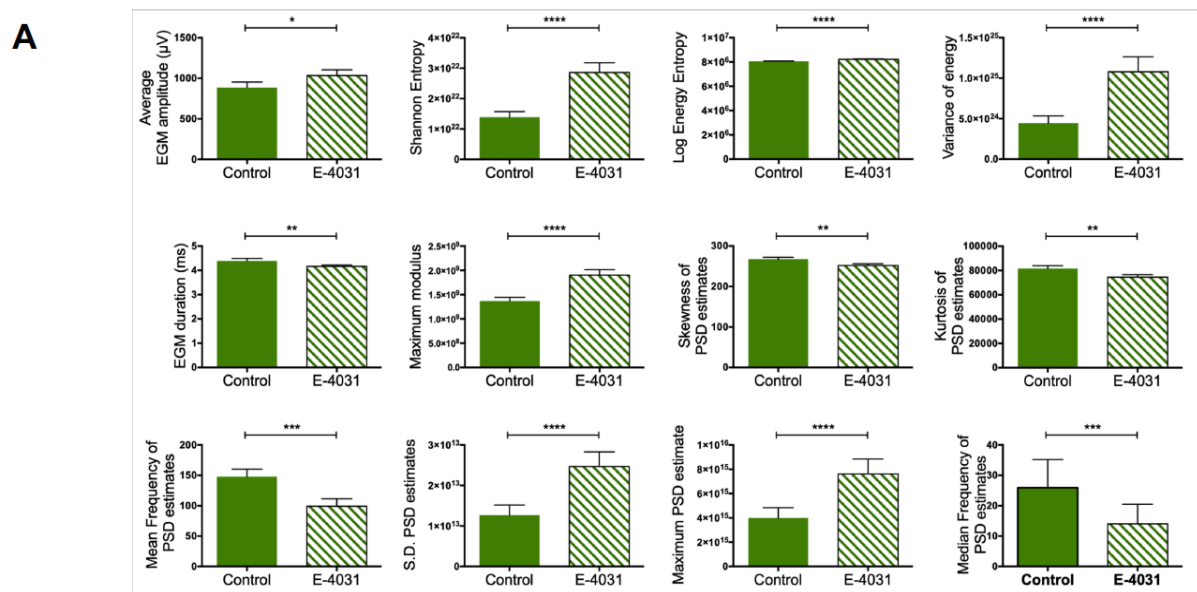


Figure A.14: EGM morphology modifications due to the I_{Kr} blockade when 100nM E-4031 were administered on NRVM/20% fibroblast co-cultures. (A) Features significantly changed due to E-4031 activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$. (B) Quantification of EGM morphology changes following I_{Kr} blockade as compared to the average control EGM morphology. (Control: $n = 47$ EGMs; E-4031: $n=73$ EGMs).

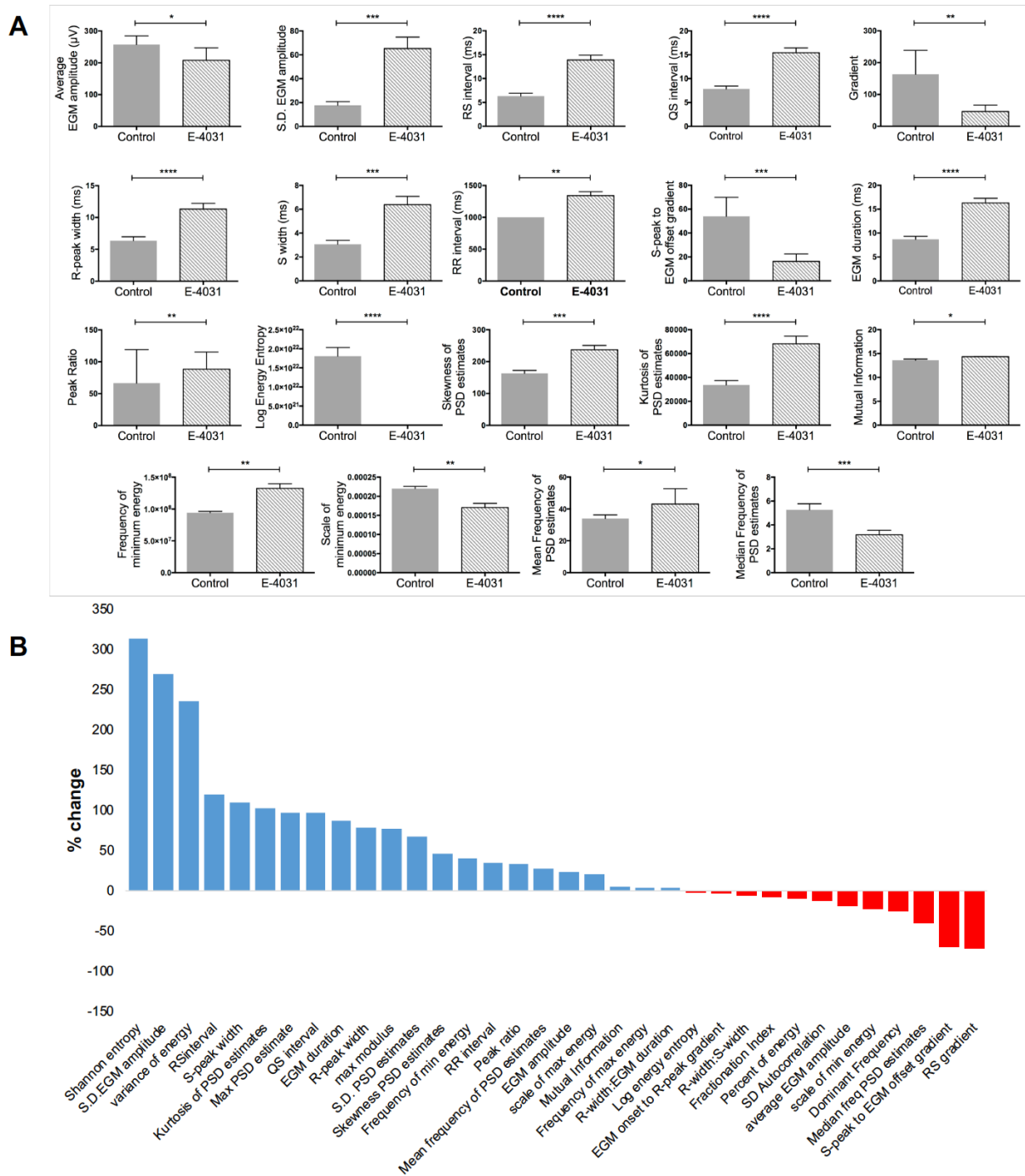


Figure A.15: Modifications on EGM morphology due to the I_{Kr} blockade when 1μM E-4031 were administered on co-cultures of NRVM and 40% cultured fibroblasts. (A) Features significantly changed due to E-4031 activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. (B) Quantification of EGM morphology changes following I_{Kr} blockade as compared to the average control EGM morphology. (Control: $n = 36$ EGMs; E-4031: $n = 32$ EGMs).

V. The effects of I_{Ks} blockade on EGM morphology

The analysis of EGMs collected from myocyte only monolayers (Control: n=27 EGMs; HMR-1556: n=23 EGMs) before and after the administration of $1\mu\text{M}$ HMR-1556 showed that 20 features significantly changed after the I_{Ks} blockade (Figure A.16A). There was a number of features (n = 17) with a limited level of change at $\pm 10\%$, but many of them changed more significantly after the I_{Ks} blockade, i.e. EGM duration, RR interval, scale of maximum energy, compared to others, i.e. R-width:S-width ratio, average EGM amplitude.

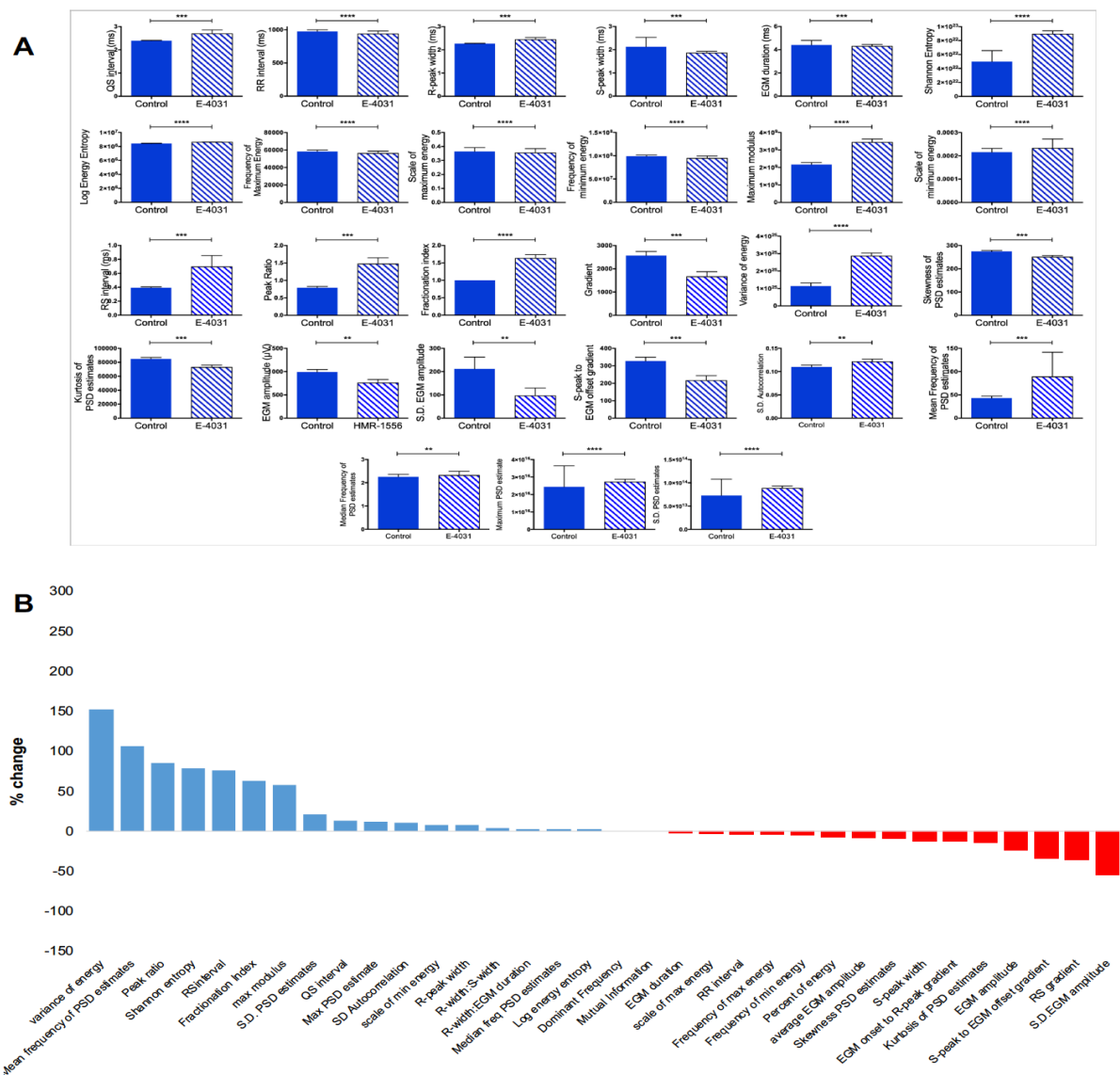


Figure A.16: EGM morphology modifications due to the I_{Ks} blockade when $1\mu\text{M}$ HMR-1556 was administered on myocyte only monolayers. (A) Features significantly changed due to HMR-1556 activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$. (B) Quantification of EGM morphology changes following I_{Ks} blockade as compared to the average control EGM morphology. (Control: n = 27 EGMs; HMR-1556: n=23 EGMs).

The EGM analysis showed that 24 features were affected by I_{Ks} blockade in NC cultures with the majority of them being decreased (Figure A.17A). These data were obtained using $1\mu\text{M}$ HMR-1556. The situation was similar with EGMs obtained from NRVM/+20% fibroblast co-cultures where 18 features were influenced by the administration of 300nM HMR-1556 (Figure A.18A).

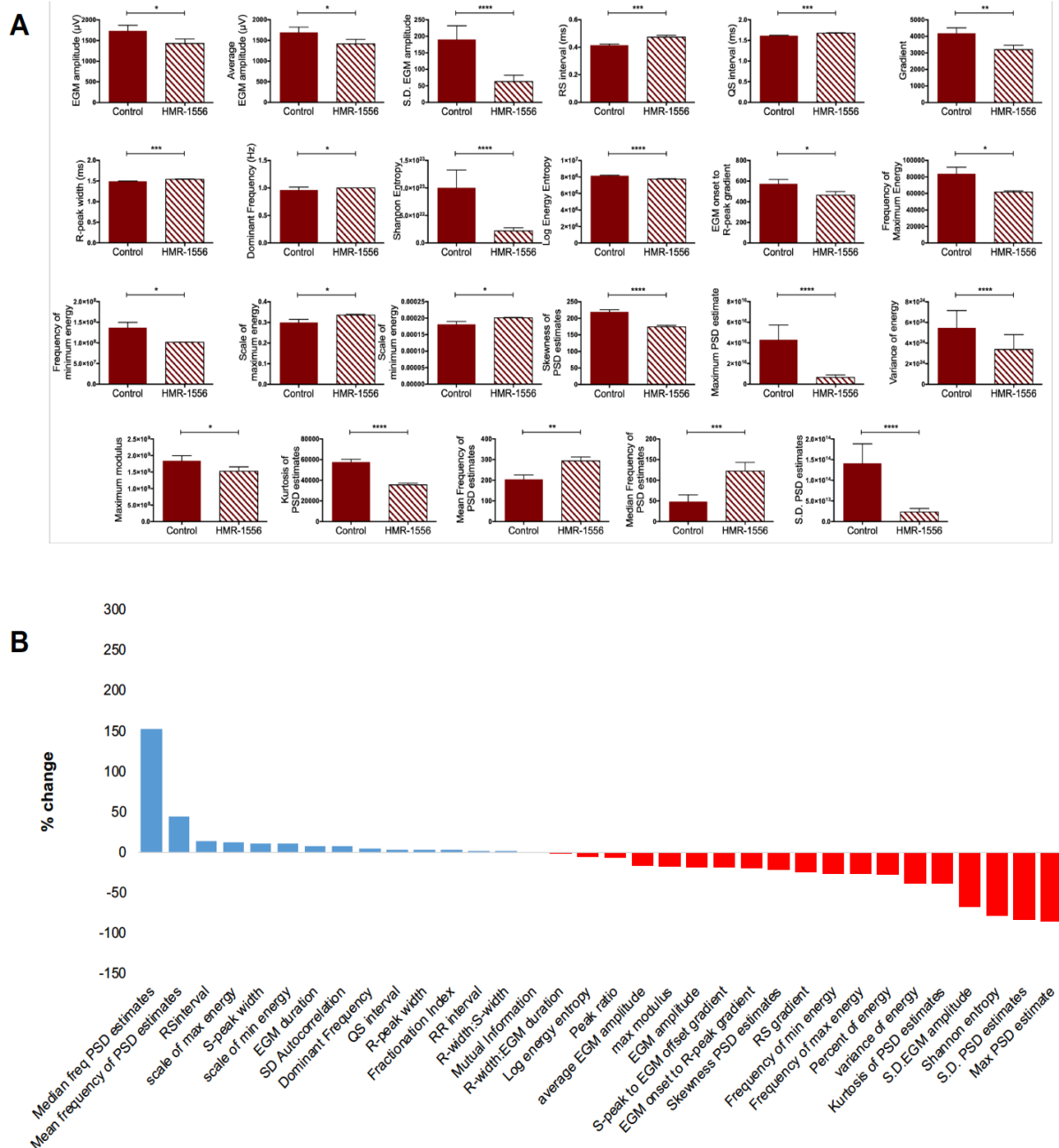


Figure A.17: Modifications on EGM morphology due to the I_{Ks} blockade when $1\mu\text{M}$ HMR-1556 was administered on NC cultures. (A) Features significantly changed due to HMR-1556 activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. (B) Quantification of EGM morphology changes following I_{Ks} blockade as compared to the average control EGM morphology. (Control: n = 54 EGMs; HMR-1556: n=74 EGMs).

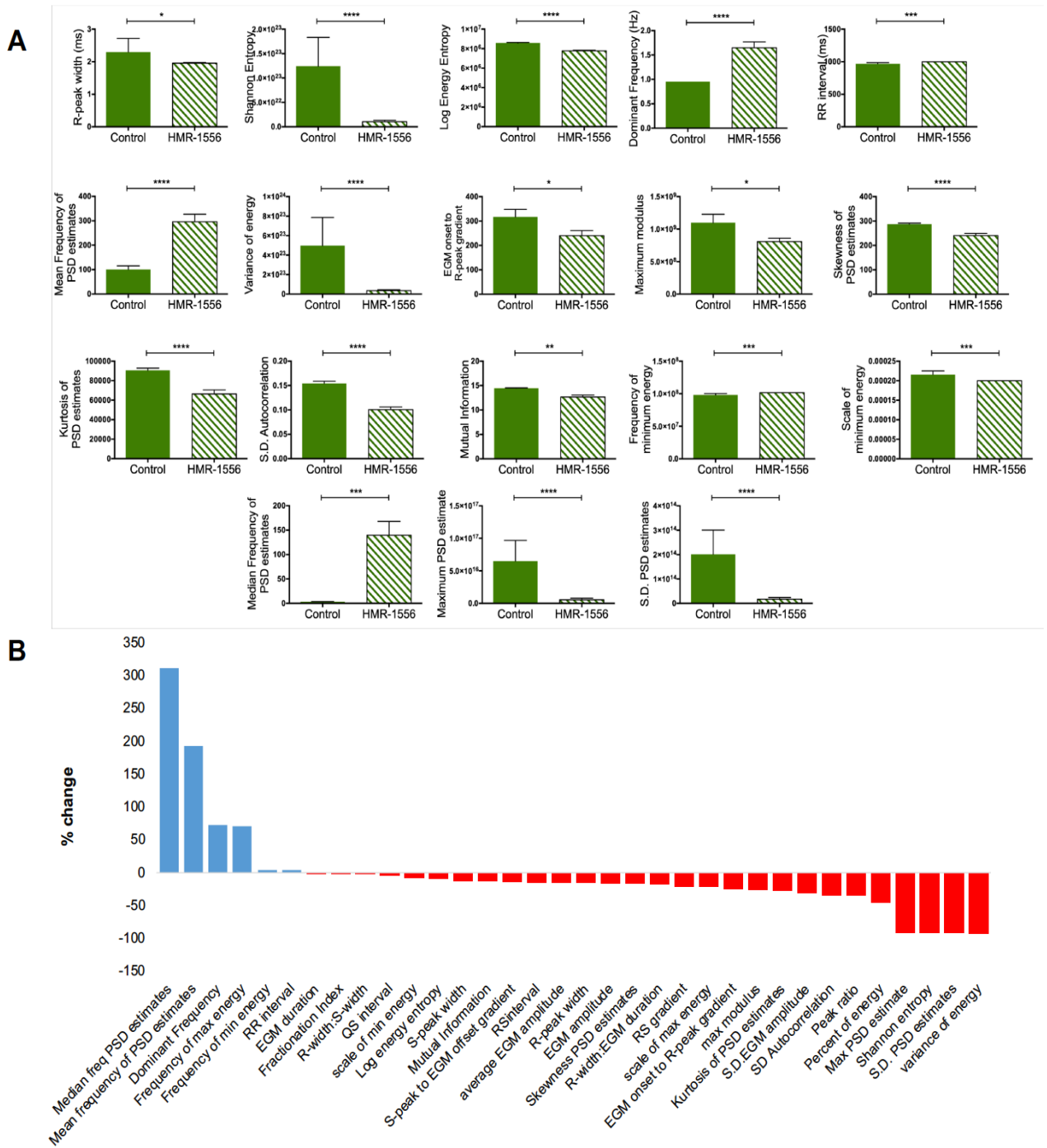


Figure A.18: EGM morphology modifications due to the I_{K_s} blockade when 300nM HMR-1556 was administered on NRVM/20% fibroblast co-cultures. (A) Features significantly changed due to HMR-1556 activity. Mann-Whitney test analysis. All bar charts represent mean±SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) Quantification of EGM morphology changes following I_{K_s} blockade as compared to the average control EGM morphology. (Control: $n = 27$ EGMs; HMR-1556: $n=23$ EGMs).

VI. EGM morphology modifications following K_{ATP} channel opening

The K_{ATP} channel opening using $3\mu\text{M}$ pinacidil resulted in significant modifications to 15 EGM features (Figure A.19A). All feature values were reduced following the enhancement of K^+ current, as opposed to the various types of K^+ current blockade where the some of the features increased. These features included the average EGM amplitude (control: $859.6\pm 76\mu\text{V}$; pinacidil: $657\pm 65\mu\text{V}$; $p=0.009$), dominant frequency (control: 1Hz ; pinacidil: $0.7\pm 0.06\text{Hz}$; $p<0.0001$) and Shannon entropy (control: $1.46e+023 \pm 2.9e+022$; pinacidil: $8.56e+022 \pm 2.3e+022$; $p<0.0001$).

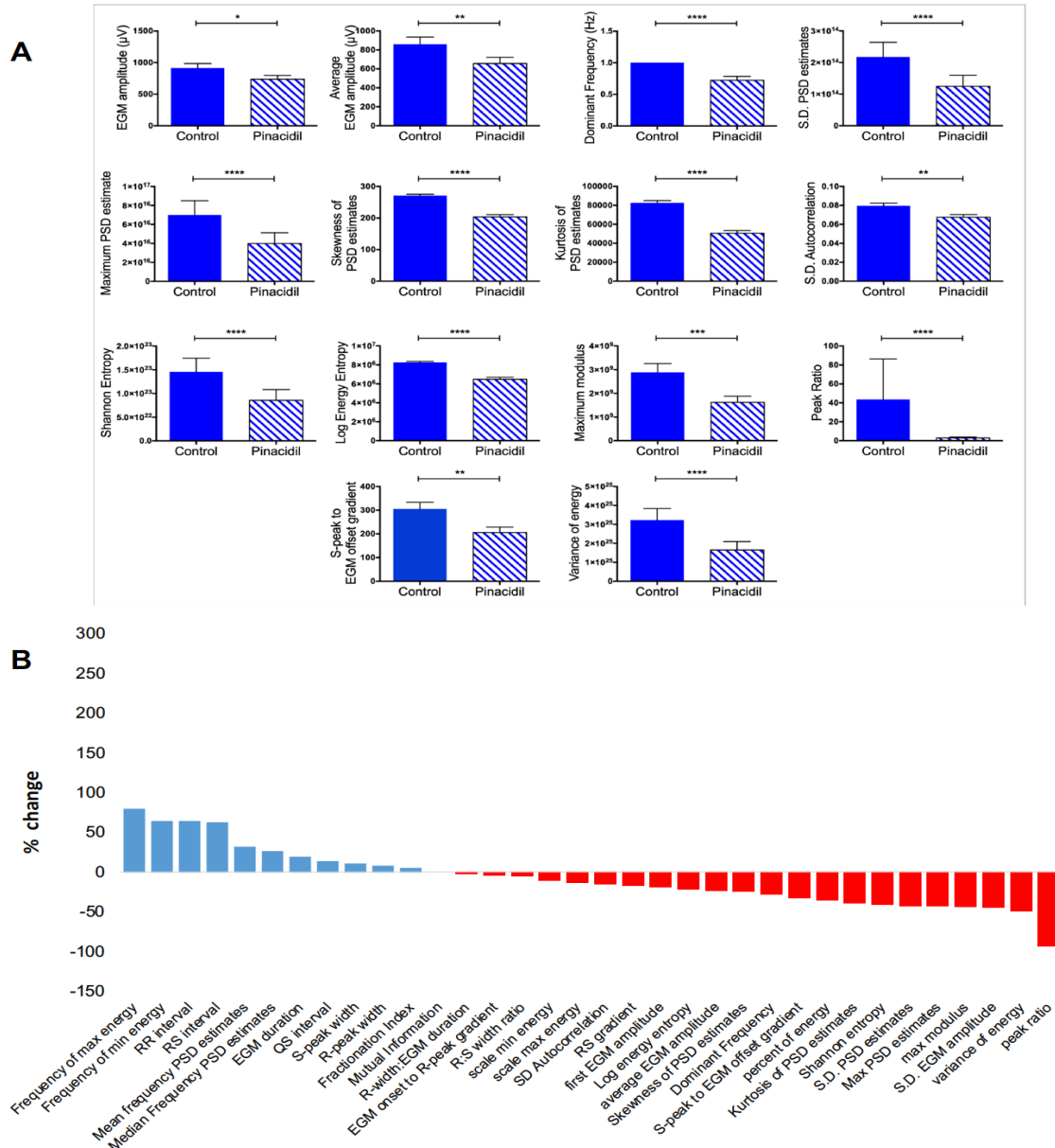


Figure A.19: EGM morphology modifications due to the K_{ATP} channel opening when $3\mu\text{M}$ pinacidil were administered on myocyte only monolayers. (A) Features significantly changed due to pinacidil activity. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$. (B) Quantification of EGM morphology changes following the enhancement of K^+ current as compared to the average control EGM morphology. (Control: $n = 46$ EGMs; Pinacidil: $n=73$ EGMs).

VII. EGM morphology manifestations of double ion channel blockade

The simultaneous Na⁺ and L-type Ca²⁺ channel blockade using 40.5μM lidocaine and 1.16μM nifedipine respectively affected 14 features (Figure A.20A). This group of changes excluded EGM duration and EGM amplitude which were influenced by Na⁺ channel blockade in isolation, but it included two non-linear dynamics features, i.e. Shannon entropy and log energy entropy, 5 time-frequency, the RR interval and 6 frequency-domain features based on power spectral density. The majority of features (n = 25) had very limited or no changes, as they were modified ±10% compared to control EGM morphology (Figure A.20B).

The concurrent Na⁺ and L-type Ca²⁺ channel blockade on NC monolayers using 740nM and 130nM nifedipine significantly affected 18 EGM features. These data were derived from three MEA dishes. Half of these features (n = 8) decreased and the rest increased. They included EGM amplitude and EGM duration, which were both reduced (average EGM amplitude - control: 2293±71μV; Ion channel blockade: 1650±75μV; p<0.0001 / EGM duration – control: 5.6±0.15ms; Ion channel blockade: 5.1±0.17ms; p=0.0005) (Figure A.21).

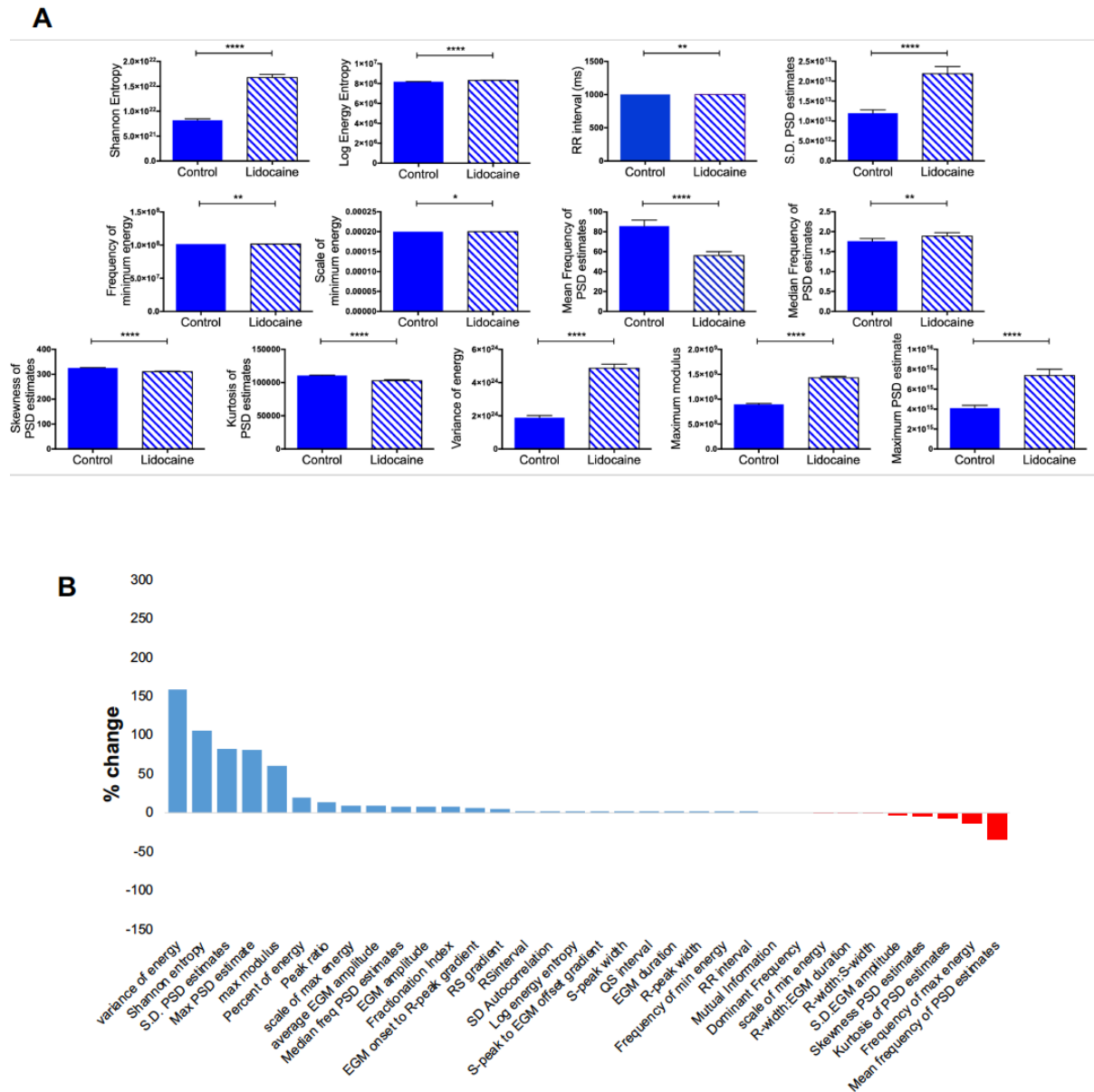


Figure A.20: Modifications on EGM morphology due to the double ion channel blockade when $40.5\mu\text{M}$ lidocaine and $1.16\mu\text{M}$ nifedipine were administered on myocyte only cultures. (A) Features significantly changed due to the simultaneous I_{Na} and I_{CaL} blockade. Mann-Whitney test analysis. All bar charts represent $\text{mean} \pm \text{SEM}$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. (B) Quantification of EGM morphology changes following ion channel blockade as compared to the average control EGM morphology. (Control: $n = 36$ EGMs; ion channel blockade = 30 EGMs).

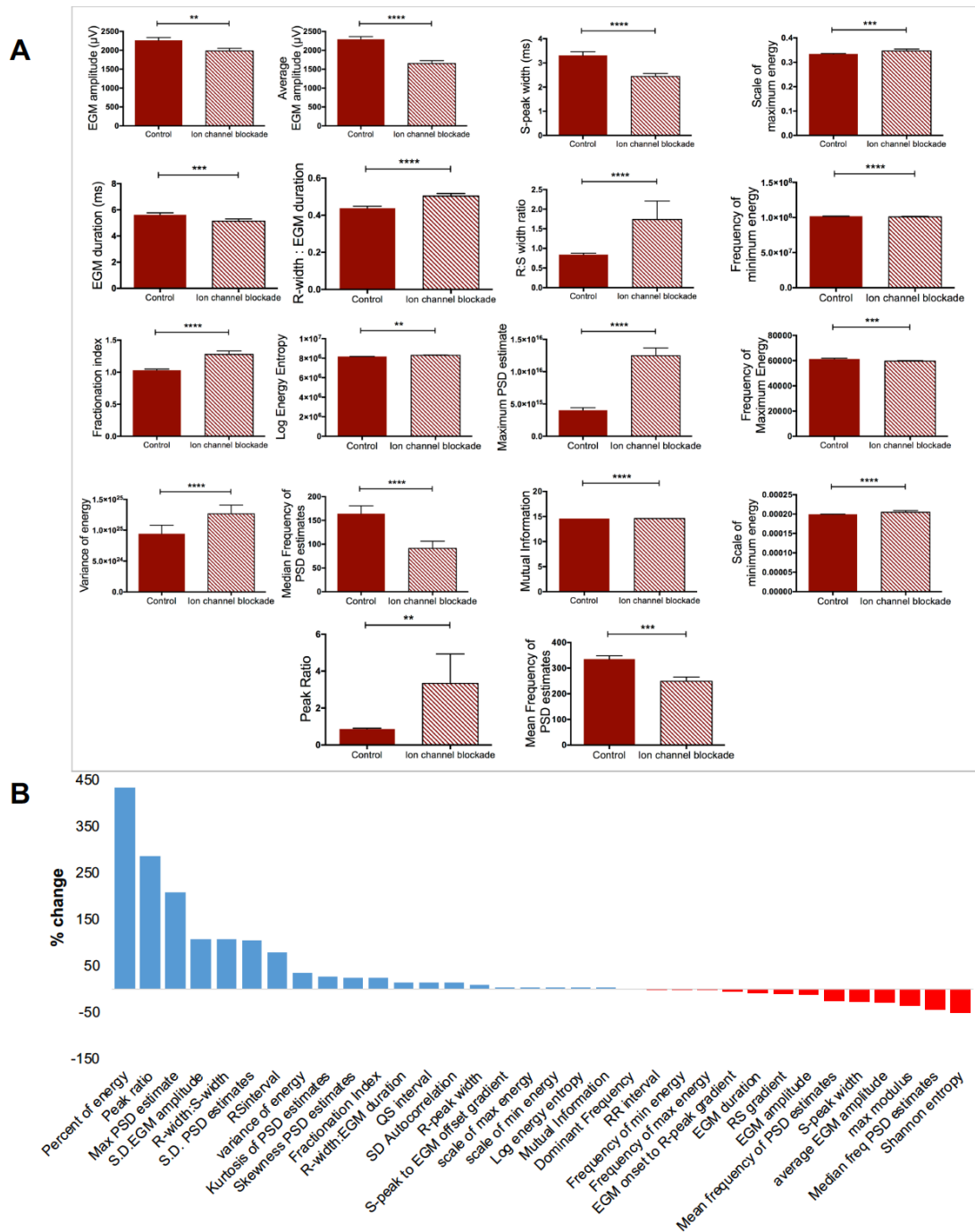


Figure A.21: Modifications on EGM morphology due to the double ion channel blockade when 740nM lidocaine and 130nM nifedipine were administered on NC cell cultures. (A) Features significantly changed due to the simultaneous I_{Na} and I_{CaL} blockade. Mann-Whitney test analysis. All bar charts represent mean \pm SEM; * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. (B) Quantification of EGM morphology changes following ion channel blockade as compared to the average control EGM morphology. (Control: n = 98 EGMs; ion channel blockade = 108 EGMs).

Appendix B

Feature subsets selected for classification training

- I. Feature subsets extracted by SFS and using the data obtained from the total of in vitro experimental conditions. SFS was optimised for each machine learning method by applying the corresponding method on feature selection process: LDA, kNN, SVM and Bagging Decision Trees (TreeBagger).

LDA	k-Nearest Neighbours	SVM	Bagging Decision Trees
Average EGM amplitude	Frequency of signal with maximum energy	RS amplitude	Frequency of signal with maximum energy
EGM duration	Frequency of signal with minimum energy	Dominant frequency	Logarithmic energy entropy
EGM amplitude	Maximum PSD estimate	Fractionation index	QS interval
Frequency of signal with maximum energy	Mutual information	EGM onset-to-Rpeak gradient	Mutual information
Frequency of signal with minimum energy	Logarithmic energy entropy	QT interval	Maximum modulus
Inter-peak signal gradient	R width	R-/S-peak width ratio	Variance of energy
Logarithmic energy entropy	RR interval	Scale of maximum energy	Frequency of signal with minimum energy
R/S amplitude ratio	S.D. PSD estimates	Logarithmic energy entropy	Median frequency of PSD estimates
Kurtosis of PSD estimates	Variance of energy	Maximum modulus	Kurtosis of PSD estimates
Maximum PSD estimate	Scale of maximum energy	Scale of maximum energy	RR interval
R width	Scale of minimum energy	Scale of minimum energy	Scale of minimum energy
Mean of PSD estimates	Shannon entropy		S.D. EGM amplitude
Scale of minimum energy			Maximum PSD estimate
Shannon entropy			
Median of PSD estimates			
Variance of energy			
Mutual information			
RR interval			
RS interval			
S.D. Autocorrelation			
Skewness of PSD estimates			
S.D. PSD estimates			

II. Feature subsets extracted by SFS for each training dataset corresponding to functional or structural modulations in isolation and the overall cell culture modifications investigated on this thesis. As a comparison, the feature subset used for defining the data obtained from the total of *in vitro* abnormalities is also presented. The TreeBagger method was implemented in SFS for all cases.

Functional modulations	Structural modulations	Total group of modulations
Kurtosis of PSD estimates	Dominant frequency	Frequency of signal with maximum energy
Maximum modulus	EGM duration	Logarithmic energy entropy
Median frequency of PSD estimates	EGM amplitude	QS interval
R width/EGM duration ratio	Fractionation index	Mutual information
Scale of minimum energy	Frequency of minimum energy	Maximum modulus
S.D. autocorrelation	RS gradient	Variance of energy
S-peak to EGM offset gradient	Mean frequency of PSD estimates	Frequency of signal with minimum energy
	EGM onset to R-peak gradient	Median frequency of PSD estimates
	R width	Kurtosis of PSD estimates
	Variance of energy	RR interval
		Scale of minimum energy
		S.D. EGM amplitude
		Maximum PSD estimate

III. Lists of feature subsets extracted by SFS and adjusted on each classification method using the 15% collagen coverage *ex vivo* data. The appropriate method was implemented for feature selection.

Bagging Decision Trees	SVM	Decision Trees	kNN	LDA
Mutual information	Frequency of minimum energy	EGM duration	Mutual information	S.V.D
RR interval	Mean frequency of PSD estimates	Frequency of maximum energy	RR interval	S.D. Autocorrelation
Logarithmic energy entropy	Median frequency of PSD estimates	Frequency of minimum energy	Scale at maximum signal energy	
Maximum modulus	EGM onset to Rpeak gradient	Maximum modulus		
EGM amplitude	Shannon entropy	Mutual information		
R-peak width	S.D. PSD estimates	QS interval		
	Speak to EGM offset gradient			
	S.D. EGM amplitude			

IV. Lists of feature subsets extracted by SFS for EGM data labelled as derived from tissue slices with collagen coverage below or above any of the percentages from the range 16-23%. The training dataset used for feature selection was the same for 16% and 17%, and it was the same for the 18% and 19% coverage. The features presented here are the ones selected following the application of a method that gives the optimum classification results, indicated in brackets.

16/17%	18/19%	20%	21%	22%	23%
<i>[Ensemble Decision Trees]</i>	<i>[Ensemble Decision Trees]</i>	<i>[Ensemble Decision Trees]</i>	<i>[SVM]</i>	<i>[SVM]</i>	<i>[SVM]</i>
EGM duration	Frequency of minimum energy	Dominant frequency	EGM amplitude	Dominant frequency	RS gradient
Kurtosis of PSD estimates	Logarithmic energy entropy	Logarithmic energy entropy	Frequency of minimum energy	Fractionation index	Maximum modulus
Fractionation index	S-width	QS interval	Kurtosis of PSD estimates	Mean frequency of PSD estimates	Mutual information
Frequency of minimum energy	Scale at maximum signal energy	RR interval	Logarithmic energy entropy	Percent of energy	Variance of energy
R-width/EGM duration ratio	Scale at minimum signal energy	Scale at maximum signal energy	Maximum modulus	S-width	
Scale at minimum signal energy	S.D. Autocorrelation	S.D. Autocorrelation	RR interval		
Speak to EGM offset gradient		S.D. EGM amplitude	Scale at minimum signal energy		

V. Feature subsets obtained by SFS using the training dataset of clinical EGMs. SFS was optimized for each classification method by applying the corresponding supervised machine learning method. The data were labelled as derived from scar or non-scar regions.

Bagging Decision Trees	SVM	Decision Trees	kNN	LDA
Frequency of maximum energy	Frequency of minimum energy	Dominant frequency	Dominant frequency	Dominant frequency
S-peak width	R/S-peak ratio	Frequency of maximum energy	Frequency of maximum energy	EGM duration
	R-peak width/EGM duration			Frequency of minimum energy
	S-peak width			Variance of energy
	Variance of energy			

Appendix C

Permissions

**THE AMERICAN PHYSIOLOGICAL SOCIETY LICENSE
TERMS AND CONDITIONS**

Mar 04, 2019

This Agreement between Imperial College London ("You") and The American Physiological Society ("The American Physiological Society") consists of your license details and the terms and conditions provided by The American Physiological Society and Copyright Clearance Center.

License Number	4541930430852
License date	Mar 04, 2019
Licensed Content Publisher	The American Physiological Society
Licensed Content Publication	Physiological Reviews
Licensed Content Title	Molecular Physiology of Cardiac Repolarization
Licensed Content Author	Jeanne M. Nerbonne, Robert S. Kass
Licensed Content Date	Oct 1, 2005
Licensed Content Volume	85
Licensed Content Issue	4
Type of Use	Thesis/Dissertation
Requestor type	author of original work
Format	print and electronic
Portion	figures/tables/images
Number of figures/tables/images	1
Will you be translating?	no
World Rights	English plus translations
Order reference number	
Title	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Institution name	n/a
Expected presentation date	Mar 2019
Portions	Figure 2
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Billing Type	Invoice
Billing Address	Imperial College London 4th floor, ICTEM building Hammersmith campus, Du Cane Road London, United Kingdom W12 0NN Attn: Imperial College London
Total	0.00 GBP

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Feb 04, 2019

This Agreement between Imperial College London ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4521990333215
License date	Feb 04, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Trends in Molecular Medicine
Licensed Content Title	The Living Scar – Cardiac Fibroblasts and the Injured Heart
Licensed Content Author	Eva A. Rog-Zielinska,Russell A. Norris,Peter Kohl,Roger Markwald
Licensed Content Date	Feb 1, 2016
Licensed Content Volume	22
Licensed Content Issue	2
Licensed Content Pages	16
Start Page	99
End Page	114
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Expected completion date	Mar 2019
Estimated size (number of pages)	300
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	GB 494 6272 12
Total	0.00 USD
Terms and Conditions	

**WOLTERS KLUWER HEALTH, INC. LICENSE
TERMS AND CONDITIONS**

Feb 05, 2019

This Agreement between Imperial College London ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	4522750977439
License date	Feb 05, 2019
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Circulation: Arrhythmia and Electrophysiology
Licensed Content Title	The Pathophysiologic Basis of Fractionated and Complex Electrograms and the Impact of Recording Techniques on Their Detection and Interpretation
Licensed Content Author	Jacques M.T. de Bakker, Fred H.M. Wittkamp
Licensed Content Date	Apr 1, 2010
Licensed Content Volume	3
Licensed Content Issue	2
Type of Use	Dissertation/Thesis
Requestor type	Individual
STM publisher name	
Portion	Figures/table/illustration
Number of figures/tables/illustrations	2
Figures/tables/illustrations used	Figure 9, Figure 10
Author of this Wolters Kluwer article	No
Title of your thesis / dissertation	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Expected completion date	Mar 2019
Estimated size(pages)	300
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	EU826013006
Total	0.00 GBP
Terms and Conditions	



RightsLink®

[Home](#)
[Account Info](#)
[Help](#)


Title: Remodelling of gap junctions and connexin expression in diseased myocardium

Author: Severs, Nicholas J.; Bruce, Alexandra F.

Publication: Cardiovascular Research

Publisher: Oxford University Press

Date: 2008-06-02

Copyright © 2008, Oxford University Press

Logged in as:
Konstantinos Nektarios Tzortzis

Account #:
3001401868

[LOGOUT](#)

Order Completed

Thank you for your order.

This Agreement between Imperial College London ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

[printable details](#)

License Number	4522760311517
License date	Feb 05, 2019
Licensed Content Publisher	Oxford University Press
Licensed Content Publication	Cardiovascular Research
Licensed Content Title	Remodelling of gap junctions and connexin expression in diseased myocardium
Licensed Content Author	Severs, Nicholas J.; Bruce, Alexandra F.
Licensed Content Date	Jun 2, 2008
Licensed Content Volume	80
Licensed Content Issue	1
Type of Use	Thesis/Dissertation
Requestor type	Author of this OUP content
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Title	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Institution name	Imperial College London
Expected presentation date	Mar 2019
Portions	Figure 1
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing address	Imperial College London 4th floor, ICTEM building Hammersmith campus, Du Cane Road London, United Kingdom W12 0NN Attn: Imperial College London
Total	0.00 GBP

**OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS**

Feb 08, 2019

This Agreement between Imperial College London ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

The publisher has provided special terms related to this request that can be found at the end of the Publisher's Terms and Conditions.

License Number	4524171315925
License date	Feb 08, 2019
Licensed content publisher	Oxford University Press
Licensed content publication	Cardiovascular Research
Licensed content title	The Lambeth Conventions: guidelines for the study of arrhythmias in ischaemia, infarction, and reperfusion
Licensed content author	WALKER, M J A; CURTIS, M J
Licensed content date	Jul 1, 1988
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Publisher of your work	Imperial College London
Expected publication date	Mar 2019
Permissions cost	0.00 GBP
Value added tax	0.00 GBP
Total	0.00 GBP
Title	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Institution name	Imperial College London
Expected presentation date	Mar 2019
Portions	Figure 3
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing Address	Imperial College London 4th floor, ICTEM building Hammersmith campus, Du Cane Road London, United Kingdom W12 0NN Attn: Imperial College London
Total	0.00 GBP
Terms and Conditions	



RightsLink®

[Home](#)
[Account Info](#)
[Help](#)


Title: A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate

Author: Koonlawee Nademanee, John McKenzie, Erol Kosar, Mark Schwab, Buncha Sunsaneewitayakul, Thaveekiat Vasavakul, Chotikorn Khunnawat, Tachapong Ngarmukos

Publication: Journal of the American College of Cardiology

Publisher: Elsevier

Date: 2 June 2004

Logged in as:
Konstantinos Nektarios Tzortzis
Account #:
3001401868

[LOGOUT](#)

Copyright © 2004 American College of Cardiology Foundation. Published by Elsevier Inc. All rights reserved.

Order Completed

Thank you for your order.

This Agreement between Imperial College London ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

[printable details](#)

License Number	4522761289745
License date	Feb 05, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of the American College of Cardiology
Licensed Content Title	A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate
Licensed Content Author	Koonlawee Nademanee, John McKenzie, Erol Kosar, Mark Schwab, Buncha Sunsaneewitayakul, Thaveekiat Vasavakul, Chotikorn Khunnawat, Tachapong Ngarmukos
Licensed Content Date	Jun 2, 2004
Licensed Content Volume	43
Licensed Content Issue	11
Licensed Content Pages	10
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1A
Title of your thesis/dissertation	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Publisher of new work	Imperial College London
Expected completion date	Mar 2019
Estimated size (number of pages)	300
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	GB 494 6272 12
Total	0.00 GBP



RightsLink®

- [Home](#)
- [Create Account](#)
- [Help](#)
-



Title: Catheter ablation of atrial fibrillation guided by electrogram fractionation and dominant frequency analysis

Author: Demosthenes G Katritsis, , Ioannis Pantos, et al

Publication: Expert Review of Cardiovascular Therapy

Publisher: Taylor & Francis

Date: May 1, 2011

Rights managed by Taylor & Francis

LOGIN

If you're a [copyright.com user](#), you can login to RightsLink using your copyright.com credentials. Already a [RightsLink user](#) or want to [learn more?](#)

Thesis/Dissertation Reuse Request

Taylor & Francis is pleased to offer reuses of its content for a thesis or dissertation free of charge contingent on resubmission of permission request if work is published.

- [BACK](#)
- [CLOSE WINDOW](#)

Copyright © 2019 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at customercare@copyright.com

**WOLTERS KLUWER HEALTH, INC. LICENSE
TERMS AND CONDITIONS**

Feb 19, 2019

This Agreement between Imperial College London ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	4532580462657
License date	Feb 19, 2019
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Circulation Research
Licensed Content Title	Subthreshold stimulation of Purkinje fibers interrupts ventricular tachycardia in intact hearts. Experimental study with voltage-sensitive dyes and imaging techniques.
Licensed Content Author	G Salama, A Kanai, I R Efimov
Licensed Content Date	Apr 1, 1994
Licensed Content Volume	74
Licensed Content Issue	4
Type of Use	Dissertation/Thesis
Requestor type	Individual
STM publisher name	
Portion	Figures/table/illustration
Number of figures/tables/illustrations	1
Figures/tables/illustrations used	1
Author of this Wolters Kluwer article	No
Title of your thesis / dissertation	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Expected completion date	Mar 2019
Estimated size(pages)	300
Requestor Location	Konstantinos Nektarios Tzortzis Apartment 30, Nautilus House 14 West Row London, Greater London W10 5QL United Kingdom Attn: Konstantinos Nektarios Tzortzis
Publisher Tax ID	EU826013006
Billing Type	Invoice
Billing Address	Imperial College London 4th floor, ICTEM building Hammersmith campus, Du Cane Road London, United Kingdom W12 0NN Attn: Imperial College London
Total	0.00 GBP

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Jul 03, 2019

This Agreement between Imperial College London ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4621531171773
License date	Jul 03, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Pharmacological and Toxicological Methods
Licensed Content Title	Validation of a guinea pig Langendorff heart model for assessing potential cardiovascular liability of drug candidates
Licensed Content Author	Liang Guo,Zedong Dong,Heather Guthrie
Licensed Content Date	September–October 2009
Licensed Content Volume	60
Licensed Content Issue	2
Licensed Content Pages	22
Start Page	130
End Page	151
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Fig.1A
Title of your thesis/dissertation	Predicting cardiac functional and structural abnormalities from electrogram morphology using supervised machine learning
Publisher of new work	Imperial College London
Expected completion date	Mar 2019
Estimated size (number of pages)	300
Requestor Location	Imperial College London 4th floor, ICTEM building Hammersmith campus, Du Cane Road London, Greater London W12 0NN United Kingdom Attn: Imperial College London
Publisher Tax ID	GB 494 6272 12
Total	0.00 GBP